Compendium of Plant Genomes
Series Editor: Chittaranjan Kole

Satoshi Tabata
Jens Stougaard
Editors

The *Lotus japonicus* Genome

Springer

Compendium of Plant Genomes

Series editor

Chittaranjan Kole
Mohanpur, Nadia, West Bengal
India

For further volumes:
http://www.springer.com/series/11805

Satoshi Tabata · Jens Stougaard
Editors

The *Lotus japonicus* Genome

Springer

Editors
Satoshi Tabata
Kazusa DNA Research Institute
Kisarazu, Chiba
Japan

Jens Stougaard
Department of Molecular Biology
and Genetics
University of Aarhus
Aarhus
Denmark

ISSN 2199-4781 ISSN 2199-479X (electronic)
ISBN 978-3-662-44269-2 ISBN 978-3-662-44270-8 (eBook)
DOI 10.1007/978-3-662-44270-8
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014949341

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

This book series is dedicated to
my wife Phullara, and our children Sourav, Carena, and Devleena

Chittaranjan Kole

Preface to the Series

Genome sequencing has emerged as the leading discipline in the plant sciences coinciding with the start of the new century. For much of the twentieth century, plant geneticists were only successful in delineating putative chromosomal location, function and changes in genes indirectly through the use of a number of 'markers' physically linked to them. These included visible or morphological, cytological, protein, and molecular or DNA markers. Among them, the first DNA marker, the RFLPs, introduced a revolutionary change in plant genetics and breeding in the mid-1980s, mainly because of their infinite number and thus potential to cover maximum chromosomal regions, phenotypic neutrality, absence of epistasis, and codominant nature. An array of other hybridization-based markers PCR-based markers, and markers based on both facilitated construction of genetic linkage maps, mapping of genes controlling simply inherited traits and even gene clusters (QTLs) controlling polygenic traits in a large number of model and crop plants. During this period a number of new mapping populations beyond F_2 were utilized and a number of computer programs were developed for map construction, mapping of genes, and for mapping of polygenic clusters or QTLs. Molecular markers were also used in studies of evolution and phylogenetic relationship, genetic diversity, DNA-fingerprinting and map-based cloning. Markers tightly linked to the genes were used in crop improvement employing the so-called marker-assisted selection. These strategies of molecular genetic mapping and molecular breeding made a spectacular impact during the last one and a half decades of the twentieth century. But still they remained 'indirect' approaches for elucidation and utilization of plant genomes since much of the chromosomes remained unknown and the complete chemical depiction of them was yet to be unraveled.

Physical mapping of genomes was the obvious consequence that facilitated development of the 'genomic resources' including BAC and YAC libraries to develop physical maps in some plant genomes. Subsequently, integrated genetic-physical maps were also developed in many plants. This led to the concept of structural genomics. Later on emphasis was laid on EST and transcriptome analysis to decipher the function of the active gene sequences leading to another concept defined as functional genomics. The advent of techniques of bacteriophage gene and DNA sequencing in the 1970s was extended to facilitate sequencing of these genomic resources in the last decade of the twentieth century.

As expected, sequencing of chromosomal regions would have led to too much data to store, characterize and utilize with the-then available computer software could handle. But development of information technology made the life of biologists easier by leading to a swift and sweet marriage of biology and informatics and a new subject was born—bioinformatics.

Thus, evolution of the concepts, strategies and tools of sequencing and bioinformatics reinforced the subject of genomics—structural and functional. Today, genome sequencing has traveled much beyond biology and involves biophysics, biochemistry and bioinformatics!

Thanks to the efforts of both public and private agencies, genome sequencing strategies are evolving very fast, leading to cheaper, quicker and automated techniques right from clone-by-clone and whole-genome shotgun approaches to a succession of second generation sequencing methods. Development of software of different generations facilitated this genome sequencing. At the same time newer concepts and strategies were emerging to handle sequencing of the complex genomes, particularly the polyploids.

It became a reality to chemically—and so directly—define plant genomes, popularly called whole-genome sequencing or simply genome sequencing.

The history of plant genome sequencing will always cite the sequencing of the genome of the model plant *Arabidopsis thaliana* in 2000 that was followed by sequencing the genome of the crop and model plant rice in 2002. Since then, the number of sequenced genomes of higher plants has been increasing exponentially, mainly due to the development of cheaper and quicker genomic techniques and, most importantly, development of collaborative platforms such as national and international consortia involving partners from public and/or private agencies.

As I write this preface for the first volume of the new series "Compendium of Plant Genomes", a net search tells me that complete or nearly-complete whole-genome sequencing of 45 crop plants, eight crop and model plants, eight model plants, 15 crop progenitors and relatives, and three basal plants are accomplished, the majority of which are in the public domain. This means that we nowadays know many of our model and crop plants chemically, i.e. directly, and we may depict them and utilize them precisely better than ever. Genome sequencing has covered all groups of crop plants. Hence, information on the precise depiction of plant genomes and the scope of their utilization is growing rapidly every day. However, the information is scattered in research articles and review papers in journals and dedicated web pages of the consortia and databases. There is no compilation of plant genomes and the opportunity of using the information in sequence-assisted breeding or further genomic studies. This is the underlying rationale for starting this book series, with each volume dedicated to a particular plant.

Plant genome science has emerged as an important subject in academia, and the present compendium of plant genomes will be highly useful both to students and teaching faculties. Most importantly, research scientists involved in genomics research will have access to systematic deliberations on the plant genomes of their interest. Elucidation of plant genomes is not only of interest for the geneticists and breeders, but also for practitioners of an array of plant science disciplines, such as taxonomy, evolution, cytology,

physiology, pathology, entomology, nematology, crop production, biochemistry, and obviously bioinformatics. It must be mentioned that information regarding each plant genome is ever-growing. The contents of the volumes of this compendium are therefore focusing on the basic aspects of the genomes and their utility. They include information on the academic and/or economic importance of the plants, description of their genomes from a molecular genetic and cytogenetic point of view, and the genomic resources developed. Detailed deliberations focus on the background history of the national and international genome initiatives, public and private partners involved, strategies and genomic resources and tools utilized, enumeration on the sequences and their assembly, repetitive sequences, gene annotation and genome duplication. In addition, synteny with other sequences, comparison of gene families and, most importantly, potential of the genome sequence information for gene pool characterization through genotyping by sequencing (GBS) and genetic improvement of crop plants have been described. As expected, there is a lot of variation of these topics in the volumes based on the information available on the crop, model or reference plants.

I must confess that as the series editor it has been a daunting task for me to work on such a huge and broad knowledge base that spans so many diverse plant species. However, pioneering scientists with life-time experience and expertise on the particular crops did excellent jobs editing the respective volumes. I myself have been a small science worker on plant genomes since the mid-1980s and that provided me the opportunity to personally know several stalwarts of plant genomics from all over the globe. Most, if not all, of the volume editors are my long-time friends and colleagues. It has been highly comfortable and enriching for me to work with them on this book series. To be honest, while working on this series I have been and will remain a student first, a science worker second and a series editor last. And I must express my gratitude to the volume editors and the chapter authors for providing me the opportunity to work with them on this compendium.

I also wish to mention here my thanks and gratitude to the Springer staff, Dr. Christina Eckey and Dr. Jutta Lindenborn in particular, for all their constant and cordial support right from the inception of the idea.

I always had to set aside additional hours to edit books besides my professional and personal commitments—hours I could and should have given to my wife, Phullara, and our kids, Sourav, Carena, and Devleena. I must mention that they not only allowed me the freedom to take away those hours from them but also offered their support in the editing job itself. I am really not sure whether my dedication of this compendium to them will suffice to do justice to their sacrifices for the interest of science and the science community.

Chittaranjan Kole

Preface to the Volume

Progress in plant genomics and genetics has been rapid and sustained in recent years. Focused research efforts on model plants have spearheaded this development and laid the foundation for subsequent investigations in the major crop species that are often less amenable. In the legume family (Fabaceae) *Lotus japonicus* (birdsfoot trefoil) was adopted as a model species more than 20 years ago and a considerable body of knowledge has since been built using genomic and genetic analyses in this species. Without being exhaustive, this volume presents some of the achievements made and provides a timely overview of topics relevant for future developments using legume genomics to improve our understanding of legume biology.

With more than 18,000 species represented, Fabaceae comprises the third largest family among the flowering plants and only grasses are more important in agriculture. Legumes are very diverse, ranging from tropical trees to temperate herbs. In addition to food and feed, they provide products from secondary metabolites and protein to oil and timber. The symbiosis with nitrogen fixing bacteria, rhizobia, enables legumes to obtain reduced dinitrogen for their own growth and is a major source of nitrogen in ecosystems and crop rotations. Like many other plants, legumes can also form symbiotic association with mycorrhizal fungi, which are important for phosphate uptake, and recent studies have identified a common symbiosis pathway for mycorrhizal and rhizobial symbiosis. Encompassing these biological and agricultural features, central topics in endosymbiosis, development, hormone regulation, carbon/nitrogen, and secondary metabolism, together with progress in high throughput genomic and genetic approaches, will be covered in this volume on the *Lotus japonicus* model system.

The world population is rapidly growing and an increase in food production is needed to match this increased food demand. Given the importance of legumes in sustainable agriculture, mining the model legume genomes and translation of knowledge from model legumes to crop legumes is important for our future. This volume provides an overview of some of the pertinent topics. We thank all the authors for their excellent contributions to this volume and hope that the expert's overview they have provided will serve as inspiration and encouragement for the future.

Satoshi Tabata
Jens Stougaard

Contents

Part I
The Importance of Lotus as a Model and a Crop

Background and History of the *Lotus japonicus* Model Legume System

1

Jens Stougaard

Abstract

The combination of favourable biological features, stable transformation procedures, application of genetics and genome-based global approaches has established *Lotus japonicus* as a model legume and provided a platform for addressing important biological questions often, but not exclusively, focusing on endosymbiosis. Several important discoveries have been made, and the *Lotus* community has contributed novel results, promoting our understanding of plant biology as well as our understanding of properties and characteristics typical for plants belonging to the legume family. Progress has been fast since *L. japonicus* was first promoted as a model plant yet there are many challenges for the coming years. This introductory chapter will set the stage for some of these challenges, while possibilities and challenges emerging from specific research projects will be addressed in the chapters that follow.

1.1 The *Lotus japonicus* Model Legume System

Mendel worked with garden peas for his groundbreaking work that established genetics as a science (Reid and Ross 2011). For many years, pea plants were also the workhorse in classical plant physiology. The ethylene-induced triple response of pea seedlings was, for example, one of the key observations leading to the identification of ethylene as a plant hormone. Continuing the genetic approaches, large collections of pea mutants and morphological variants were isolated, and substantial effort was invested in their phenotypic characterisation. Included in this collection was a sizeable subset of symbiotic plant mutants, with phenotypes ranging from non-nodulation to hypernodulation (Borisov et al. 2007; Tsyganov et al. 2002). From a historical perspective, the need for a model legume may therefore not have been obvious when the quest for a model legume started. However, prospects for combining genetics with stable transformation and emerging methodologies for genome-based studies inspired a search for a legume better suited to these global approaches. One of the outcomes was the proposal of *Lotus japonicus* as a model legume in 1992 (Handberg

J. Stougaard (✉)
Centre for Carbohydrate Recognition and Signalling, Department of Molecular Biology and Genetics, Aarhus University, Gustav Wieds Vej 10, 8000 Aarhus C, Denmark
e-mail: stougaard@mb.au.dk

S. Tabata and J. Stougaard (eds.), *The Lotus japonicus Genome*, Compendium of Plant Genomes,
DOI 10.1007/978-3-662-44270-8_1,

and Stougaard 1992). Without aiming to be exhaustive, this volume highlights some of the achievements reached within the 20 years that followed and sketches the possibilities lying ahead.

Early botanical work on morphological features of the *Loteae* tribe in the 1950s led to the proposal of *L. japonicus* as a separate species (Larsen 1955). Further biological studies and karyotyping of chromosomes found *L. japonicus* to be self-fertile and diploid with a chromosome number of $n = 6$ (Cheng and Grant 1973). Subsequently, fluorescent measurements of 1C values for DNA content in nuclei of individual root cells indicated a genome size among the lowest in the legume family (Bennett and Smith 1976). These features distinguished *L. japonicus* from the morphologically very similar tetraploid outbreeder *Lotus corniculatus* ($n = 12$) that had previously been used for investigating regulation and promoter function of nodulin genes in transgenic roots and transgenic plants (Stougaard Jensen et al. 1986; Stougaard et al. 1990). Fortunately, some of the tissue culture and transformation techniques developed in *Lotus corniculatus* could be refined and transferred to *L. japonicus* (Stougaard et al. 1987; Hansen et al. 1989). A list of these model plant features was published previously (Handberg and Stougaard 1992).

Model features are to some extent technology and time dependent; however, it appears that *L. japonicus* has passed the test of time. An updated version of this list of "raison d'être" is shown in Table 1.1. Almost all of the features in the list have in one way or another been exploited in experimental procedures addressing important biological questions often, but not exclusively, focusing on endosymbiosis. Several different transformation procedures for regeneration of transgenic and composite plants have been established and used experimentally (Handberg and Stougaard 1992; Hansen et al. 1989). The number of selectable markers that can be used has been expanded, and both positive and negative selection schemes have been developed on this basis (Lohar et al. 2001; Lombari et al. 2003; Stougaard 1993). RNAi technologies have been used successfully (Kumagai et al. 2006; Soyano et al. 2013), and stable lines, such as pNin-GUS that inducibly express promoter reporter fusion for use as symbiotic response markers, have been made available (Radutoiu et al. 2003). Exploiting the favourable culture characteristics of *L. japonicus*, grafting procedures for root–shoot grafts and Y grafts have been used for investigating systemic plant responses mainly in the context of autoregulation of nodulation (Magori et al. 2009; Takahara et al. 2013). The small size of *L. japonicus* plantlets allowed for the development of in vitro mycorrhization in petri dishes using a filter sandwich set-up (Novero et al. 2002). Taking a whole plant approach, the vegetative growth pattern has been described and the role of strigolactone investigated. In contrast to Arabidopsis, *L. japonicus* develops multiple axillary shoots, and the ontogeny of these cotyledonary shoot meristems has been characterised and the influence of strigolactone on shoot architecture described (Alvarez et al. 2006; Lui et al. 2013). The reproductive life phase has also been studied, and analysis of the genetic background for the development of asymmetric flowers is ongoing (Xu et al. 2013). Another line of investigation has taken advantage of easy access to seeds in the simple straight seedpods of *L. japonicus* to follow seed development and the seed proteome from early-stage green seeds to mature dry seeds (Dam et al. 2009; Credali et al. 2013).

Forward genetic approaches based on mutant populations and gene discovery starting from interesting phenotypes have been a core activity for the *L. japonicus* community (Kouchi et al. 2010; Kistner et al. 2005; Sandal et al. 2006). Several breakthroughs have been achieved, and combined with the parallel efforts in *Medicago truncatula,* this has, in a relatively short time span, revealed the molecular backbone of both rhizobial and mycorrhizal endosymbioses. Key components of the legume signal perception/transduction genetic network mediating the rhizobial and endomycorrrhizal interactions have been defined and the functional aspects of symbiosis opened for analysis (Madsen et al. 2010; Desbrosses and Stougaard 2011; Oldroyd 2013).

Table 1.1 Features and characteristics of *Lotus japonicus*

Growth characteristics	Small primary plant Auxiliary shoots, bushy plant architecture Perennial 7-week period from seed to flowering Generation time from seed to seed, 3–4 months Small seeds: ~1.2 g per 1,000 seeds Fast regrowth from stem base/tap root Fast plant multiplication from nodal sections Root/shoot grafting and Y grafts possible
Propagation	Continuous flowering Large flowers allow for controlled crossings Self-fertile Simple spikeless and straight seedpod—like soybean and pea Approximately 20 seeds per pod Ample seed production, up to 6,000 seeds per plant Relative humidity above 65 % prevents seed shattering Hand pollination possible
Genome characteristics	Diploid, $2n = 12$ Genome size of ~478 Mb Cytogenetics developed Genespace fully sequenced Gene models based on mRNA and small RNAseq Genome re-sequenced in different ecotypes and diploid Lotus species High-resolution genetic maps available Recombinant inbred populations available Large collection of ecotypes available Diploid Lotus species for interspecific crosses available
Tissue culture	Regeneration from callus Stable transformation with *Agrobacterium tumefaciens* Positive selection: Hygromycin, kanamycin, geneticin and Basta Negative selection: 5-fluorocytosine Composite plants with *Agrobacterium rhizogenes*
Nodulation	Primary symbiont: *Mesorhizobium loti* Alternative often less-efficient symbionts: *Azorhizobium coulinodans*, *Sinorhizobium fredii*, IRGB74, NGR234 Several symbiont genomes sequenced Determinate nodules Sequential nodule development Primarily invasion mode via infection threads Crack entry and intercellular invasion observed in absence of infection threads
Mycorrhiza	Mycorrhized by *Rhizophagus irregularis* and *Gigantia margarita* and more
Pathogens	Leaf rust, *Uromyces loti* Clover rot, *Sclerotinia trifoliorum* Root-knot nematodes, *Meloidogyne incognita*
Parasites	Weed parasites, compatible and incompatible *Striga* and *Orobanche* spp
Insect interactions	Burnet moth, *Zygaena* filipendulae

Interestingly, the pea mutant collection has frequently been drawn into this work upon identification of causative genes using the model legume discovery tools (Madsen et al. 2003: Zhukov et al. 2008; Borisov et al. 2003). Taking a broader view of plant interactions, *L. japonicus* has been used for studies of nematode invasion (Poch et al. 2007; Weerasinghe et al. 2005), emerging investigations of root colonisation by parasitic weeds like Striga spp (Hiraoka et al. 2009) and specialised insect interactions (Zagrobelsky et al. 2007).

Aiming at continuing this success, development of additional genetic resources has remained a focal point. To enable reverse genetics, a TILLING population was established from EMS mutagenised *L. japonicus* seeds and made available to the plant community (Perry et al. 2009). Later, an endogenous retrotransposon called *Lotus retrotransposon 1* (LORE1) enabled the organisation of an insertion mutant population for reverse genetics. LORE1 has several unique characteristics making it particularly suitable for this (Urbanski et al. 2012; Fukai et al. 2012). The element was initially found to be activated by tissue culture; however, it is only transposed in the pollen line. In regenerated plants, these features give rise to seeds with independent patterns of insertions (Fukai et al. 2010). This has paved the way for identification of insertions in genes of interest by a simple sequence search, and together with the annotated genome sequence available, this resource was a quantum leap in legume research and a resource matching the best among model plant systems. The already established studies of primary and secondary metabolism that can be difficult to approach using forward genetics are likely to benefit from this resource (Vriet et al. 2010; Clemente et al. 2012; Perez-Delgado et al. 2013).

1.2 Challenges Ahead

Much has been accomplished, yet there are many challenges for the coming years. The *L. japonicus* genespace has been sequenced and re-sequenced in different ecotypes and *Lotus* species to uncover the biodiversity, and a well-annotated genome has been established as a basis for comparative genome analysis within the genus and the legume family. So far, around 30 *L. japonicus* ecotypes and related species such as *Lotus burttii* have been re-sequenced providing single nucleotide polymorphisms and thereby setting the stage for genome-wide association studies accessing natural variation and biodiversity (Kai et al. 2010; Andersen and Sato, pers.com). Epigenetic regulation is another level of control that can now be addressed on a comparative basis. Further improvements in the annotation are likely to come from participatory genome annotation, and this will be useful for functional analysis in the more complex genomes of crop legumes.

Reverse genetic resources are available and the gene coverage is high. However, inactivation of small genes that by nature have a limited target size could still be improved. Likewise, genetic linkage is also an obstacle for functional analysis of individual members of gene families. Redundancy may shield the effects of inactivation, and because of the linkages, double mutants may be difficult or impossible to obtain by crossing. Gene-specific inactivation procedures based on transcription activator-like effector nucleases (TALEN), Zinc finger nucleases (ZFN) or clustered regulatory interspaced short palindrome repeat-based technologies (CRISPR) could nicely supplement TILLING and LORE1 mutants for studies of small genes and gene families. Studies of miRNAs and other small RNAs that do not lend themselves easily to molecular genetic studies may particularly benefit from these technologies (De Luis et al. 2012). Biochemistry and physiology are the brothers of genetics, and it is now time to bring biochemical and physiological analysis back to centre stage. Molecular genetics is a powerful tool for the identification of central components in processes of interest. However, other approaches are needed for detailed understanding of cellular processes and pathways. It is thus important to advance approaches integrating genetic, biochemical and physiological analyses. Finally, the *L. japonicus* model system with all the resources available and the knowledge generated from analysis of fungal and bacterial endosymbiosis should be in a prime position to contribute to a better understanding of plant–endophyte interactions as well as interactions with microbial populations in the rhizosphere.

Acknowledgements Supported by the Danish National Research Foundation grant no. DNRF79 and the ERC Advanced Grant 268523.

References

Alvarez NDG, Meeking RJ, White DRW (2006) The origin, initiation and development of axillary shoot meristems in *Lotus japonicus*. Ann Bot 98:953–963

Bennett MD, Smith JB (1976) Nuclear DNA amounts in Angiosperms. Phil Trans Royal Soc B 274:224–274

Borisov AY, Danilova TN, Koroleva TA et al (2007) Regulatory genes of garden pea (*Pisum sativum* L.) controlling the development of nitrogen-fixing nodules and arbuscular mycorrhiza: a review of basic and applied aspects. Appl Biochem Microbiol 43 (3):237–243

Borisov AY, Madsen LH, Tsyganov VE et al (2003) The Sym35 gene required for root nodule development in pea is an ortholog of *Nin* from *Lotus japonicus*. Plant Phys 131:1009–1017

Cheng RI-J, Grant WF (1973) Species relationships in the *Lotus corniculatus* group as determined by karyotype and cytophotometric analysis. Can J Genet Cytol 15:101–115

Clemente MR, Bustos-Sanmamed P, Loscos J et al (2012) Thiol synthetases of legumes: immunogold localization and differential regulation by phytohormones. J Exp Botany 63(10):3923–3934

Credali A, Garcia-Calderon M, Dam S et al (2013) The K +-dependent asparaginase, NSE1, is crucial for plant growth and seed production in *Lotus japonicus*. Plant Cell Physiol 54(1):107–118

De Luis A, Markmann K, Cognat V et al (2012) Two MicroRNAs linked to nodule infection and nitrogen-fixing ability in the legume *Lotus japonicus*. Plant Phys 160:2137–2154

Dam S, Laursen BS, Ørnfelt JH et al (2009) The proteome of seed development in the model legume *Lotus japonicus*. Plant Phys 149:1325–1340

Desbrosses GJ, Stougaard J (2011) Root nodulation: a paradigm for how plant-microbe symbiosis influences host developmental pathways. Cell Host Microbe 10:348–358

Fukai E, Umehara Y, Sato S et al (2010) Derepression of the plant chromovirus LORE1 induces germline transposition in regenerated plants. PloS Genet 6(3): e1000868

Fukai E, Soyano T, Umehara U et al (2012) Establishment of a *Lotus japonicus* gene tagging population using the exon-targeting endogenous retrotransposon LORE1. Plant J 69:720–730

Handberg K, Stougaard J (1992) *Lotus japonicus*, an autogamous, diploid legume species for classical and molecular genetics. Plant J 2(4):487–496

Hansen J, Jørgensen J-E, Stougaard J et al (1989) Hairy roots—a short cut to transgenic root nodules. Plant Cell Rep 8:12–15

Hiraoka Y, Ueda H, Sugimoto Y (2009) Molecular responses of *Lotus japonicus* to parasitism by the compatible species *Orobanche aegyptiaca* and the incompatible species *Striga hermonthica*. J Exp Botany 60(2):641–650

Kai S, Tanaka H, Hashiguchi M et al (2010) Analysis of genetic diversity and morphological traits of Japanese *Lotus japonicus* for establishment of a core collection. Breeding Science 60:436–446

Kistner C, Winzer T, Pitzschke A et al (2005) Seven *Lotus japonicus* genes required for transcriptional reprogramming of the root during fungal and bacterial symbiosis. Plant Cell 17:2217–2229

Kouchi H, Imaizumi-Anraku H, Hayashi M et al (2010) How many peas in a pod? Legume genes responsible for mutualistic symbioses underground. Plant Cell Physiol 51(9):1381–1397

Kumagai H, Kinoshita E, Ridge RW et al (2006) RNAi knock-down of ENOD40s leads to significant suppression of nodule formation in *Lotus japonicus*. Plant Cell Physiol 47(8):1102–1111

Larsen K (1955) Cyto-taxonomical studies in Lotus II. Somatic chromosomes and chromosome numbers. Botanisk Tidsskrift 52:8–17

Lohar DP, Schuller K, Buzas DM et al (2001) Transformation of *Lotus japonicus* using the herbicide resistance bar gene as selectable marker. J Exp Bot 52 (361):1697–1702

Lombari P, Ercolano E, El Alaouni H et al (2003) A new transformation-regeneration procedure in the model legume *Lotus japonicus*: root explants as a source of large numbers of cells susceptible to *Agrobacterium*-mediated transformation. Plant Cell Rep 21:771–777

Lui J, Novero M, Charnikhova T et al (2013) CAROTENOID CLEAVAGE DIOXYGENASE 7 modulates plant growth, reproduction, senescence, and determinate nodulation in the model legume *Lotus japonicus*. J Exp Biol 64(7):1967–1981

Madsen EB, Madsen LH, Radutoiu S et al (2003) A receptor kinase gene of the LysM type is involved in legume perception of rhizobial signals. Nature 425:637–640

Madsen LH, Tirichine L, Jurkiewicz A et al (2010) The molecular network governing nodule organogenesis and infection in the model legume *Lotus japonicus*. Nature Comms 10. doi:10.1038/ncomms1009

Magori S, Oka-Kira E, Shibata S et al (2009) TOO MUCH LOVE, a root regulator associated with the long-distance control of nodulation in *Lotus japonicus*. MPMI 22(3):259–268

Novero M, Faccio A, Genre A et al (2002) Dual requirements of the *LjSym4* gene for mycorrhizal development in epidermal and cortical cells of *Lotus japonicus* roots. New Phytol 154:741–749

Oldroyd GED (2013) Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. Nature Rev Microbiol 11:252–263

Perez-Delgado CM, Garcia-Calderon M, Sanchez DH et al (2013) Transcriptomic and metabolic changes associated to photorespiratory ammonium accumulation in the model legume *Lotus japonicus*. Plant Phys 162(4):1834–1848

Perry J, Brachmann A, Welham T et al (2009) TILLING in *Lotus japonicus* identified large allelic series for symbiosis genes and revealed a bias in functionally

defective ethyl methanesulfonate alleles toward glycine replacements. Plant Phys 151:1281–1291

Poch HLC, Lopez RHM, Clark SJ (2007) Ecotypes of the model legume *Lotus japonicus* vary in their interaction phenotypes with the root-knot nematode *Meloidogyne incognita*. Ann Bot 99:1223–1229

Radutoiu S, Madsen LH, Madsen EB et al (2003) Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases. Nature 425:585–592

Reid JB, Ross JJ (2011) Mendel's genes: towards a full molecular characterization. Genetics 189:3–10

Sandal N, Petersen TR, Murray J et al (2006) Genetics of symbiosis in *Lotus japonicus*: recombinant inbred lines, comparative genetic maps, and map position of 35 symbiotic loci. MPMI 19(1):80–91

Soyano T, Kouchi H, Hirota A et al (2013) NODULE INCEPTION directly targets NF-Y subunit genes to regulate essential processes of root nodule development in *Lotus japonicus*. PLoS Genet 9(3): e1003352

Stougaard J (1993) Substrate-dependent negative selection in plants using a bacterial cytosine deaminase gene. Plant J 3(5):755–761

Stougaard Jensen J, Marcker KA, Otten L et al (1986) Nodule specific expression of a chimaeric soybean leghaemoglobin gene in transgenic *Lotus corniculatus*. Nature 321(6071):669–674

Stougaard J, Abildsten D, Marcker KA (1987) The *Agrobacterium rhizogenes* pRi TL-DNA segment as a gene vector system for transformation of plants. Mol Gen Genet 207:251–255

Stougaard J, Jørgensen J-E, Christensen T et al (1990) Interdependence and nodule specificity of *cis*-acting regulatory elements in the soybean leghemoglobin *lbc3* and N23 gene promoters. Mol Gen Genet 220:353–360

Takahara M, Magori S, Soyano T et al (2013) TOO MUCH LOVE, a novel Kelch repeat-containing F-box protein, functions in the long-distance regulation of the legume-Rhizobium symbiosis. Plant Cell Phys 54 (4):433–447

Tsyganov VE, Voroshilova VA, Priefer UB et al (2002) Genetic dissection of the initiation of the infection process and nodule tissue development in the *Rhizobium*-pea (*Pisum sativum* L.) symbiosis. Ann Bot 89:357–366

Urbanski DF, Malolepszy A, Stougaard J et al (2012) Genome-wide LORE1 retrotransposon mutagenesis and high-throughput insertion detection in *Lotus japonicus*. Plant J 69:731–741

Vriet C, Welham T, Brachmann A et al (2010) A suite of *Lotus japonicus* starch mutants reveals both conserved and novel features of starch metabolism. Plant Phys 154:643–655

Weerasinghe RR, Bird DMcK, Allen N (2005) Root-knot nematodes and bacterial Nod factors elicit common signal transduction events in *Lotus japonicus*. Proc Natl Acid Sci USA 102(8):3147–3152

Xu S, Luo Y, Cai Z et al (2013) Functional diversity of CYCLOIDEA-like TCP genes in the control of zygomorphic flower development in *Lotus japonicus*. J Intgr Plant Biol 55(3):221–231

Zagrobelsky M, Bak S, Ekstrøm CT et al (2007) The cyanogenic glucoside composition of *Zygaena filipendulae* (Lepidoptera: Zygaenidae) as effected by feeding on wild-type and transgenic lotus populations with variable cyanogenic glucoside profiles. Insect Biochem Mol Biol 37(1):10–18

Zhukov V, Radutoiu S, Madsen LH et al (2008) The pea *Sym37* receptor kinase gene controls infection-thread initiation and nodule development. MPMI 21 (12):1600–1608

Lotus Cytogenetics

2

Joana Ferreira and Andrea Pedrosa-Harand

Abstract

Most *Lotus* species have the basic chromosome number $x = 7$. The basic number $x = 6$ is, however, characteristic for the Corniculatus group and the other species from the section *Lotus*. Polyploidy, especially tetraploidy ($2n = 4x$), is recurrent in the genus with many species showing diploid and tetraploid accessions and others known as tetraploids only, such as *L. corniculatus*, the major forage crop. Genomes are relatively small, which, together with other interesting features, led to the choice of *L. japonicus* as a model legume species. Since then, advances in molecular cytogenetics, with the mapping of repetitive and single-copy sequences, enabled the integration of chromosomes to genetic maps and genome sequence information. Comparative cytogenetic maps were established for species from the section *Lotus*, mostly from the Corniculatus groups, and have demonstrated the importance of inversions and translocations, in addition to descending dysploidy and polyploidy, to the karyotype evolution of the genus.

2.1 Introduction

The first report on *Lotus* chromosomes was from 1924 (reviewed by Grant 1965). Since then, chromosome numbers have been reported for most of its species (reviewed by Grant 1995). The economic importance of *L. corniculatus* and related species has led to more detailed analyses of *Lotus* chromosomes, especially for understanding the origin of *L. corniculatus*, a polyploid crop species (Grant 1995). More recently, with the proposal of *L. japonicus* as a legume model, the fluorescent in situ hybridization (FISH) technique was applied to *Lotus* chromosomes (Ito et al. 2000), marking the transition from the classical to the molecular cytogenetic age (Jiang and Gill 2006).

In this chapter, we review the major advances in *Lotus* cytogenetics and its contribution to understanding *Lotus* genome organization and evolution.

J. Ferreira · A. Pedrosa-Harand (✉)
Laboratory of Plant Cytogenetics and Evolution, Department of Botany, Universidade Federal de Pernambuco, Recife-PE, Brazil
e-mail: andrea.pedrosaharand@pesquisador.cnpq.br

S. Tabata and J. Stougaard (eds.), *The Lotus japonicus Genome*, Compendium of Plant Genomes,
DOI 10.1007/978-3-662-44270-8_2,

2.2 Relationship Among *Lotus* Species

The genus *Lotus* comprises approximately 120–130 species and belongs to *Loteae*, a tribe of herbaceous species from temperate climates that was expanded by the inclusion of *Coronilleae* (Allan and Porter 2000). *Lotus* is the largest genus of the tribe and has the most complex taxonomic delimitation, mostly due to its high morphological and biogeographical diversity (Grant and Small 1996; Kramina and Sokoloff 2004; Kramina 2006). The circumscription of species and sections, as well as the genus itself, is controversial, but Degtjareva et al. (2006, 2008) considered the genus to be restricted to species native to Europe, Asia, Africa, and Australia, accepted the segregation of three Old World monotypic genera (*Kebirita*, *Podolotus,* and *Pseudolotus*) and included species commonly placed in *Dorycnium* and *Tetragonolobus* in *Lotus*. In this circumscription, 14 sections are recognized.

Phylogenetic analyses have contributed to elucidate the relationships among its species (Allan and Porter 2000; Arrambari 2000a, b; Allan et al. 2003; Degjareva et al. 2006, 2008). In general, those analyses have been congruent with major classical groups defined by morphological, reproductive, and cytotaxonomic approaches (Cheng and Grant 1973; Ross and Jones 1985; Arrambari et al. 2005; Barykina and Kramina 2006; Kramina 2006; Sokoloff et al. 2007).

The most investigated species of the genus belongs to the *L. corniculatus* group (Grant 1995), due to the fact that *L. corniculatus*, birdsfoot trefoil, is widely used as forage and for soil bioremediation in temperate regions (Díaz et al. 2005; Banuelos et al. 1992). Three other species were also domesticated: *L. glaber* Mill. (also known as *L. tenuis* Wald and Kit.), *L. uliginosus* Schkuhr (also considered synonymous with *L. pedunculatus* Cav.), and *L. subbiflorus* Lag. (Grant 1995; Gonnet and Diaz 2000; Scheffer-Basso et al. 2005). *Lotus glaber* and *L. uliginosus* are classically included in the Corniculatus group, together with *L. alpinus*, *L. borbassi*, *L. burttii*, *L. filicaulis*, *L. japonicus*, *L. krylovii*, *L. schoeleri,* and other species (Grant 1995). The phylogenetic analysis, based on ribosomal nuclear ITS (Internal Transcribed Spacer) and on morphologic characters, included in the same clade of *L. corniculatus* (also denominated Corniculatus group) almost all species cited above, plus *L. delortii*, *L. palustris*, *L. peczoricus*, *L. preslii,* and *L. stepposus* (Degtjareva et al. 2006, 2008). *Lotus uliginosus*, greater lotus, big trefoil or marsh birdsfoot trefoil, was, however, grouped with other species in the sister clade of the Corniculatus group, and *L. subbiflorus*, hairy birdsfoot trefoil, is now recognized as a less related species (Degtjareva et al. 2006).

2.3 Classic Cytogenetics

The species from the Corniculatus group were often investigated using classical cytogenetic methods, which were mainly aimed at contributing to the understanding of the origin of *L. corniculatus* and to its improvement (Sz-Borsos 1973; Ross and Jones 1985; Pupilli et al. 1990; Grant 1995; Grant and Small 1996; Gauthier et al. 1997). *Lotus corniculatus* is a tetraploid, with $2n = 4x = 24$ (Grant 1995). The other species of the group are diploids, also with basic chromosome number $x = 6$, which thus constitute a shared, derived character (synapomorphy) of the section *Lotus*, to which those species belong (Degtjareva et al. 2006).

Classic cytogenetics also has a long tradition in the genus *Lotus* outside the Corniculatus group, predominantly with cytotaxonomic studies comprising chromosome counts and karyotype descriptions (Cheng and Grant 1973; Freed and Grant 1976; Grant 1995). It was shown that in addition to $x = 6$ the genus also presents basic numbers $x = 5$ and 7. The basic number $x = 5$ is present in a single species of the section *Lotus*, while $x = 7$ is the most common and probably the ancestral basic chromosome number (reviewed by Grant 1995), observed in the ten sections with cytologically investigated species (Table 2.1). It probably gave rise to $x = 6$ and 5 by descending dysploidy. Supernumerary B-chromosomes have been reported in few species (Table 2.1).

Table 2.1 Basic chromosome number, ploidy level, and C-value of *Lotus* species represented in the genus phylogeny (Degtjareva et al. 2006, 2008)

Species[a]	Name status	Basic	Ploidy	1C (pg)[b]	References
***Lotus* sect. *Benedictella* (Maire) Kramina and D.D. Sokoloff (1/0)**					
***Lotus* sect. *Bonjeanea* (Rchb.) D.D. Sokoloff (3/3)**					
L. hirsutus L. [= *Dorycnium hirsutum* (L.) Ser.]	Synonym (ILDIS)	7	2*x*		IPCN (2013)
L. rectus L. [= *Dorycnium rectum* (L.) Ser.]	Synonym (ILDIS)	7	2*x*		IPCN (2013)
L. strictus Fisch. and C.A. Mey. [= *Dorycnium strictum* (Fisch. and C.A. Mey.) Lassen]	Synonym (ILDIS)	7	2*x*		Grant (1995)
***Lotus* sect Canaria (Rikli.) D.D. Sokoloff (3/0)**					
***Lotus* sect. Chamaelotus Kramina and D.D. Sokoloff (3/2)**					
L. glinoides Del. [= *L. trigonelloides* Webb and Berth.]	Accepted (ILDIS)	7	2*x*		Grant (1995)
L. schimperi Steud. ex Boiss	Accepted (ILDIS)	7	2*x*		IPCN (2013)
***Lotus* sect. Dorycnium (Mill.) D.D. Sokoloff (5/2)**					
L. dorycnium L. s.l.[= *Dorycnium herbaceum* Vill.]	Synonym (ILDIS)	7	2*x*		IPCN (2013)
L. graecus L. [= *Dorycnium graecum* (L.) Ser.]	Synonym (ILDIS)	7	2*x*		IPCN (2013)
***Lotus* sect. *Erythrolotus* Brand (0/0)**					
***Lotus* sect. *Heinekenia* Webb and Berth. (23/9)**					
Lotus arabicus group					
L. arabicus L.	Accepted (ILDIS)	6, 7	2*x*		Grant (1995)
L. lanuginosus Vent.	Accepted (ILDIS)	7	2*x*		Grant (1995)
L. laricus Rech.f., Aellen and Esfand	Accepted (ILDIS)	7	2*x*		IPCN (2013)
Lotus australis group					
L. australis Andrews	Accepted (ILDIS)	7	4*x*		Grant (1995)
L. cruentus Court	Accepted (ILDIS)	7	4*x*		Grant (1995)
Lotus discolor group					
L. discolor E. Mey	Accepted (ILDIS)	7	2*x*		Grant (1995)
Lotus gebelia group					
L. aegaeus (Griseb.) Nym	Accepted (ILDIS)	6, 7	4*x*		Grant (1995)
L. gebelia Vent.	Accepted (ILDIS)	7	2*x*		Grant (1995), IPCN (2013)
L. michauxianus Ser.	Accepted (ILDIS)	7	2*x*		IPCN (2013)

(continued)

Table 2.1 (continued)

Species[a]	Name status	Basic	Ploidy	1C (pg)[b]	References
***Lotus* sect. *Krokeria* (Moench) Ser (1/1)**					
L. edulis L.	Accepted (ILDIS)	7	2*x*	1.10	Grant (1995), IPCN (2013)
***Lotus* sect. *Lotea* (Medik.) DC. (10/8)**					
L. cytisoides L.	Accepted (ILDIS)	7	2*x*	1.40	IPCN (2013)
L. halophilus Boiss. and Spruner	Accepted (ILDIS)	7	2*x*, 4*x*		Grant (1995), IPCN (2013)
L. longiseliquosus R. Roem. [= *L. collinus* (Boiss.) Heldr.]	Accepted (ILDIS)	7	2*x*, 4*x*		Grant (1995), IPCN (2013)
L. ornithopodioides L.	Accepted (ILDIS)	7	2*x*	1.30[c]	Grant (1995), IPCN (2013)
L. peregrinus L.	Accepted (ILDIS)	7	4*x*		Grant (1995), IPCN (2013)
L. polyphyllos Clarke	Accepted (ILDIS)	6, 7	2*x*		Grant (1995)
L. tetraphyllus Murr.	Accepted (ILDIS)	7	2*x*		Grant (1995)
L. weilleri Maire	Accepted (ILDIS)	7	2*x*		Grant (1995)
***Lotus* sect. *Lotus* (31/22)**					
L. angustissimus group					
L. angustissimus L. [= *L. praetermissus* Kuprian.]	Accepted (ILDIS)	6	2*x*, 4*x*		Grant (1995), IPCN (2013)
L. castellanus Boiss. and Reut. [= *L. subbiflorus* Lag.]	Synonym (ILDIS)	6	2*x*		IPCN (2013)
L. castellanus Boiss. and Reut. [= *L. glareosus* Boiss. and Reut.]	Synonym (ILDIS)	6	2*x*		Grant (1995), IPCN (2013)
L. parviflorus Desf.	Accepted (ILDIS)	6	2*x*		Grant (1995), IPCN (2013)
L. subbiflorus Lag. [= *L. suaveolens* Pers.]	Accepted (ILDIS)	6	2*x*, 4*x*		Grant (1995), IPCN (2013)
Lotus corniculatus group					
L. alpinus (DC.) Schleicher ex Ramond	Accepted (ILDIS)	6 + B	2*x*, 4*x*, 6*x*	0.48	Grant (1995), IPCN (2013)
L. borbasii Ujhelyi	Accepted (ILDIS)	6	2*x*	0.50	Grant (1995)
L. burttii Borsos	Accepted (ILDIS)	6	2*x*	0.53	Grant (1995)
L. corniculatus L.	Accepted (ILDIS)	6	4*x*[d]	0.48, 1.05	Grant (1995), IPCN (2013)
L. delortii Timb.-Lagr. ex F.W. Schultz [= *L. pilosus* Jordan]	Accepted (ILDIS)	6	4*x*		Grant (1995)
L. filicaulis Durieu[= *L. tenuis* Waldst. and Kit. ex Willd.]	Synonym (ILDIS)	6	2*x*	0.50	Grant (1995)

(continued)

Table 2.1 (continued)

Species[a]	Name status	Basic	Ploidy	1C (pg)[b]	References
L. glaber Mill. [= *L. tenuis* Waldst. and Kit]	Accepted (ILDIS)	6[e]	2*x*, 4*x*	0.48	Grant (1995), IPCN (2013)
L. japonicus (Regel) K. Larsen 'Gifu' [= *L. corniculatus* subsp. *corniculatus* L.]	Synonym (ILDIS)	6	2*x*	0.48	Grant (1995), IPCN (2013)
L. japonicus (Regel) K. Larsen 'Miyakojima' [= *L. corniculatus* subsp. *corniculatus* L.]	Synonym (ILDIS)	6	2*x*		Grant (1995), IPCN (2013)
L. krylovii Schischk. and Serg.	Accepted (ILDIS)	6	2*x*	0.53	Grant (1995), IPCN (2013)
L. palustris Willd.	Accepted (ILDIS)	6, 7	2*x*, 4*x*	0.75	Grant (1995)
L. peczoricus Miniaev and Ulle	Accepted (ILDIS)	6	2*x*		Grant (1995)
L. preslli Tem.	Accepted (ILDIS)	6	2*x*, 4*x*		Grant (1995), IPCN (2013)
L. schoelleri Schweinf.	Accepted (ILDIS)	6	2*x*	0.50	Grant (1995)
L. conimbricensis Brot. [= *L. coimbrensis* Brot. ex Willd.]	Accepted (ILDIS)	6	2*x*	0.45	Grant (1995), IPCN (2013)
Lotus pedunculatus group					
L. pedunculatus Cav.	Accepted (ILDIS)	6	2*x*, 4*x*	0.55	Grant (1995), IPCN (2013)
L. uliginosus Schkuhr [= *L. pedunculatus* Cav.]	Synonym (ILDIS)	6	2*x*, 4*x*	0.55	Grant (1995), IPCN (2013)
***Lotus* sect. *Ononidium* Boiss. (4/0)**					
***Lotus* sect. *Pedrosia* (Lowe) Christ (29/10)**					
L. arenarius Brot.	Accepted (ILDIS)	7	2*x*, 4*x*	1.13	Grant (1995), IPCN (2013)
L. azoricus P.W. Ball [= *L. macranthus* Lowe]	Accepted (ILDIS)	7[f]	2*x*		Grant (1995), IPCN (2013)
L. campylocladus Webb and Berth	Accepted (ILDIS)	7	2*x*	0.62	Grant (1995), IPCN (2013)
L. creticus L.	Accepted (ILDIS)	7 + B	2*x*, 4*x*		Grant (1995), IPCN (2013)
L. emeroides R.P. Murray	Accepted (ILDIS)	7	2*x*, 4*x*		Grant (1995), IPCN (2013)
L. jacobaeus L.	Accepted (ILDIS)	7	2*x*		Grant (1995), IPCN (2013)
L. jolyi Battand	Accepted (ILDIS)	7	2*x*		Grant (1995), IPCN (2013)
L. lancerottensis Webb and Berth	Accepted (ILDIS)	7	2*x*		Grant (1995), IPCN (2013)
L. maroccanus Ball	Accepted (ILDIS)	7	2*x*		Grant (1995), IPCN (2013)
L. mascaensis Burchd	Accepted (ILDIS)	7	4*x*	1.25	Grant (1995), IPCN (2013)

(continued)

Table 2.1 (continued)

Species[a]	Name status	Basic	Ploidy	1C (pg)[b]	References
***Lotus* sect. *Rhyncholotus* (Manod) D.D. Sokoloff (3/2)**					
L. berthelotii Masf	Accepted (ILDIS)	7	4*x*	1.22	Grant (1995), IPCN (2013)
L. maculatus Breitf	Accepted (ILDIS)	7	4*x*		Grant (1995), IPCN (2013)
***Lotus* sect. *Tetragonolobus* (Scop.) Benth. and Hook.f. (5/2)**					
L. maritimus L. [= *Tetragonolobus maritimus* (L.) Roth.]	Accepted (ILDIS)	7 [g]	2*x*		Grant (1995), IPCN (2013)
L. tetragonolobus L. [= *T. purpureus* Moench.]	Accepted (ILDIS)	7	2*x*		Grant (1995), IPCN (2013)

[a] Species names and name status are based on The Plant List (2010). Version 1. Sections of *Lotus* are based on Degtjareva et al. (2006, 2008). Numbers after sectional names show total number of species in a section/number of species included here

[b] C-values from Bennett and Leitch (2012)

[c] C-value for *L. ornithopoides*

[d] 2*x* was reported, but is not anymore accepted

[e] Chromosome number for *L. tenuis*

[f] Chromosome number for *L. macranthus*

[g] Chromosome number for *T. maritimus*

Genome sizes are relatively small and have been estimated for 26 species (Bennett and Leitch 2012), even before the C-value was considered for estimating genome coverage in genome sequencing projects. Estimates are available for around 20 % of the species of the genus, comprising representatives from five out of the fourteen sections (see Table 2.1). Minimum and maximum genome sizes were 0.45 pg/1C for *L. conimbricensis* and 1.40 pg/1C for *L. cytisoides*, an approximate threefold difference in genome size at the diploid level within the genus.

Chromosome differential staining techniques, such as C-banding, which allows the differentiation between euchromatin and heterochromatin, have been applied to three species: *L. pedunculatus*, *L. tenuis* and *L. japonicus* (Shankland and Grant 1976; Falistocco and Piccirilli 1989; Pedrosa et al. 2002). Because heterochromatic regions remain condensed during most of the cell cycle, they appear as more condensed regions during mitotic prometaphase. Thus, imaging analysis of prometaphase chromosomes has also been used to construct idiograms for *L. japonicus* (Ito et al. 2000; Ohmido et al. 2007). Both approaches revealed that the heterochromatin is mainly located at pericentromeric regions, with terminal and intercalary blocks in few chromosomes and variation in heterochromatin distribution between genotypes of *L. japonicus* (Ito et al. 2000; Hayashi et al. 2001).

2.4 Molecular Cytogenetics in *Lotus*

Various repetitive DNA sequences have been used as probes in FISH experiments to investigate their distribution along *Lotus* chromosomes. The FISH technique consists of denaturing the chromosomes on microscopic preparations to separate the two complimentary DNA strands, followed by their renaturation in the presence of a probe, a labeled DNA fragment. The excess of available probe will compete against the chromosomal DNA strands, allowing its localization on chromosomes (Jiang and Gill 2006). For example, probes for ribosomal RNA coding sequences 5S and 45S rDNA were applied to several plants because these sequences are

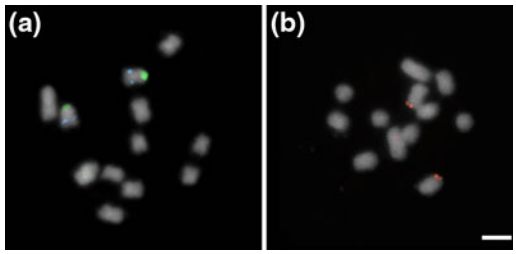


Fig. 2.1 Fluorescent in situ hybridization on mitotic metaphase chromosomes of *Lotus japonicus* 'Gifu.' **a** TAC 28L17/TM0153 (*blue*) is positioned on the opposite chromosome arm of 45S rDNA (green). **b** TAC 15K21/TM0088 (*orange*). Both TACs are located on the second largest chromosome and identify the chromosome 2. Chromosomes were counterstained with DAPI and are shown in gray. Bar in **b** = 5 μm

conserved and repeated in tandem, generating signals that are usually easily visualized on chromosomes (reviewed by Kato et al. 2005).

In *L. japonicus*, the 5S rDNA site was located interstitially in the short arm of chromosome 2, linked to a 45S rDNA site that was terminally located in the same chromosome arm (Hayashi et al. 2001; Pedrosa et al. 2002). In addition to this major 45S rDNA site on chromosome 2 (Fig. 2.1a), minor 45S rDNA sites were observed in the smallest chromosomes pairs, 5 and 6, in interstitial positions. Both probes have also been applied to other species of the Corniculatus group, showing that the linkage between 5S and 45S rDNA sites on chromosome 2 is conserved in *L. filicaulis* (Pedrosa et al. 2002), *L. burttii* (Kawaguchi et al. 2005), *L. glaber*, and *L. krilovii* (Fig. 2.2a, c). Except for *L. krilovii*, the 45S rDNA site on chromosome 6 was also present in the investigated species, but the weakest site on chromosome 5 has only been detected in *L. japonicus* 'Gifu' and 'Miyakojima'. Mapping of 5S and 45S on *L. uliginosus*, however, revealed more pronounced differences, although the rDNA sites on chromosome 2 were maintained. An additional 5S rDNA site was observed on chromosome 6, and two additional 45S rDNA sites were present on chromosomes 4 and 5, both in terminal positions (Ferreira et al. 2012).

Other repetitive DNA sequences have also been identified and localized to *Lotus* chromosomes. The *Ljcen1* repeat was identified because of its similarity to the Arabidopsis-type telomeric repeat and turned out to be centromeric, not only in *L. japonicus*, but also in other investigated species from the Corniculatus group, such as *L.*

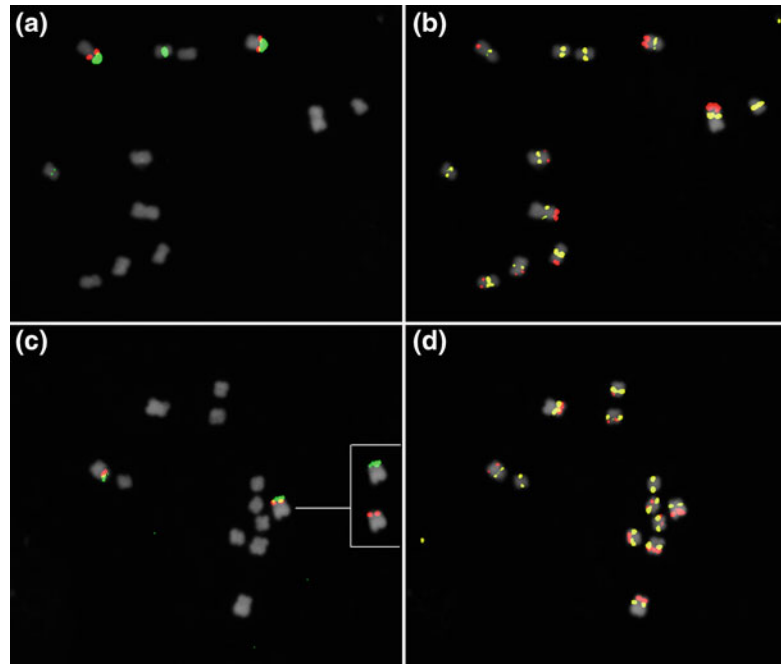


Fig. 2.2 Fluorescent in situ hybridization of repetitive sequences on mitotic metaphase chromosomes of diploids *L. glaber* (**a**, **b**) and *L. krilovii* (**c**, **d**). (**a**, **c**) 45S (*green*) and 5S (*orange*) rDNA, and (**b**, **d**) *Ljcen1* (*yellow*) and LJTR1 (*red*). Chromosomes were counterstained with DAPI and are shown in gray. Bar in (**d**) = 5 μm

filicaulis (Pedrosa et al. 2002), *L. burttii* (Kawaguchi et al. 2005), *L. glaber,* and *L. krilovii* (Fig. 2.2b, d). Later, a Ty3-*gypsy* LTR-retrotransposon, named LjRE2, was shown to have the same distribution as *Ljcen1* (Sato et al. 2008), as *Ljcen1* shows high sequence similarity to the LTR region of LjRE2 (Ohmido et al. 2010). The other characterized LTR-retrotransposon, LjRE1, a Ty1-*copia* type, showed a dispersed labeling of all chromosomes (Sato et al. 2008). Four tandem repeat sequences, LjTR1-4, were distributed in specific chromosomal regions, forming blocks associated with eu- or heterochromatin in prometaphase or pachytene chromosomes (Sato et al. 2008; Ohmido et al. 2010). LjTR1 has also been localized to *L. glaber* and *L. krilovii* mitotic metaphase chromosomes, showing similar patterns of terminal blocks of varying intensities in the short or the long chromosome arm, except for chromosome 5 (Fig. 2.2b, d).

2.5 Integrated Genetic and Cytogenetic Maps in *Lotus*

After *L. japonicus* had been chosen as a model legume, genetic maps were established as a first step toward positional cloning (Handberg and Stougaard 1992; Sato and Tabata 2006). The first maps, which included AFLPs, RAPDs, RFLPs, SSRs, and dCAPS markers, as well as mutant phenotypes, were based on mapping populations obtained from crosses between *L. japonicus* ecotypes, 'Gifu' and 'Miyakojima,' or between *L. japonicus* and a closely related species from the Corniculatus group, *L. filicaulis* (Hayashi et al. 2001; Sandal et al. 2002). The first version of these maps, however, presented distortions in the recombination frequencies, leading to maps with five or seven linkage groups, instead of the expected six.

In parallel to the genetic mapping efforts, cytogenetic maps were built using genomic DNA clones with large, single-copy inserts, such as BACs (bacterial artificial chromosomes) and TACs (transformation-competent artificial chromosomes). Cytogenetic maps are physical maps in which DNA sequences are localized on the chromosomes and positioned in relation to centromeres, telomeres, and the heterochromatin and are usually developed by FISH. The *Lotus* BACs and TACs used as probes were anchored to the genetic maps, allowing the integration of linkage groups and chromosomes (Fig. 2.1). These integrated cytogenetic maps helped to establish six linkage groups in each map, which were named according to the six chromosome pairs. Furthermore, they revealed chromosome rearrangements between the parental accessions or species, which were responsible for the observed segregation distortions (Hayashi et al. 2001; Pedrosa et al. 2002). TACs have later been used to mitotic prometaphase and meiotic pachytene chromosomes for higher resolution mapping (Sato et al. 2008; Ohmido et al. 2010). The availability of those BACs and TACs as chromosome markers and the indication of rearrangements among closely related genotypes stimulated the investigation of chromosome evolution in the genus.

2.6 Comparative Cytogenetics in *Lotus*

The establishment of cytogenetic maps for *L. japonicus* made available a set of chromosome-specific markers that could be used to build similar maps in related species. These comparative maps allow exploration of the macrosynteny and collinearity among genomes and investigation of karyotype evolution in more detail.

In *Lotus*, paracentric and pericentric inversions and translocations could be clearly demonstrated between *L. japonicus* ecotypes 'Gifu' and 'Miyakojima' and between *L. japonicus* and *L. burttii* and *L. filicaulis* (Hayashi et al. 2001; Pedrosa et al. 2002; Kawaguchi et al. 2005). Between 'Gifu' and 'Miyakojima', a reciprocal translocation has exchanged the terminal portions of chromosome 1 short arm and chromosome 2 long arm. When the same chromosome markers were mapped in *L. burttii* and *L. filicaulis*, synteny with 'Gifu' was observed, what indicates that 'Gifu' chromosomes 1 and 2 represent the ancestral (plesiomorphic) condition. On the other hand, the inversion in a small portion of the long

arm of *L. japonicus* chromosome 1, when compared to the other two species, seemed to be the derived (apomorphic) condition, as well as a pericentric inversion on *L. filicaulis* chromosome 3, which is acrocentric and has so far only been observed as acrocentric in this species.

Lotus japonicus ecotypes 'Miyakojima' and 'Gifu' present other cytogenetic differences. The TAC 28L17, mapped on 'Miyakojima' between the 5S and 45S rDNA sites on the short arm of chromosome 2, is positioned on the opposite chromosomal arm on 'Gifu' (Fig. 2.1a). Furthermore, terminal heterochromatic blocks are more frequent in 'Miyakojima' than in 'Gifu.' These ecotypes appear to have not only enough genomic differences, but also distinct morphological characters to be considered two species: *L. japonicus* (Regel) K. Larsen and *L. miyakojimae* Kramina (Barykina and Kramina 2006). In fact, it was also suggested in the first phylogeny (Degtjareva et al. 2006) and considered in the last update (Degtjareva et al. 2008).

More recently, the comparative map was expanded to *L. uliginosus*, a phylogenetically more distant species (Degtjareva et al. 2006), which does not belong to the Corniculatus group (Ferreira et al. 2012). A different translocation was observed, involving chromosomes 3 and 5. Karyotypic differences were more pronounced between *L. uliginosus* and *L. japonicus* than between any Corniculatus species, reflecting their phylogenetic distances (Fig. 2.3).

2.7 *Lotus* Polyploids

Although most *Lotus* species are diploids, polyploids, particularly tetraploids, are of relevance in the genus because polyploidy is observed in at least five sections and most of the cultivated accessions are polyploids. *Lotus corniculatus* is the classical example, but even in species known as diploid, such as *L. uliginosus*, its cultivars may be polyploid, such as 'Maku,' with $2n = 4x = 24$. Indeed, several species are reported to have diploid and tetraploid accessions, such as *Lotus subbiflorus* (see Table 2.1).

Lotus subbiflorus also belongs to the section *Lotus*, but is placed in clade A, a sister clade to clade B, where *L. corniculatus* is present (Degtjareva et al. 2006). One polyploid accession has been recently investigated using rDNA and *Ljcen1* probe and this analysis gave support for an allopolyploid origin for this species. The first evidence came from the number and distribution of 5S and 45S rDNA sites. One chromosome pair showed linked 5S and 45S rDNA sites, as observed for chromosome 2 in the Corniculatus group, but the possible homeologous pair showed a 45S rDNA cluster only. A second 5S rDNA site was in one smaller chromosome pair (Fig. 2.4a). In addition, *Ljcen1* only strongly labeled one set of chromosomes (Fig. 2.4b), suggesting that the two diploid species that hybridized to form the *L. subbiflorus* genome showed remarkable karyotype differences. Because its closely related, diploid species have not been investigated to date, it is still not possible to suggest putative ancestral species.

The origin of *L. corniculatus* has been investigated in more detail. Classical cytogenetic analysis, as well as biochemical and morphological markers, have been employed. The most recent hypothesis considered this an allotetraploid species originating from the crossing of *L. tenuis* and *L. uliginosus* (Ross and Jones 1985; Grant and Small 1996). Other possible diploids considered to be involved in the origin of *L. corniculatus* are *L. alpinus* and *L. japonicus* (Grant and Small 1996) or *L. schoelleri*, *L. stepposus*, *L. peczoricus*, *L. borbasii*, *L. krylovii,* and *L. japonicus* (Degtjareva et al. 2006).

From these, *L. glaber* (a synonym of *L. tenuis*), *L. uliginosus*, *L. japonicus,* and *L. krylovii* have been investigated cytogenetically in more detail and compared to *L. corniculatus*. *L. glaber,* and *L. japonicus* 'Gifu' have the most similar karyotypes, with 5S and 45S rDNA sites in chromosome 2 and a 45S rDNA site in chromosome 6. *L. corniculatus* chromosomes, when analyzed with the same probes, showed double the number of rDNA sites in similar positions (Fig. 2.4c). *L. krylovii* apparently lacks the 45S rDNA site in chromosome 6 and *L. uliginosus* is

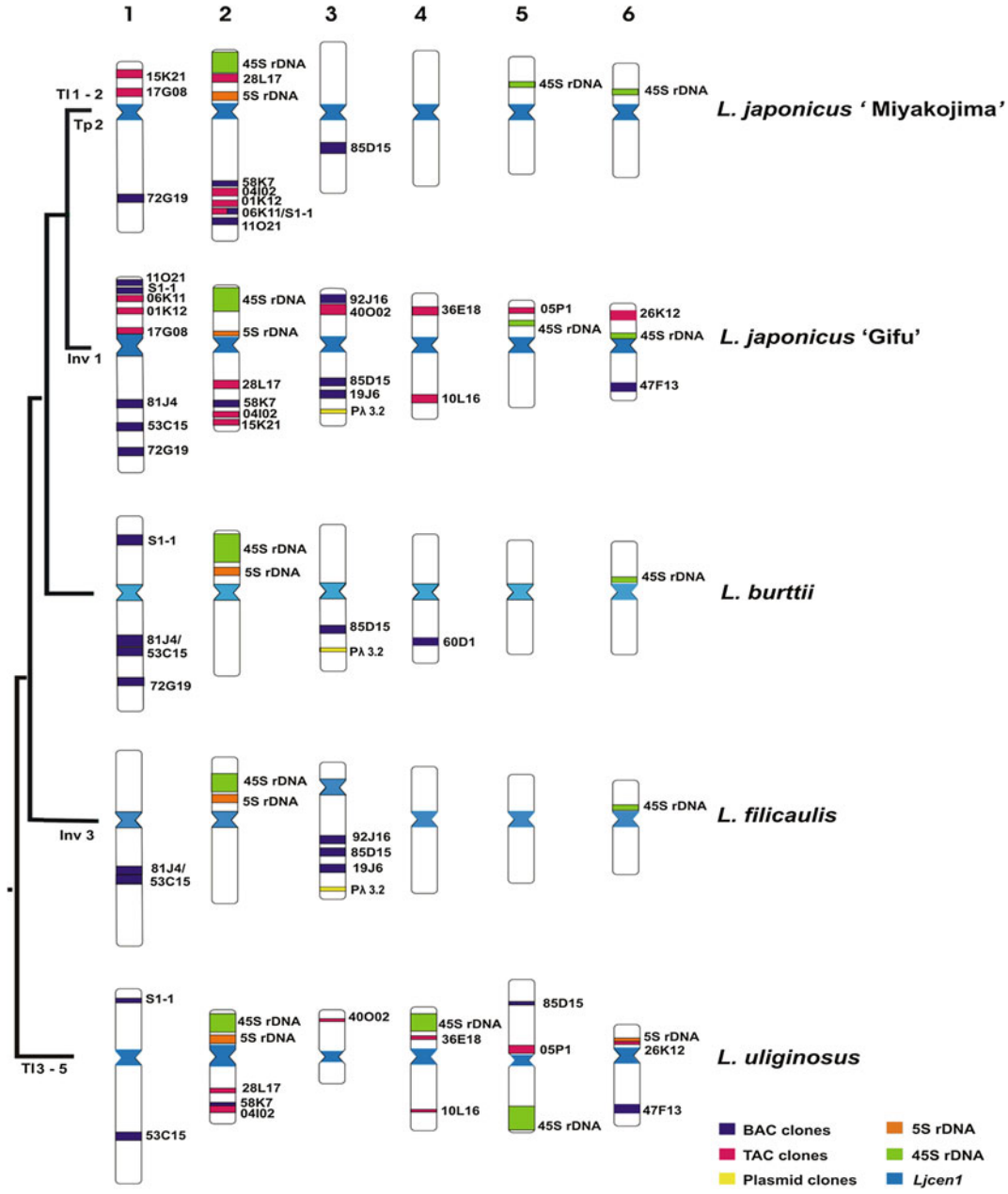


Fig. 2.3 Comparative schematic representation of the chromosome complement of *L. japonicus* 'Miyakojima' and 'Gifu', *L. burttii*, *L. filicaulis* (modified from Hayashi et al. 2001; Pedrosa et al. 2002; Sato et al. 2008), and *L. uliginosus*. Approximated positions of rDNA sites, pericentromeric repeat *Ljcen1*, and mapped TAC/BAC clones are represented. TACs are visualized in red and BACs in dark blue (thin blocks represent weaker signals in *L. uliginosus*). *Lotus uliginosus* chromosomes 3 and 5 were rotated (short arm down) to facilitate comparison. Phylogenetic relationships are based on Degtjareva et al. (2006, 2008). The proposed rearrangements (Tl = translocation, Tp = transposition, and Inv = inversion) are indicated (Ferreira et al. 2012)

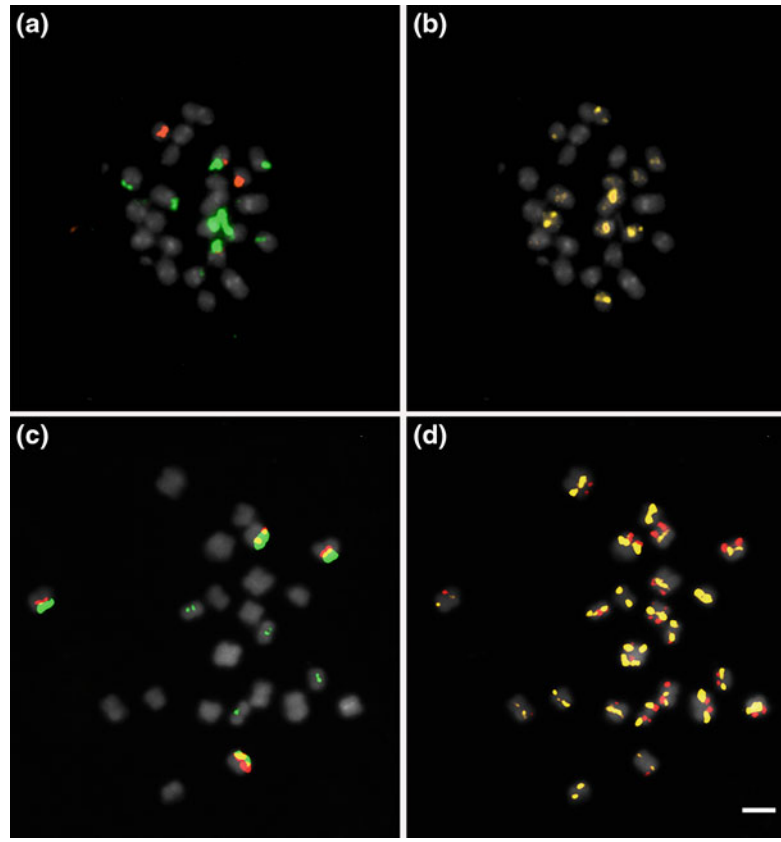


Fig. 2.4 Fluorescent in situ hybridization of repetitive sequences on mitotic metaphase chromosomes of polyploids *L. subbiflorus* (**a**, **b**) and *L.corniculatus* (**c**, **d**). (**a**, **c**) 45S (*green*) and 5S (*orange*) rDNA, (**b**, **d**) *Ljcen1* (*yellow*) and (**d**) LJTR1 (*red*). Note that *Ljcen1* signals are present in only one set of chromosomes of *L. subbiflorus*, suggesting an allotetraploid origin. Chromosomes were counterstained with DAPI and are shown in gray. Bar in (**d**) = 5 μm

clearly very different in rDNA distribution. Current cytogenetic evidence would suggest *L. glaber* and *L. japonicus* as possible ancestral species of *L. corniculatus*, or other closely related species with similar karyotypes (Fig. 2.4c–d).

Acknowledgments We thank Federico Condón (Instituto Nacional de Investigación Agropecuária – INIA, Uruguay), Miguel Dall'Agnol (Universidade Federal do Rio Grande do Sul, Brazil), Niels Sandal (Aarhus University, Denmark) and Shusei Sato (Kazusa DNA Research Institute, Japan) for seeds and probes and Sandra Mendes (Universidade Federal de Pernambuco, Brazil) for drawing the original Fig. 2.3. Andrea Pedrosa-Harand thanks the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Brazil, for financial support.

References

Allan GJ, Porter JM (2000) Tribal delimitation and phylogenetic relationships of Loteae and Coronilleae (Faboideae: Fabaceae) with special reference to *Lotus:* evidence from nuclear ribosomal ITS sequences. Am J Bot 87:1871–1881

Allan GJ, Zimmer EA, Wagner WL, Sokoloff DD (2003) Molecular phylogenetic analyses of the tribe Lotae (Leguminosae), implications for classification and biogeography. In: Klitgaard BB, Bruneau A (eds) Advances in Legume Systematics, parte 10. High Level Systematics, Royal Botanic Gardens, Kew, pp 371–393

Arrambari AM (2000a) A cladistic analysis of the New World species of *Lotus* L. (Fabaceae, Lotae). Cladistics 16:283–297

Arrambari AM (2000b) A cladistic analysis of the Old World species of *Lotus* L. (Fabaceae: Lotae). Can J Bot 78:351–360

Arrambari AM, Stenglein AS, Colares MN, Novoa MC (2005) Taxonomy of the New World species of *Lotus* (Leguminosae: Lotae). Aust J Bot 53:797–812

Banuelos G, Cardon G, Mackey B et al (1992) Boron and selenium removal in boron-laden soil by birdsfoot trefoil. Lotus Newsletter 23:32–35

Barykina RP, Kramina T (2006) A comparative morphological and anatomical study of the model legume *Lotus japonicus* and related species. Wulfenia 13:33–56

Bennett MD, Leitch IJ (2012) Plant DNA C-values database (release 6.0, December 2012). http://data.kew.org/cvalues/. Accessed 29 June 2013

Cheng RI-J, Grant W (1973) Species relationships in the *Lotus corniculatus* group as determined by karyotype and cytophotometric analyses. Genome 15:101–115

Degtjareva GV, Kramina TE, Sokoloff DD et al (2006) Phylogeny of the genus *Lotus* (Leguminosae, Loteae): evidence from nrITS sequences and morphology. Can J Bot 84:813–830

Degtjareva GV, Kramina TE, Sokoloff DD et al (2008) New data on nrITS phylogeny of *Lotus* (Leguminosae, Lotae). Wulfenia 15:35–49

Díaz P, Borsani O, Monza J (2005) *Lotus*-related species and their agronomic importance. In: Márquez (ed) *Lotus japonicus* Handbook, Springer, Berlin, pp 25–37

Falistocco E, Piccirilli M (1989) The basic karyotype of *Lotus tenuis* C-banding and Feulgen studies. Ann Bot 63:401–404

Ferreira J, Mendes S, Dall'Agnol M et al (2012) Comparative analyses in *Lotus*: the cytogenetic map of *Lotus uliginosus* Schkuhr. Cytogenet Genome Res 137:42–49. doi:10.1159/000339617

Freed HJ, Grant WF (1976) Polytene chromosomes in the suspensor cells of *Lotus* (Fabaceae). Caryologia 29:387–390

Gauthier P, Lumaret R, Bedecarrats A (1997) Chloroplast-DNA variation in the genus *Lotus* (Fabaceae) and further evidence regarding the maternal parentage of *Lotus corniculatus* L. Theor Appl Genet 95:629–636

Gonnet S, Díaz P (2000) Glutamine synthetase and glutamate synthase activities in relation to nitrogen fixation in *Lotus* Spp. R Bras Fisiol Veg 12: 195–202

Grant WF (1965) A chromosome atlas and interspecific hybridization index for the genus *Lotus* (Leguminosae). Can J Genet Cytol 7:457–471

Grant WF (1995) A chromosome atlas and interspecific-intergenic index for *Lotus* and *Tetragonolobus* (Fabaceae). Can J Bot 73:1787–1809

Grant WF, Small E (1996) The origin of the *Lotus corniculatus* (Fabaceae) complex: a synthesis of diverse evidence. Can J Bot 74:975–989

Handberg K, Stougaard J (1992) *Lotus japonicus*, an autogamous, diploid legume species for classical and molecular genetics. Plant J 2:487–496

Hayashi M, Miyahara A, Sato S et al (2001) Construction of a genetic linkage map of the model legume *Lotus japonicus* using an intraspecific F2 population. DNA Res 8:301–310

IPCN (2013) Index to plant chromosome numbers (1979–2013) In: Goldblatt P, Johnson DE (eds) Missouri botanical garden, St. Louis. http://www.tropicos.org/Name/13010001. Accessed 29 Jun 2013

Ito M, Miyamoto J, Mori Y et al (2000) Genome and chromosome dimensions in *Lotus japonicus*. J Plant Res 113:435–442

Jiang J, Gill BS (2006) Current status and the future of fluorescence *in situ* hybridization (FISH) in plant genome research. Genome 49:1057–1068

Kato A, Vega JM, Han F et al (2005) Advances in plant chromosome identification and cytogenetic techniques. Curr Opin Plan Biol 8:148–154

Kawaguchi M, Pedrosa-Harand A, Yano K et al (2005) *Lotus burttii* takes a position of the third corner in the *Lotus* molecular genetics triangle. DNA Res 12:69–77

Kramina T (2006) A contribution to the taxonomic revision of the *Lotus angustissimus*-complex (Leguminosae, Lotae). Wulfenia 13:57–92

Kramina T, Sokoloff D (2004) A taxonomic study of *Lotus australis* complex (Leguminosae), with special emphasis on plants from Pacific Ocean islands. Adansonia 26:171–197

Ohmido N, Sato S, Tabata S, Fukui K (2007) Chromosome maps of legumes. Chromosome Res 15:97–103

Ohmido N, Ishimaru A, Kato S et al (2010) Integration of cytogenetic and genetic linkage maps of *Lotus japonicus*, a model plant for legumes. Chromosome Res 18:287–299

Pedrosa A, Sandal N, Stougaard J et al (2002) Chromosomal map of the model legume *Lotus japonicus*. Genetics 161:1661–1672

Pupilli F, Arcioni A, Damiani F, Pezzoti M (1990) Plant regeneration from callus and protoplast cultures of *Lotus pedunculatus* Car. Plant Cell Tiss Org 23:193–199

Ross MD, Jones WT (1985) The origin of *Lotus corniculatus*. Theor Appl Genet 71:284–288

Sandal N, Krusell L, Radutoiu S et al (2002) An AFLP based genetic linkage map of the model legume *Lotus japonicus* developed from an interspecific F2 mapping population. Genetics 161:1673–1683

Sato S, Tabata S (2006) *Lotus japonicus* as a platform for legume research. Curr Opin Plant Biol 9:128–132

Sato S, Nakamura Y, Kaneko T et al (2008) Genome structure of the legume, *Lotus japonicus*. DNA Res 15 (4):227–239. doi:10.1093/dnares/dsn008

Scheffer-Basso SM, Vendrúscolo MC, Cecchetti D (2005) Desempenho de leguminosas nativas (*Adesmia*) e exóticas (*Lotus, Trifolium*), em função do estádio fenológico no primeiro corte. R Bras Zootec 34:1871–1880

Shankland NE, Grant WF (1976) Localization of Giemsa bands in *Lotus pedunculatus* chromosomes. Can J Gen Cytol 18:239–244

Sokoloff DD, Degtjareva GV, Endress PK et al (2007) Inflorescence and early flower development in Lotae (Leguminosae) in a phylogenetic and taxonomic context. Int J P Sci 168:801–833

Sz.-Borsos O (1973) Cytophotometric studies on the DNA contents of diploid *Lotus* species. Acta Bot Acad Sci Hungar 18:49–58

The Plant List (2010) Version 1. http://www.theplantlist.org/. Accessed 29 June 2013

Genetic Linkage Maps, Synteny and Map-based Cloning

3

Niels Sandal and Shusei Sato

Abstract

Nitrogen fixation is a very important trait in agriculture and nature. It is made possible through symbiosis between plants, mainly legumes, and microorganisms such as rhizobia. Like most plants, legumes have symbiosis with mycorrhizal fungi. In order to isolate the plant genes that are important for symbiosis with nitrogen-fixing organisms and mycorrhizal symbiosis, *Lotus japonicus* was suggested as a model legume by Handberg and Stougaard (1992).

3.1 Mapping with Crosses Between *L. japonicus* Gifu and *L. japonicus* Funakura and *L. filicaulis*

Mapping and map-based cloning are a very powerful procedure to isolate genes based on phenotypes of mutants and species/ecotypes and molecular markers. Therefore, mutant isolation and mapping in *Lotus* were initiated. For mapping, you need differences at the DNA and phenotypic levels. Handberg and Stougaard elected to work on the ecotype Gifu B-129 (growing in the Gifu prefecture in the middle of Japan) (Stougaard and Beuselinck 1996). Many of the *Lotus japonicus* mutants have been made in the Gifu ecotype. To initiate mapping, polymorphic (with differences at the DNA level) ecotype(s)/species had to be found. The level of polymorphisms in the different ecotypes known at that time did not appear to be high, but Funakura was used as a mapping partner by Jiang and Gresshoff (1997). Grant and co-workers from Canada showed that several diploid species of *Lotus* can be crossed with *L. japonicus* and give fertile offspring (de Nettancourt and Grant 1964; Somaroo and Grant 1971). One of the best crossing partners turned out to be *L. filicaulis,* which originates from Algeria (de Nettancourt and Grant 1964; Sandal et al. 2002, 2006). This was the basis of one of the first F2 mapping populations in *Lotus*. When no sequence data are available, the amplified fragment length polymorphism (AFLP) technique is very powerful as many polymorphic AFLP markers can be found

N. Sandal (✉)
Centre for Carbohydrate Recognition and Signalling, Department of Molecular Biology and Genetics, Aarhus University, Gustav Wieds Vej 10, 8000 Aarhus C, Denmark
e-mail: nns@mb.au.dk

S. Sato
Graduate School of Life Sciences, Tohoku University, Katahira 2-1-1, Aoba-ku, Sendai 980-8577, Japan
e-mail: shuseis@ige.tohoku.ac.jp

S. Tabata and J. Stougaard (eds.), *The Lotus japonicus Genome*, Compendium of Plant Genomes,
DOI 10.1007/978-3-662-44270-8_3,

in a short time (Vos et al. 1995). The level of polymorphisms is very high between Gifu and *L. filicaulis*. In that way, a number of genetic markers were found and they could be grouped into six linkage groups corresponding to the six chromosomes of *L. japonicus* (Sandal et al. 2002). Recombinant inbred lines (RILs), generation S8, have been developed from *L. filicaulis* x *L. japonicus* Gifu (Sandal et al. 2006). The advantage of RILs is that they are almost completely homozygous, and therefore, mapping information can be shared between groups working with these lines.

3.2 Ecotype Miyakojima MG-20

The ecotype *L. japonicus* Miyakojima MG-20 was found on one of the southern islands of Japan (Kawaguchi 2000). Because of the large geographical distance to most of the other ecotypes, it has a high level of polymorphisms to the Gifu ecotype. *L. japonicus* MG-20 also has the advantage of growing very well in indoor systems (Kawaguchi et al. 2001) and was therefore chosen for genome sequencing by Kazusa DNA Research Institute in Japan with a BAC/TAC sequencing strategy (Sato et al. 2008 and references therein). With the sequences of BAC or TAC clones, it became possible to look for microsatellite sequences such as TATATATA that are highly polymorphic in, for example, TA repeat number and can therefore be used to develop microsatellite size markers that are easier to work with than AFLP markers. Microsatellite markers can be tested in high percentage (2–4%) agarose gels to see small size differences. The microsatellite markers were developed to work with the same PCR conditions, such as an annealing temperature of 55 °C. AFLP markers were also found in the cross between Gifu and MG-20. The AFLP and microsatellite markers were used to develop an F2 genetic map for this cross (Hayashi et al. 2001). Some of the microsatellite markers were included in both maps, and therefore, it was possible to see the colinearity between the two maps. With the substantial increase in the number of microsatellite markers, the high level of colinearity between the maps of *L. filicaulis* x Gifu and Gifu x MG-20 was clearly seen (Sandal et al. 2006). Only the genetic distances between markers varied in the two crosses. So, with the combination of the mapping information from the two crosses, most of the BAC and TAC sequences from the genome sequence efforts could be ordered on the six linkage groups of *L. japonicus*. The number of AFLP markers in the Gifu x MG-20 population has been substantially increased to 2,053 markers by Wang et al. (2008). In that case, the mapping was done on RILs of Gifu x MG-20 developed at Kazusa DNA Research Institute. This made it possible to look for codominant markers as the RILs are almost 100 % homozygotic. These RILs are available through LegumeBase (http://www.legumebase.brc.miyazaki-u.ac.jp/lotus/rilStrainListAction.do). RILs are developed by selfing up to S8 from F2 populations from the crosses. The genotypes of the different lines can be downloaded from http://www.kazusa.or.jp/lotus/RIline/index.html. These lines have been used for the QTL analysis of 13 agronomic traits by Gondo et al. (2007). There are so many differences between Gifu and MG-20 that MG-20 has been suggested as a new species, *Lotus miyakojimae* (Degtjareva et al. 2008).

As more and more BAC and TAC sequences became available, more microsatellite markers were developed and microsatellite markers became the chosen method for the mapping of, for example, symbiotic genes known from mutants. A set of such markers distributed over the six chromosomes (Table 3.1) can quickly give an approximate map position of the gene. Further, microsatellite markers from the region can then be used to narrow down the region of the gene. It should be noted that in the cross between Gifu and MG-20, the markers on the top of chromosome I and the bottom of chromosome II appear linked because of the translocation occurring between these ecotypes. Information about primer sequences, map positions, and the quality of the marker for hundreds of microsatellite markers can be found at http://www.kazusa.or.jp/lotus/markerdb_index.html. When the region is narrowed down to less than 200 kb in a

sequenced region, a candidate gene approach can be used with sequencing in the mutant of the candidate gene. Alternatively, one could do a whole genome sequence of the mutant and look in the mapped region for a mutated gene(s) (Liao et al. 2012). This approach also led to the identification of a candidate gene for *snf4* (Sandal and Andersen, unpublished). To prove that you found the right gene you need to sequence several alleles and/or complement the mutant with transformation with a wild-type gene construct in *Agrobacterium rhizogenes* or *Agrobacterium tumefaciens*. For most symbiotic mutants, the phenotype is determined by the root genotype, and therefore, hairy root transformation with *A. rhizogenes* can be used. For more information about the map-based cloning procedure, see Sandal et al. (2005).

In Fig. 3.1, the map positions of 49 symbiotic (*Sym*) genes from *L. japonicus* MG-20 are shown. In Fig. 3.2, the same map is shown for Gifu. Notice that the map position in Gifu is different on chromosome I and the bottom of chromosome II because of a translocation of the region corresponding to the top 10 cm of Gifu chromosome I to the bottom of MG-20 chromosome II. Most of these genes have now been isolated with map-based cloning. Information on some of the map positions can be found in Sandal et al. (2006). For *Sym105* see Hossain et al. (2006). Information about the gene isolation/map-based cloning projects can be found for many genes (Table 3.2). *Nin* was isolated by transposon tagging and *Nap* by a combination of retrotransposon tagging and mapping.

3.3 Physical Mapping

Fluorescent in situ hybridisation (FISH) analysis with BAC clones with known genetic map positions gave the link to the physical chromosomes (Pedrosa et al. 2002; Hayashi et al. 2001; Ohmido et al. 2010). Therefore, the chromosomes and linkage groups could be numbered with chromosome I as the largest Gifu chromosome corresponding to linkage group I and chromosome II as the second largest chromosome corresponding to linkage group II and so on. For the Gifu x MG-20 map, chromosomes I and II presented a problem. This was solved by a comparison to the genetic map developed from *L. filicaulis* x *L. japonicus* Gifu and FISH analysis. It turned out that there is a translocation between the lower part of Gifu chromosome II to the upper part of MG-20 chromosome I. On the other hand, *L. filicaulis* x *L. japonicus* Gifu has problems with low seed set, distorted segregation and several regions with suppressed recombination. Therefore, the alignment of the two genetic maps was of great help in ordering the markers and sequences from the BAC and TAC sequencing made by Kazusa DNA Research Institute.

3.4 Additional *Lotus* Species

L. burttii (Sz.-Borsos et al. 1972; Kawaguchi et al. 2005) originates from Pakistan. It is a good crossing partner for *L. japonicus* as it provides a better seed set, shorter generation times and has fewer chromosome regions with distorted segregation than *L. filicaulis* x *L. japonicus* Gifu. For some regions with suppression of recombination in Gifu x MG-20, such as the top of MG-20 chromosome I, the cross to *L. burttii* could be advantageous (Kawaguchi et al. 2005). *L. burttii* has a number of phenotypic differences to *L. japonicus,* and therefore, the developed RILs are of use in QTL detection for a number of traits. One example is a QTL for *Sinorhizobium fredii* nodulation (Sandal et al. 2012).

Evolutionarily, *L. krylovii* appears to be as close to *L. japonicus* as *L. burttii* and closer than *L. filicaulis* (Degtjareva et al. 2008). *L. krylovii* has been found in Russia and China, and an F2 population of *L. japonicus* Gifu x *L. krylovii* (Russian accession) has been obtained (Sandal, unpublished). This population could be used for mapping various phenotypic traits and be further developed to RILs.

The evolutionary relationship between the different Lotus species based on nrITS phylogeny was determined by Degtjareva et al. (2008). As part of the Lotassa collaboration, we have further investigated the level of polymorphism

Table 3.1 Selected *L. japonicus* microsatellite (simple sequence repeat, *SSR*) markers distributed over the six chromosomes (Chr.). The position is shown in centiMorgan (cM). The quality of the markers is shown on the right side (*g* good, *vg* very good)

			SSR pattern		Product size (bp)						
Chr.	Position	Clone	Motif	Repeat numbers in MG-20	MG-20	Gifu	Heteroduplex	Extra bands	Separation	Amplify	Quality
1	0.0	TM0027	CT	33	178	158	–	–	vg	g	g
1	8.8	TM0050	GT	11	149	165	180	–	g	g	g
1	31.4	TM0575	AT	23	124	110	–	–	g	g	g
1	50.1	TM0113	AAT	16	135	123	–	–	g	g	g
1	69.0	TM0295	CT	14	116	140	–	–	vg	g	g
2	15.6	TM0065	AT	12	114	124	130	–	g	g	g
2	29.0	TM0377	CT	32	140	122	–	–	g	f	g
2	42.8	TM0608	AT	18	153	141	–	–	g	g	g
2	61.4	TM0522	AT	24	173	153	–	–	vg	g	g
2	70.6	TM0002	CT	15	159	179	–	–	vg	g	vg
3	10.5	TM0436	ATG	8	142	160	–	–	g	g	g
3	30.0	TM0035	AAG	17	110	95	–	70	g	g	g
3	56.7	TM0049	AT	14	141	159	–	–	vg	g	g
3	84.8	TM0127	AAG	7 + 5	131	119	–	–	g	g	g
4	5.6	TM0256	AAAT	7	142	166	–	–	vg	g	g
4	19.3	TM0194	CT	17	122	156	–	–	vg	g	g
4	33.0	TM0030	AAT	18	142	121	–	100 (weak)	vg	g	vg
4	55.4	TM0046	CT	16	155	143	–	–	g	g	g
4	69.3	TM0617	CT	13	134	122	–	–	g	g	g
5	10.4	TM0596	CT	9	106	116	–	–	g	g	g
5	29.0	TM0048	AT	12	162	146	–	–	g	g	g
5	48.6	TM0428	GT	15	143	157	161	–	g	g	g
6	16.0	TM0302	CT	17	143	169	–	–	vg	g	g
6	38.5	TM0013	AT	20	206	216	–	176(Gifu), 166(MG-20)	g	g	g
6	48.6	TM0367	CT	27	172	152	–	–	vg	g	g
6	59.6	TM0336	CT	11(GT) + 14	164	180	–	–	g	g	g

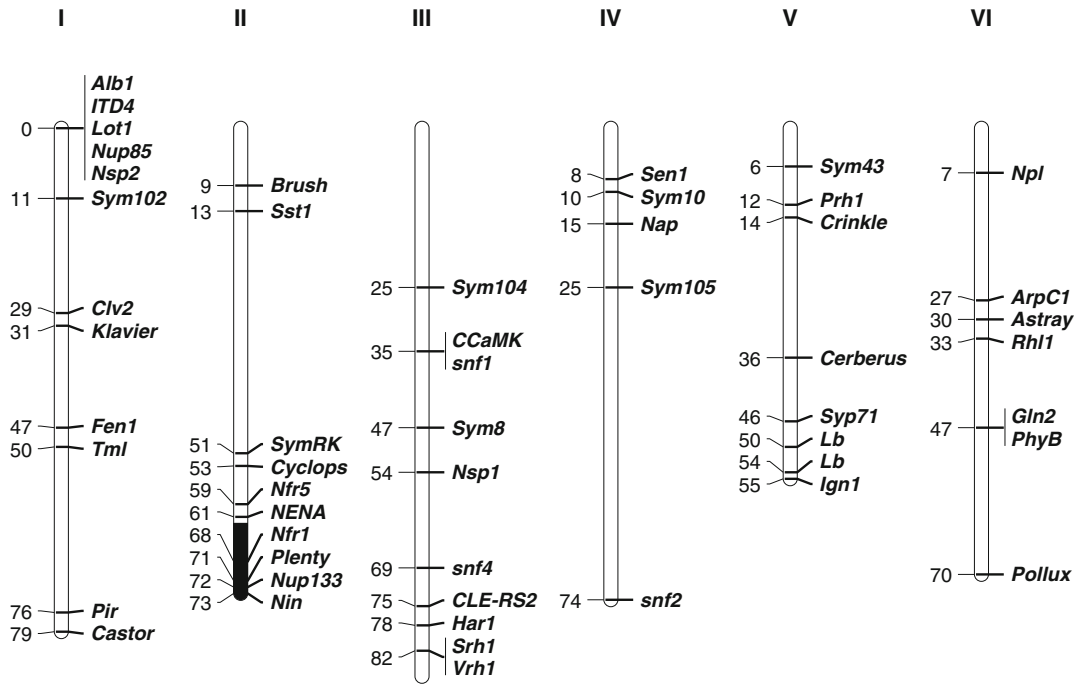


Fig. 3.1 Position of symbiotic (*Sym*) genes on the genetic map of *Lotus japonicus* MG-20. The *black* area shows the region that is translocated between MG-20 and Gifu

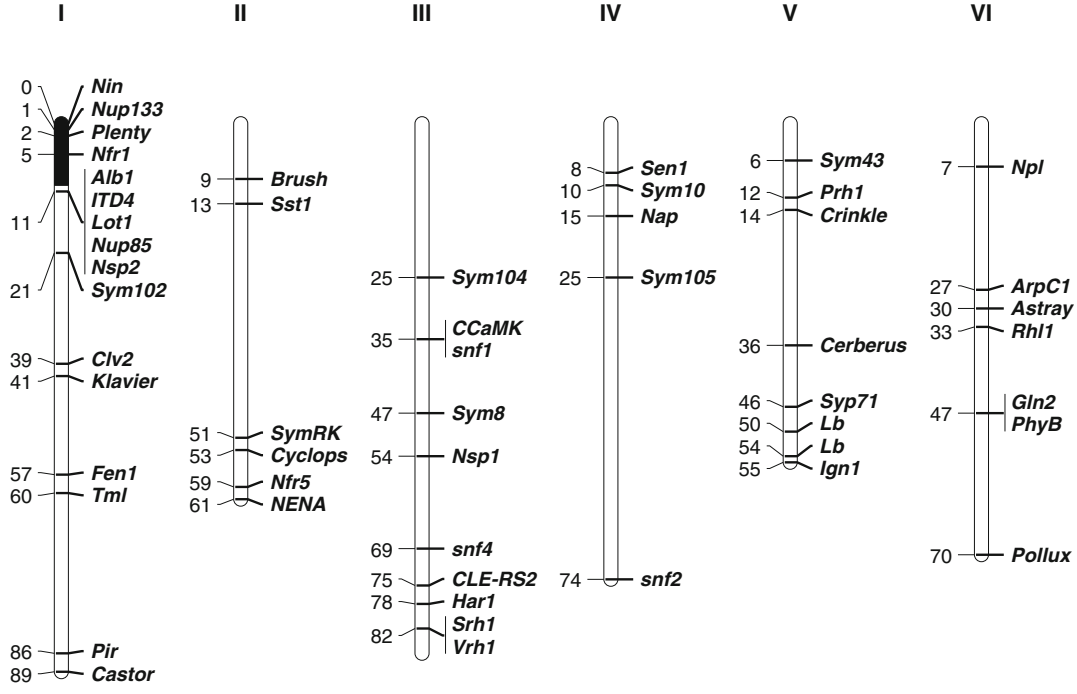


Fig. 3.2 Position of symbiotic (*Sym*) genes on the genetic map of *Lotus japonicus* Gifu. The *black* area shows the region that is translocated between MG-20 and Gifu

Table 3.2 Gene isolation/map-based cloning projects

Genes	References	Genes	References
ArpC1	Hossain et al. (2012)	*Nfr5*	Madsen et al. (2003)
Astray	Nishimura et al. (2002b)	*Nin*	Schauser et al. (1999)
Brush	Maekawa-Yoshikawa et al. (2009)	*Npl*	Xie et al. (2012)
Castor	Imaizumi-Anraku et al. (2005)	*Nsp1* and *Nsp2*	Heckmann et al. (2006), Murakami et al. (2006)
CCaMK/snf1	Tirichine et al. (2006a, b)	*Nup133*	Kanamori et al. (2006)
Cerberus	Yano et al. (2009)	*Nup85*	Saito et al. (2007)
CLE-RS2	Okamoto et al. (2009)	*PhyB*	Suzuki et al. (2011)
Clv2	Krusell et al. (2011)	*Plenty*	Yoshida et al. (2010)
Cyclops	Yano et al. (2008)	*Pollux*	Imaizumi-Anraku et al. (2005)
Fen1	Hakoyama et al. (2009)	*Rhl1*	Karas et al. (2009)
Gln2	Orea et al. (2002), García-Calderón et al. (2012)	*Sen1*	Hakoyama et al. (2012a)
Har1	Krusell et al. (2002), Nishimura et al. (2002a)	*snf2/hit1*	Tirichine et al. (2007), Murray et al. (2007)
Ign1	Kumagai et al. (2007)	*snf4*	Tirichine et al. (2006a, b)
ITD4	Lombardo et al. (2006)	*Srh1* and *sym105*	Karas et al. (2005), Hossain et al. (2006)
Klavier	Miyazawa et al. (2010)	*SymRK*	Stracke et al. (2002)
Lot1	Ooki et al. (2005)	*Sst1*	Krusell et al. (2005)
Nap and *Pir*	Yokota et al. (2009)	*Syp71*	Hakoyama et al. (2012b)
NENA	Groth et al. (2010)	*Tml*	Takahara et al. (2013)
Nfr1	Radutoiu et al. (2003)	*Vrh1*	Karas et al. (2005)

between *L. japonicus* and *L. tenuis* by sequencing 17 full length *L. tenuis* cDNA clones isolated by INIA/UACH in Chile. These cDNA clones represent genes that are differentially expressed during drought stress of *L. tenuis* root nodules (Tapia et al. 2010; Tapia et al. unpublished). The sequence identity between the coding regions of these genes in *L. tenuis* and *L. japonicus* is above 95 %, showing how closely these *Lotus* species are related. Again, this emphasises that the genetic information obtained in the model species *L. japonicus* is of use in other *Lotus* species such as *L. tenuis*.

Many of the microsatellite markers developed from *L. japonicus* Gifu x *L. japonicus* MG-20 have been tested on DNA of *Lotus corniculatus, L. tenuis, Lotus uliginosus* and *Lotus subbiflorus* using the following PCR conditions: 94° for 3 min followed by 40 rounds of 94° for 15 s, 54° for 20 s and 72° for 30 s, and finally 72° for 7 min before cooling (Fig. 3.3). In many cases, products are amplified from these species. Such molecular markers are potentially useful for breeding with the *Lotus* species used in agriculture. The map positions in *L. japonicus* are already known for these microsatellite markers (Sato et al. 2008).

In total, several hundred microsatellite markers have been developed by Kazusa DNA Research Institute and mapped to a specific position on the *Lotus* genome.

The cytogenetic map has been developed for *L. uliginosus* (Ferreira et al. 2012). It shows a high degree of macrosynteny to *L. japonicus,* but it was interrupted by a translocation involving chromosomes III and V, a new rearrangement for the genus.

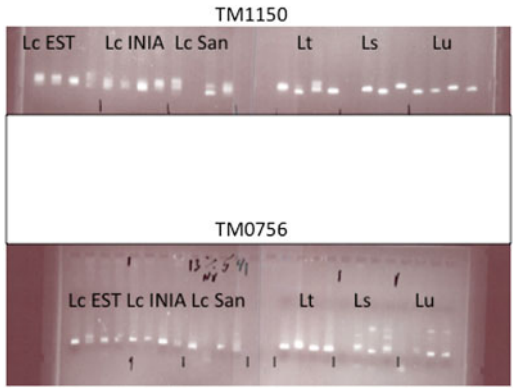


Fig. 3.3 *L. japonicus* microsatellite markers (TM1150 and TM0756) tested in individual plants of *L. corniculatus* (Lc), *L. tenuis* (Lt), *L. subbiflorus* (Ls) and *L. uliginosus* (Lu)

3.4.1 Conclusion About Microsatellite Markers

Most of the *L. japonicus* microsatellite markers amplify well in *L. corniculatus* and *L. tenuis* and several of them amplify well in *L. uliginosus* and *L. subbiflorus*. This analysis confirms that *L. japonicus* is closest to *L. burttii* and *L. filicaulis,* followed by *L. corniculatus* and *L. tenuis* and then finally *L. subbiflorus* and *L. uliginosus*. This is in agreement with the evolutionary tree suggested by Degtjareva et al. (2008).

The level of polymorphism is high between breeding lines and even within breeding lines. Some of these markers could therefore be used in breeding. Furthermore, these markers are useful for testing of crosses between different *Lotus* species and lines.

Additional information can be obtained by searching NCBI for sequences from the other *Lotus* species. From the comparison of *L. uliginosus* sequences available at NCBI to *L. japonicus*, a marker was developed giving a product that is 12 bp larger in *L. uliginosus* than in *L. japonicus* using the following primers:

ulig-EH380069-fw(49) GCTTCTTTCCTTGT TTGAACAT

ulig-EH380069-rv(49) AAGCAAAAGAAAA GTAAAATATGCAT

The accession number of the *L. uliginosus* sequence is EH380069. As this marker is developed from a sequence that is known from both species, it performs better than most markers developed only from the knowledge of the *L. japonicus* sequence such as the microsatellite markers. Such a marker can be used for testing crosses between *L. japonicus* and *L. uliginosus*.

3.5 Synteny

One of the ideas behind the model legume concept was to exploit synteny between models and crops to accelerate isolation and comparative characterisation of genes from the less characterised crop legumes. This approach requires that the target genome regions contain the same genes in approximately the same order in models and crops. In our legume anchor project, we have identified a number of well-conserved single copy genes from legumes based on EST sequences from soybean, *Medicago truncatula* and *L. japonicus*. A criterium for selection of a gene as a legume anchor marker candidate (Leg marker) is also that it is a single copy gene in *Arabidopsis*. Such genes are ideal for the analysis of homologous/orthologous genes from different legumes to look for the level of synteny.

Using Leg markers, we have anchor tagged loci covering 758 cm of the bean (*Phaseolus vulgaris*) genetic map and a set of 99 shared loci made it possible to compare this map with the genetic map of *Lotus.* All of the 11 bean linkage groups had non-interrupted regions of at least two markers also showing linkage in the *Lotus* genome, and in several of the bean linkage groups synteny spanned entire linkage groups. On *Pv*LG7, all nine Leg markers with a known position on the genetic linkage map of *Lotus* map to *Lj*LG5. On *Pv*LG11 11 out of 12 markers map to *Lj*LG3, and on *Pv*LG2 11 out of 20 markers map to *Lj*LG4, whereas the remaining nine markers map to *Lj*LG2. All in all, the legume anchor markers revealed a broad conservation of gene linkage (macrosynteny) between bean and *Lotus* (Hougaard et al. 2008).

This is in agreement with previous data based on genome sequences of *L. japonicus* and *M. truncatula* showing ten areas with extensive

synteny covering a major part of the total genomes of these species (Cannon et al. 2006). In addition, we used the legume anchor approach to look for synteny to the distantly related legume groundnut (*Arachis hypogeae*). In this case, we were also able to show synteny between *Arachis species, Lotus* and *M. truncatula*, but less pronounced compared to beans as expected (Bertioli et al. 2009).

It is therefore clear that synteny can be used as an analytical tool in both scientific and practical aspects of legume research. The substantial genome information from *Lotus*, *Medicago* and soybean can be used to help with genetic mapping and gene isolation in other legumes. Recently, a comparative mapping approach was used for a complex disease resistance gene locus in bean (Perrine et al. 2008) and the virus resistance gene Rsv4 in soybean (Hwang et al. 2009). In addition, a number of genes for basic traits, in addition to symbiotic genes, will be easier to isolate from model legumes, and afterwards the corresponding gene can be identified and followed in breeding programmes in crop legumes.

The conserved marker order (synteny) will be very high to the other *Lotus* species. We expect that markers that are linked in *L. japonicus* will also be linked in most of the other *Lotus* species.

Acknowledgments This work was supported by Danish National Research Foundation, by the European Union (Lotassa FP6-517617) and by the Kazusa DNA Research Institute Foundation.

References

Bertioli DJ, Moretzsohn MC, Madsen LH, Sandal N, Leal-Bertioli SC, Guimarães PM, Hougaard BK, Fredslund J, Schauser L, Nielsen AM, Sato S, Tabata S, Cannon SB, Stougaard J (2009) An analysis of synteny of *Arachis* with *Lotus* and *Medicago* sheds new light on the structure, stability and evolution of legume genomes. BMC Genomics 10:45

Cannon SB, Sterck L, Rombauts S, Sato S, Cheung F, Gouzy J, Wang X, Mudge J, Vasdewani J, Schiex T, Spannagl M, Monaghan E, Nicholson C, Humphray SJ, Schoof H, Mayer KF, Rogers J, Quétier F, Oldroyd GE, Debellé F, Cook DR, Retzel EF, Roe BA, Town CD, Tabata S, Van de Peer Y, Young ND (2006) Legume genome evolution viewed through the *Medicago truncatula* and *Lotus japonicus* genomes. Proc Natl Acad Sci USA 103:14959–14964

Degtjareva GV, Kramina TE, Sokoloff DD, Samigullin TH, Sandral G, Valiejo-Roman CM (2008) New data on nrITS phylogeny of *Lotus* (Leguminosae, Loteae). Wulfenia 15:35–49

de Nettancourt D, Grant WF (1964) Gene inheritance and linkage relationships in interspecific diploid hybrids closely related to *Lotus corniculatus* L. Can J Genet Cytol 6:277–287

Ferreira J, Mendes S, Dall'agnol M, Sandal N, Sato S, Pedrosa-Harand A (2012) Comparative analyses in *Lotus*: the cytogenetic map of *Lotus uliginosus* Schkuhr. Cytogenet Genome Res 137:42–49

García-Calderón M, Chiurazzi M, Espuny MR, Márquez AJ (2012) Photorespiratory metabolism and nodule function: behavior of *Lotus japonicus* mutants deficient in plastid glutamine synthetase. Mol Plant Microbe Interact 25:211–219

Gondo T, Sato S, Okumura K, Tabata S, Akashi R, Isobe S (2007) Quantitative trait locus analysis of multiple agronomic traits in the model legume *Lotus japonicus*. Genome 50:627–637

Groth M, Takeda N, Perry J, Uchida H, Dräxl S, Brachmann A, Sato S, Tabata S, Kawaguchi M, Wang TL, Parniske M (2010) NENA, a *Lotus japonicus* homolog of Sec13, is required for rhizodermal infection by arbuscular mycorrhiza fungi and rhizobia but dispensable for cortical endosymbiotic development. Plant Cell 22:2509–2526

Hakoyama T, Niimi K, Watanabe H, Tabata R, Matsubara J, Sato S, Nakamura Y, Tabata S, Jichun L, Matsumoto T, Tatsumi K, Nomura M, Tajima S, Ishizaka M, Yano K, Imaizumi-Anraku H, Kawaguchi M, Kouchi H, Suganuma N (2009) Host plant genome overcomes the lack of a bacterial gene for symbiotic nitrogen fixation. Nature 462:514–517

Hakoyama T, Niimi K, Yamamoto T, Isobe S, Sato S, Nakamura Y, Tabata S, Kumagai H, Umehara Y, Brossuleit K, Petersen TR, Sandal N, Stougaard J, Udvardi MK, Tamaoki M, Kawaguchi M, Kouchi H, Suganuma N (2012a) The integral membrane protein SEN1 is required for symbiotic nitrogen fixation in *Lotus japonicus* nodules. Plant Cell Physiol 53:225–236

Hakoyama T, Oi R, Hazuma K, Suga E, Adachi Y, Kobayashi M, Akai R, Sato S, Fukai E, Tabata S, Shibata S, Wu GJ, Hase Y, Tanaka A, Kawaguchi M, Kouchi H, Umehara Y, Suganuma N (2012b) The SNARE protein SYP71 expressed in vascular tissues is involved in symbiotic nitrogen fixation in *Lotus japonicus* nodules. Plant Physiol 160:897–905

Handberg K, Stougaard J (1992) *Lotus japonicus*, an autogamous, diploid legume species for classical and molecular genetics. Plant J 2:487–496

Hayashi M, Miyahara A, Sato S, Kato T, Yoshikawa M, Taketa M, Hayashi M, Pedrosa A, Onda R, Imaizumi-Anraku H, Bachmair A, Sandal N, Stougaard J, Murooka Y, Tabata S, Kawasaki S, Kawaguchi M, Harada K (2001) Construction of a genetic linkage

map of the model legume *Lotus japonicus* using an intraspecific F2 population. DNA Res 8:301–310

Heckmann AB, Lombardo F, Miwa H, Perry JA, Bunnewell S, Parniske M, Wang TL, Downie JA (2006) *Lotus japonicus* nodulation requires two GRAS domain regulators, one of which is functionally conserved in a non-legume. Plant Physiol 142:1739–1750

Hossain MS, Umehara Y, Kouchi H (2006) A novel fix-symbiotic mutant of *Lotus japonicus*, *Ljsym105*, shows impaired development and premature deterioration of nodule infected cells and symbiosomes. Mol Plant Microbe Interact 19:780–788

Hossain MS, Liao J, James EK, Sato S, Tabata S, Jurkiewicz A, Madsen LH, Stougaard J, Ross L, Szczyglowski K (2012) *Lotus japonicus* ARPC1 is required for rhizobial infection. Plant Physiol 160:917–928

Hougaard BK, Madsen LH, Sandal N, de Carvalho Moretzsohn M, Fredslund J, Schauser L, Nielsen AM, Rohde T, Sato S, Tabata S, Bertioli DJ, Stougaard J (2008) Legume anchor markers link syntenic regions between Phaseolus vulgaris, Lotus japonicus, Medicago truncatula and Arachis. Genetics 179:2299–2312

Hwang TY, Moon JK, Yu S, Yang K, Mohankumar S, Yu YH, Lee YH, Kim HS, Kim HM, Maroof MA, Jeong SC (2009) Application of comparative genomics in developing molecular markers tightly linked to the virus resistance gene Rsv4 in soybean. Genome 49:380–388

Imaizumi-Anraku H, Takeda N, Charpentier M, Perry J, Miwa H, Umehara Y, Kouchi H, Murakami Y, Mulder L, Vickers K, Pike J, Downie JA, Wang T, Sato S, Asamizu E, Tabata S, Yoshikawa M, Murooka Y, Wu GJ, Kawaguchi M, Kawasaki S, Parniske M, Hayashi M (2005) Plastid proteins crucial for symbiotic fungal and bacterial entry into plant roots. Nature 433:527–531

Jiang Q, Gresshoff P (1997) Classical and molecular genetics of the model legume *Lotus japonicus*. Mol Plant-Microbe Interact 10:59–68

Kanamori N, Madsen LH, Radutoiu S, Frantescu M, Quistgaard EMH, Miwa H, Downie JA, James EK, Felle HH, Haaning LL, Jensen TH, Sato S, Nakamura Y, Tabata S, Sandal N, Stougaard J (2006) A nucleoporin is required for induction of Ca2+ spiking in legume nodule development and essential for rhizobial and fungal symbiosis. Proc Natl Acad Sci USA 103:359–364

Karas B, Murray J, Gorzelak M, Smith A, Sato S, Tabata S, Szczyglowski K (2005) Invasion of *Lotus japonicus root hairless 1* by *Mesorhizobium loti* involves the nodulation factor-dependent induction of root hairs. Plant Physiol 137:1331–1344

Karas B, Amyot L, Johansen C, Sato S, Tabata S, Kawaguchi M, Szczyglowski K (2009) Conservation of *Lotus* and *Arabidopsis* basic helix-loop-helix proteins reveals new players in root hair development. Plant Physiol 151:1175–1185

Kawaguchi M (2000) *Lotus japonicus* 'Miyakojima' MG-20: an early-flowering accession suitable for indoor handling. J Plant Res 133:507–509

Kawaguchi M, Motomura T, Imaizumi-Anraku H, Akao S, Kawasaki S (2001) Providing the basis for genomics in *Lotus japonicus*: the accessions Miyakojima and Gifu are appropriate crossing partners for genetic analyses. Mol Genet Genomics 266:157–166

Kawaguchi M, Pedrosa-Harand A, Yano K, Hayashi M, Murooka Y, Saito K, Nagata T, Namai K, Nishida H, Shibata D, Sato S, Tabata S, Hayashi M, Harada K, Sandal N, Stougaard J, Bachmair A, Grant WF (2005) *Lotus burttii* takes a position of the third corner in the *Lotus* molecular genetics triangle. DNA Res 12:69–77

Krusell L, Madsen LH, Sato S, Aubert G, Genua A, Szczyglowski K, Duc G, Kaneko T, Tabata S, de Bruijn F, Pajuelo E, Sandal N, Stougaard J (2002) Shoot control of root development and nodulation is mediated by a receptor-like kinase. Nature 420:422–426

Krusell L, Krause K, Ott T, Desbrosses G, Krämer U, Sato S, Nakamura Y, Tabata S, James EK, Sandal N, Stougaard J, Kawaguchi M, Miyamoto A, Suganuma N, Udvardi MK (2005) The sulfate transporter SST1 is crucial for symbiotic nitrogen fixation in *Lotus japonicus* root nodules. Plant Cell 17:1625–1636

Krusell L, Sato N, Fukuhara I, Koch BE, Grossmann C, Okamoto S, Oka-Kira E, Otsubo Y, Aubert G, Nakagawa T, Sato S, Tabata S, Duc G, Parniske M, Wang TL, Kawaguchi M, Stougaard J (2011) The *Clavata2* genes of pea and *Lotus japonicus* affect autoregulation of nodulation. Plant J 65:861–871

Kumagai H, Hakoyama T, Umehara Y, Sato S, Kaneko T, Tabata S, Kouchi H (2007) A novel ankyrin-repeat membrane protein, IGN1, is required for persistence of nitrogen-fixing symbiosis in root nodules of *Lotus japonicus*. Plant Physiol 143:1293–1305

Liao J, Singh S, Hossain MS, Andersen SU, Ross L, Bonetta D, Zhou Y, Sato S, Tabata S, Stougaard J, Szczyglowski K, Parniske M (2012) Negative regulation of CCaMK is essential for symbiotic infection. Plant J 72:572–584

Lombardo F, Heckmann AB, Miwa H, Perry JA, Yano K, Hayashi M, Parniske M, Wang TL, Downie JA (2006) Identification of symbiotically defective mutants of *Lotus japonicus* affected in infection thread growth. Mol Plant Microbe Interact 19:1444–1450

Madsen EB, Madsen LH, Radutoiu S, Olbryt M, Rakwalska M, Szczyglowski K, Sato S, Kaneko T, Tabata S, Sandal N, Stougaard J (2003) A receptor kinase gene of the LysM type is involved in legume perception of rhizobial signals. Nature 425:637–640

Maekawa-Yoshikawa M, Müller J, Takeda N, Maekawa T, Sato S, Tabata S, Perry J, Wang TL, Groth M, Brachmann A, Parniske M (2009) The temperature-sensitive *brush* mutant of the legume *Lotus japonicus* reveals a link between root development and nodule infection by Rhizobia. Plant Physiol 149:1785–1796

Miyazawa H, Oka-Kira E, Sato N, Takahashi H, Wu GJ, Sato S, Hayashi M, Betsuyaku S, Nakazono M, Tabata S, Harada K, Sawa S, Fukuda H, Kawaguchi M (2010) The receptor-like kinase KLAVIER mediates systemic regulation of nodulation and non-symbiotic shoot development in *Lotus japonicus*. Development 137:4317–4325

Murakami Y, Miwa H, Imaizumi-Anraku H, Kouchi H, Downie JA, Kawaguchi M, Kawasaki S (2006) Positional cloning identifies *Lotus japonicus* NSP2, a putative transcription factor of the GRAS family, required for NIN and ENOD40 gene expression in nodule initiation. DNA Res 13:255–265

Murray JD, Karas BJ, Sato S, Tabata S, Amyot L, Szczyglowski K (2007) A cytokinin perception mutant colonized by *Rhizobium* in the absence of nodule organogenesis. Science 315:101–104

Nishimura R, Hayashi M, Wu GJ, Kouchi H, Imaizumi-Anraku H, Murakami Y, Kawasaki S, Akao S, Ohmori M, Nagasawa M, Harada K, Kawaguchi M (2002a) HAR1 mediates systemic regulation of symbiotic organ development. Nature 420:426–429

Nishimura R, Ohmori M, Fujita H, Kawaguchi M (2002b) A *Lotus* basic leucine zipper protein with a RING-finger motif negatively regulates the developmental program of nodulation. Proc Natl Acad Sci USA 99:15206–15210

Ooki Y, Banba M, Yano K, Maruya J, Sato S, Tabata S, Saeki K, Hayashi M, Kawaguchi M, Izui K, Hata S (2005) Characterization of the *Lotus japonicus* symbiotic mutant *lot1* that shows a reduced nodule number and distorted trichomes. Plant Physiol 137:1261–1271

Ohmido N, Ishimaru A, Kato S, Sato S, Tabata S, Fukui K (2010) Integration of cytogenetic and genetic linkage maps of *Lotus japonicus*, a model plant for legumes. Chromosome Res 18:287–299

Okamoto S, Ohnishi E, Sato S, Takahashi H, Nakazono M, Tabata S, Kawaguchi M (2009) Nod factor/nitrate-induced *CLE* genes that drive HAR1-mediated systemic regulation of nodulation. Plant Cell Physiol 50:167–177

Orea A, Pajuelo P, Pajuelo E, Quidiello C, Romero JM, Márquez AJ (2002) Isolation of photorespiratory mutants from *Lotus japonicus* deficient in glutamine synthetase. Physiol Plant 115:352–361

Pedrosa A, Sandal N, Stougaard J, Schweizer D, Bachmair A (2002) Chromosomal map of the model legume *Lotus japonicus*. Genetics 161:1661–1672

Perrine D, Sevignac M, Thareau V, Catillon Y, Kami J, Gepts P, Langin T, Geffroy V (2008) BAC end sequences corresponding to the B4 resistance gene cluster in common bean: a resource for markers and synteny analyses. Mol. Genet. Genomics 280:521–533

Radutoiu S, Madsen LH, Madsen EB, Felle HH, Umehara Y, Grønlund M, Sato S, Nakamura Y, Tabata S, Sandal N, Stougaard J (2003) Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases. Nature 425:585–592

Saito K, Yoshikawa M, Yano K, Miwa H, Uchida H, Asamizu E, Sato S, Tabata S, Imaizumi-Anraku H, Umehara Y, Kouchi H, Murooka Y, Szczyglowski K, Downie JA, Parniske M, Hayashi M, Kawaguchi M (2007) NUCLEOPORIN85 is required for calcium spiking, fungal and bacterial symbioses, and seed production in *Lotus japonicus*. Plant Cell 19:610–624

Sandal N, Krause K, Frantescu M, Radutoiu S, Krusell L, Madsen LH, Kanamori N, Madsen E, Tirichine L, Stougaard J (2005) Mapping and map-based cloning in *Lotus japonicus*. In: Márquez AJ, Stougaard J, Udvardi M, Parniske M, Spaink H, Saalbach G, Webb J, Chiurazzi M, Márquez AJ (eds) *Lotus japonicus* handbook. Springer, Netherlands, pp 217–232

Sandal N, Krusell L, Radutoiu S, Olbryt M, Pedrosa A, Stracke S, Sato S, Kato T, Tabata S, Parniske M, Bachmair A, Ketelsen T, Stougaard J (2002) A genetic linkage map of the model legume *Lotus japonicus* and strategies for fast mapping of new loci. Genetics 161:1673–1683

Sandal N, Petersen TR, Murray J, Umehara Y, Karas B, Yano K, Kumagai H, Yoshikawa M, Saito K, Hayashi M, Murakami Y, Wang X, Hakoyama T, Imaizumi-Anraku H, Sato S, Kato T, Chen W, Hossain MS, Shibata S, Wang TL, Yokota K, Larsen K, Kanamori N, Madsen E, Radutoiu S, Madsen LH, Radu TG, Krusell L, Ooki Y, Banba M, Betti M, Rispail N, Skøt L, Tuck E, Perry J, Yoshida S, Vickers K, Pike J, Mulder L, Charpentier M, Müller J, Ohtomo R, Kojima T, Ando S, Marquez AJ, Gresshoff PM, Harada K, Webb J, Hata S, Suganuma N, Kouchi H, Kawasaki S, Tabata S, Hayashi M, Parniske M, Szczyglowski K, Kawaguchi M, Stougaard J (2006) Genetics of symbiosis in *Lotus japonicus*: recombinant inbred lines, comparative genetic maps, and map position of 35 symbiotic loci. Mol Plant Microbe Interact 19:80–91

Sandal N, Jin H, Rodriguez-Navarro DN, Temprano F, Cvitanich C, Brachmann A, Sato S, Kawaguchi M, Tabata S, Parniske M, Ruiz-Sainz JE, Andersen SU, Stougaard J (2012) A set of *Lotus japonicus* Gifu x *Lotus burttii* recombinant inbred lines facilitate map-based cloning and QTL mapping. DNA Res. 19:317–323

Sato S, Nakamura Y, Kaneko T, Asamizu E, Kato T, Nakao M, Sasamoto S, Watanabe A, Ono A, Kawashima K, Fujishiro T, Katoh M, Kohara M, Kishida Y, Minami C, Nakayama S, Nakazaki N, Shimizu Y, Shinpo S, Takahashi C, Wada T, Yamada M, Ohmido N, Hayashi M, Fukui K, Baba T, Nakamichi T, Mori H, Tabata S (2008) Genome structure of the legume. Lotus Japonicus DNA Res 15:227–239

Schauser L, Roussis A, Stiller J, Stougaard J (1999) A plant regulator controlling development of symbiotic root nodules. Nature 402:191–195

Somaroo BH, Grant WF (1971) Interspecific hybridization between diploid species of *Lotus* (Leguminosae). Genetica 42:353–367

Stougaard J, Beuselinck PR (1996) Registration of GIFU B-129-S *Lotus japonicus* germplasm. Crop Sci 36:476

Stracke S, Kistner C, Yoshida S, Mulder L, Sato S, Kaneko T, Tabata S, Sandal N, Stougaard J,

Szczyglowski K, Parniske M (2002) A plant receptor-like kinase required for both bacterial and fungal symbiosis. Nature 417:959–962

Suzuki A, Suriyagoda L, Shigeyama T, Tominaga A, Sasaki M, Hiratsuka Y, Yoshinaga A, Arima S, Agarie S, Sakai T, Inada S, Jikumaru Y, Kamiya Y, Uchiumi T, Abe M, Hashiguchi M, Akashi R, Sato S, Kaneko T, Tabata S, Hirsch AM (2011) *Lotus japonicus* nodulation is photomorphogenetically controlled by sensing the red/far red (R/FR) ratio through jasmonic acid (JA) signaling. Proc Natl Acad Sci USA 108:16837–16842

Sz.-Borsos O, Somaroo BH, Grant WF (1972) A new diploid species of *Lotus* (Leguminosae) in Pakistan. Can J Bot 50:1865–1870

Takahara M, Magori S, Soyano T, Okamoto S, Yoshida C, Yano K, Sato S, Tabata S, Yamaguchi K, Shigenobu S, Takeda N, Suzaki T, Kawaguchi M (2013) Too much love, a novel Kelch repeat-containing F-box protein, functions in the long-distance regulation of the legume-*Rhizobium* symbiosis. Plant Cell Physiol 22:259–268

Tapia G, Morales-Quintana L, Inostroza L, Acuna H (2010) Molecular characterization of *Ltchi7*, a gene encoding a class III endochitinase induced by drought stress in *Lotus* spp. Plant Biol 13:69–77

Tirichine L, James EK, Sandal N, Stougaard J (2006a) Spontaneous root-nodule formation in the model legume *Lotus japonicus*: a novel class of mutants nodulates in the absence of rhizobia. Mol Plant Microbe Interact 19:373–382

Tirichine L, Imaizumi-Anraku H, Yoshida S, Murakami Y, Madsen LH, Miwa H, Nakagawa T, Sandal N, Albrektsen AS, Kawaguchi M, Downie A, Sato S, Tabata S, Kouchi H, Parniske M, Kawasaki S, Stougaard J (2006b) Deregulation of a Ca2+/calmodulin-dependent kinase leads to spontaneous nodule development. Nature 441:1153–1156

Tirichine L, Sandal N, Madsen LH, Radutoiu S, Albrektsen AS, Sato S, Asamizu E, Tabata S, Stougaard J (2007) A gain-of-function mutation in a cytokinin receptor triggers spontaneous root nodule organogenesis. Science 315:104–107

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414

Wang X, Sato S, Tabata S, Kawasaki S (2008) A high-density linkage map of *Lotus japonicus* based on AFLP and SSR markers. DNA Res 15:323–332

Xie F, Murray JD, Kim J, Heckmann AB, Edwards A, Oldroyd GE, Downie JA (2012) Legume pectate lyase required for root infection by rhizobia. Proc Natl Acad Sci USA 109:633–638

Yano K, Yoshida S, Müller J, Singh S, Banba M, Vickers K, Markmann K, White C, Schuller B, Sato S, Asamizu E, Tabata S, Murooka Y, Perry J, Wang TL, Kawaguchi M, Imaizumi-Anraku H, Hayashi M, Parniske M (2008) CYCLOPS, a mediator of symbiotic intracellular accommodation. Proc Natl Acad Sci USA 105:20540–20545

Yano K, Shibata S, Chen WL, Sato S, Kaneko T, Jurkiewicz A, Sandal N, Banba M, Imaizumi-Anraku H, Kojima T, Ohtomo R, Szczyglowski K, Stougaard J, Tabata S, Hayashi M, Kouchi H, Umehara Y (2009) CERBERUS, a novel U-box protein containing WD-40 repeats, is required for formation of the infection thread and nodule development in the legume-*Rhizobium* symbiosis. Plant J 60:168–180

Yokota K, Fukai E, Madsen LH, Jurkiewicz A, Rueda P, Radutoiu S, Held M, Hossain MS, Szczyglowski K, Morieri G, Oldroyd GE, Downie JA, Nielsen MW, Rusek AM, Sato S, Tabata S, James EK, Oyaizu H, Sandal N, Stougaard J (2009) Rearrangement of actin cytoskeleton mediates invasion of *Lotus japonicus* roots by *Mesorhizobium loti*. Plant Cell 21:267–284

Yoshida C, Funayama-Noguchi S, Kawaguchi M (2010) Plenty, a novel hypernodulation mutant in *Lotus japonicus*. Plant Cell Physiol 51:1425–1435

Part II
Genomics and Functional Genomics of Lotus and the Microsymbionts

Genome Sequencing

4

Shusei Sato and Stig U. Andersen

Abstract

The current *Lotus japonicus* reference genome sequence is based on a hybrid assembly of Sanger TAC/BAC, Sanger shotgun and Illumina shotgun sequencing data generated from the Miyakojima-MG20 accession. It covers nearly all expressed *L. japonicus* genes and has been annotated mainly based on transcriptional evidence. Analysis of repetitive sequences suggests that they are underrepresented in the reference assembly, reflecting an enrichment of gene-rich regions in the current assembly. Characterization of Lotus natural variation by resequencing of *L. japonicus* accessions and diploid Lotus species is currently ongoing, facilitated by the MG20 reference sequence.

4.1 Sequencing and Assembly of Gene-rich Regions

Large scale sequencing of the *Lotus japonicus* genome was launched in 2001 at the Kazusa DNA Research Institute. Miayakojima-MG20 was chosen as the reference accession because of its favorable features, including very early flowering and robust growth under both growth chamber and greenhouse conditions. In addition, MG20 is polymorphic with respect to the widely used experimental accession Gifu B-129, and offspring from MG20 x Gifu crosses can be used to generate genetic linkage maps, which are essential for contig ordering and assignment to linkage groups (chromosomes). By flow cytometry analysis, the genome size of MG20 was estimated to be 472 Mb (Ito et al. 2000), and the preliminary genomic sequencing revealed the presence of large amounts of various types of repetitive sequences in the *L. japonicus* genome, estimated to make up at least 20 % of the total genome length. Aiming at a cost-efficient characterization of gene-rich regions, two independent approaches for focused sequencing of gene-rich regions were implemented: Clone-by-clone sequencing from seed points and shotgun sequencing of selected genomic regions.

For the clone-by-clone approach, genome libraries were constructed using mainly transformation-competent artificial chromosome (TAC)

S. Sato (✉)
Graduate School of Life Sciences, Tohoku University, Katahira 2-1-1, Aoba-Ku, Sendai 980-8577, Japan
e-mail: shuseis@ige.tohoku.ac.jp

S.U. Andersen
Centre for Carbohydrate Recognition and Signalling, Department of Molecular Biology and Genetics, Aarhus University, Gustav Wieds Vej 10, 8000 Aarhus C, Denmark

S. Tabata and J. Stougaard (eds.), *The Lotus japonicus Genome*, Compendium of Plant Genomes,
DOI 10.1007/978-3-662-44270-8_4,

vectors. The average insert sizes of the main TAC libraries were around 100 kb. In addition to the main libraries, TAC libraries with shorter inserts and BAC libraries were constructed as sublibraries. The total of these libraries was 20 times haploid genome equivalents. Three-dimensional DNA pools for PCR screening were prepared from the main libraries and BAC sublibraries. End sequences of these clones were accumulated to facilitate walking clone selection from the seed sequences.

Aiming to identify seed points for sequencing as well as to generate a catalogue of expressed genomic regions, a large scale analysis of expressed sequence tags (ESTs) was carried out in the initial phase of the project. Based on the EST information, TAC/BAC clones were selected from the genomic libraries as seed points for the clone-by-clone approach. The nucleotide sequence of each clone was determined using a the shotgun strategy with three to five times redundancy, and the sequenced clones were anchored onto six chromosomes using a total of 788 microsatellite markers derived from the clone sequences. Since there are chromosomal inversions presumably caused by translocation between the top arm of MG-20 chromosome 1 and the bottom arm of B-129 chromosome 2 the genetic distance of these regions is nearly zero, with limited recombination. Therefore, the order of the clones placed within the chromosomal inversion regions was assigned based on the corresponding markers on the genetic linkage maps of *L. filicaulis* x *L. japonicus* Gifu and *L. burttii* x *L. japonicus* Gifu as well as the results of fluorescent in situ hybridisation (FISH) analysis. The constructed pseudomolecules represented the physical form of MG20 genome. In parallel with the clone-by-clone approach, shotgun sequencing of selected genomic regions was used to accumulate draft sequence information for the remaining gene-rich regions (Sato et al. 2008).

The total length of the v. 1.0 assembly was 315 Mbp, consisting of 594 anchored supercontigs with a total length of 130 Mbp and 110,000 unanchored contigs with a total length of 184 Mbp. While this assembly corresponded to 67 % of the reported *L. japonicus* genome (472 Mb) (Ito et al. 2000), it was estimated that it covered ~91 % of the gene space because 11,404 out of 12,485 tentative consensus (TC) sequences of the *L. japonicus* Gene Index provided by the Gene Index Project [http://compbio.dfci.harvard.edu/tgi/plant.html] could be mapped to the v. 1.0 assembly.

In 2010, an updated genome sequence, v. 2.5, was constructed by adding genome sequence information from 460 TAC/BAC clones analyzed after the release of v. 1.0, increasing the total length of anchored contigs to 195 Mbp. This Sanger-based sequence information is available through the web database at http://www.kazusa.or.jp/lotus/build2.5.

4.2 Whole-genome Sequencing

The advent of high-throughput and low cost next-generation sequencing (NGS) technology led to a general change in sequencing strategy from clone-by-clone Sanger sequencing to whole-genome shotgun approaches. In 2009, the NGS shotgun approach was implemented in the *L. japonicus* genome sequencing project using two emerging NGS platforms, 454 GS FLX (Margulies et al. 2005) and Illumina (Bentley et al. 2008). Since there was no assembly program available that could carry out hybrid assembly of short NGS reads and longer Sanger-based contigs, a step-by-step approach for combined assembly was used (Fig. 4.1).

Using this hybrid assembly approach, a total of 132 scaffolds covering 232 Mbp of the genome were aligned to the six *L. japonicus* chromosomes. Thus, in v 2.5, the total number of scaffolds has been decreased to one-fifth of the previous 646 scaffolds, and the total length of anchored contigs has been increased by 20 % of the original 195 Mbp. The remaining 23,572 unanchored contigs, corresponding to 162 Mbp were assigned to a virtual chromosome 0 (Table 4.1). Version 3.0 of the Lotus genome (Sato et al. 2014) is thus comparable to other published legume reference genomes, including *Medicago truncatula* v. 3.5 and soybean (Schmutz et al. 2010; Young et al. 2011). The gene

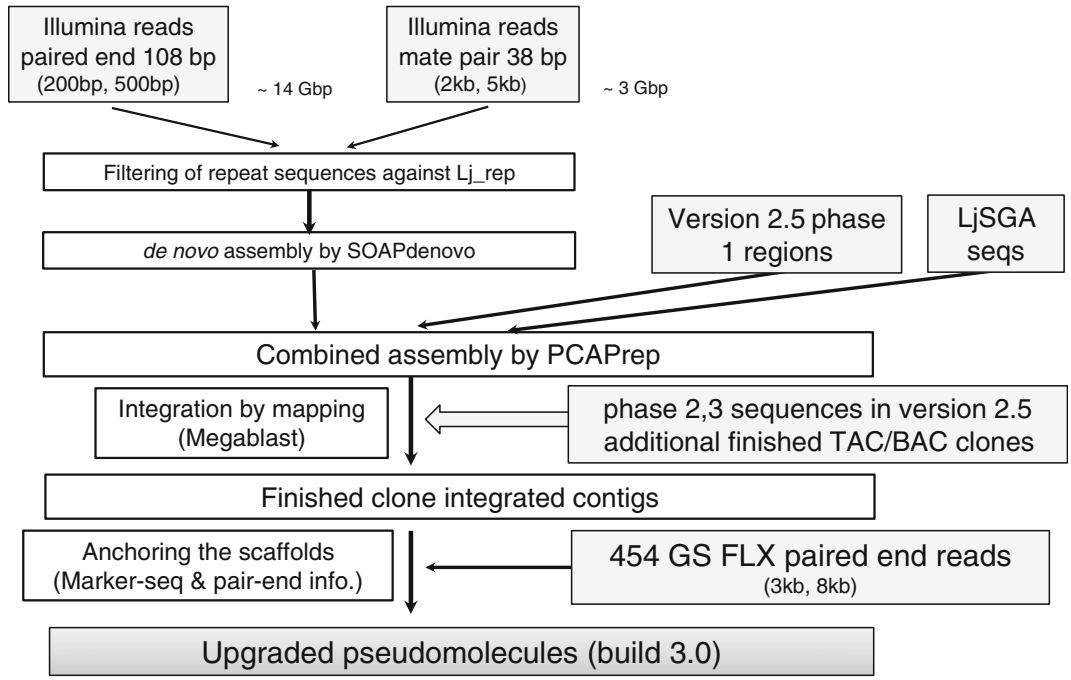


Fig. 4.1 Summary of the hybrid assembly strategy used for *L. japonicus* genome sequence version 3.0

Table 4.1 Assembly statistics of *L. japonicus* genome sequences in version 1.0, 2.5 and 3.0

	Version 1.0	Version 2.5	Version 3.0
Total length of scaffolds (bases)	315,073,275	296,886,266	393,918,449
Anchored scaffolds			
Number of anchored scaffolds	594	647	132
Length of anchored scaffolds (bases)	130,251,279	184,542,525	231,615,632
Unanchored scaffolds			
Number of unanchored scaffolds	110,346	67,902	23,572
Length of unanchored scaffolds (bases)	184,821,996	112,343,741	162,302,817

space coverage of v. 3.0 is ~98 % based on the placement of 57,916 de novo assembled RNA-seq contigs and 25,694 tentative contigs (TCs, *L. japonicus* Gene Index v. 6.0) (Sato et al. 2014).

4.3 Gene Annotation

In early versions of the genome assembly, 1.0 and 2.5, annotation relied on ab initio predictions and homology to available protein sequences. In version 3.0 a hierarchical approach, prioritizing transcription evidence was used instead (Sato et al. 2014). The differences in annotation approaches did not greatly influence the number of annotated protein coding genes, which remained around 40,000. However, the number of annotated amino acids increased by 20 % between v. 1.0 and 3.0, and this increase was reflected by an increase in the number of assigned peptides from proteomics data (Sato et al. 2014).

4.4 Repetitive Sequences

Since the highly repetitive sequences in the *L. japonicus* genome tend to be excluded from scaffolds/contigs during NGS and hybrid assembly, these regions of the genome could be less represented in the latest assembly. Therefore, less biased TAC end sequences were used to survey the highly repetitive sequences in the entire genome of *L. japonicus*. Based on analysis of 37,000 TAC end sequences, 34 types of highly repetitive sequences were identified (Sato et al. 2008). These included sequences localized to highly condensed heterochromatic regions of the genome known as chromosome knobs and centromere-associated sequences (Pedrosa et al. 2002; Sato et al. 2008).

In parallel, a combined computer-assisted and experimental analysis of transposable elements (TEs) on the v. 1.0 assembly was carried out (Holligan et al. 2006). The analysis revealed that the *L. japonicus* genome is rich in Pack-MULEs, nonautonomous MULEs (*Mutator*-like elements), and Copia-like elements with an additional ORF (Holligan et al. 2006). Transposon display indicated a significant level of insertion polymorphism between Miyakojima MG-20 and Gifu B-129, suggesting recent element activity. Indeed, the transposition activity of two gypsy-like retrotranspons, *LORE1* and *LORE2*, was confirmed (Fukai et al. 2010; Madsen et al. 2005). Later, the construction of a large scale *L. japonicus* insertion mutant collection was initiated based on the transposition activity of *LORE1* (Fukai et al. 2012; Urbanski et al. 2012, 2013).

As a further analysis of repetitive sequences RECON (Bao and Eddy, 2002) was used to identify dispersed repetitive sequences in v. 3.0. As a result, a variety of repeat elements including class I and class II TE subfamilies and those that are difficult to classify into known subfamilies were stored in repetitive sequence libraries in addition to the highly repetitive sequences identified by TAC end sequence analysis. Using RepeatMasker analysis based on the *L. japonicus* repetitive sequence libraries, ~32 % of the *L. japonicus* genome sequence (version 3.0) was classified as repetitive (Table 4.2). About half of the entire length of the repetitive sequences identified was Class I TEs (retrotransposons), while the nonautonomous class II TEs including Pack-MULEs and MITEs were the most abundant with nearly 130,000 copies. A substantial portion of these nonautonomous class II TEs were found in introns and UTRs. A short insert size class I retroelement, a member of the short-interspersed nucleotide elements (SINE), was also preferentially observed in introns and the 3'UTRs (Fawcett et al. 2006). The fraction of

Table 4.2 Repetitive sequences in *L. japonicus* genome sequence v. 3.0

Repeat type	Number of elements	Coverage (kb)	Percentage of sequence (%)
Class I			
SINEs	294	36.9	0.01
LINEs	11,650	4105.7	1.04
LTR: Copia	56,317	30773.5	7.80
LTR: Gypsy	36,819	26773.3	6.79
LTR: other	4,915	2410.2	0.61
Total class I	109,995	64099.6	16.25
Class II			
Autonomous class II	57,146	15214.7	3.86
Nonautonomous class II	129,182	30833.1	7.81
Total class II	186,328	46047.8	11.67
Short tandem repeats	650	612.7	0.16
Unclassified	68,567	17327.5	4.39

short tandem repeats in v. 3.0 was lower than the fraction in TAC end sequences, reflecting that the highly repetitive heterochromatic regions are underrepresented in the v. 3.0 genome sequence.

4.5 Resequencing of Wild Accessions and Species

Polymorphism information from resequencing of both *L. japonicus* accessions and different Lotus species is a valuable resource that will facilitate genetic mapping, including linkage mapping in bi-parental populations and genome-wide association studies (GWAS). Currently, a whole-genome resequencing strategy is being used to sequence 130 Japanese *L. japonicus* accessions and a number of, mainly diploid, different Lotus species including *L. burttii*, *L. corniculatus*, *L. filicaulis*, *L. glaber*, *L. japonicus*, *L. pedunculatus*, *L. preslii*, *L. subbiflorus*, *L. tenuis*, *L. uliginosus* and *L. krylovii*. The polymorphism density and rate linkage disequilibrium decay within the Japanese population sample appear compatible with GWAS studies (Sato et al. 2014), and the large number of inter-species polymorphism will allow global studies of molecular gene evolution.

4.6 Future Perspectives

By integrating sequencing data from NGS platforms, the gene space coverage reached 98 %, and thus the genome sequence information of version 3.0 should serve as a solid basis for gene annotation, expression analysis, and insertion mutant identification. Further improvement of the genome sequence information is planned to be carried out in the following two ways: (1) assignment of unanchored contigs onto pseudomolecules by resequencing recombinant inbred lines to generate high-resolution genetic map information for the SNPs located on the unanchored contigs; (2) filling the gaps between the contigs by applying long read data generated by the Pacific Biosciences RS third generation sequencer.

References

Bao Z, Eddy SR (2002) Automated de novo identification of repeat sequence families in sequenced genomes. Genome Res 12:1269–1276

Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, Hall KP, Evers DJ, Barnes CL, Bignell HR et al (2008) Accurate whole human genome sequencing using reversible terminator chemistry. Nature 456:53–59

Fawcett JA, Kawahara T, Watanabe H, Yasui Y (2006) A SINE family widely distributed in the plant kingdom and its evolutionary history. Plant Mol Biol 61:505–514

Fukai E, Soyano T, Umehara Y, Nakayama S, Hirakawa H, Tabata S, Sato S, Hayashi M (2012) Establishment of a *Lotus japonicus* gene tagging population using the exon-targeting endogenous retrotransposon LORE1. Plant J 69:720–730

Fukai E, Umehara Y, Sato S, Endo M, Kouchi H, Hayashi M, Stougaard J, Hirochika H (2010) Derepression of the plant Chromovirus LORE1 induces germline transposition in regenerated plants. PLoS Genet 6: e1000868

Holligan D, Zhang X, Jiang N, Pritham EJ, Wessler SR (2006) The transposable element landscape of the model legume *Lotus japonicus*. Genetics 174:2215–2228

Ito M, Miyamoto J, Mori Y, Fujimoto S, Uchiumi T, Abe M, Suzuki A, Tabata S, Fukui K (2000) Genome and chromosome dimensions of *Lotus japonicus*. J Plant Res 113:435–442

Madsen LH, Fukai E, Radutoiu S, Yost CK, Sandal N, Schauser L, Stougaard J (2005) LORE1, an active low-copy-number TY3-gypsy retrotransposon family in the model legume *Lotus japonicus*. Plant J 44:372–381

Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z et al (2005) Genome sequencing in microfabricated high-density picolitre reactors. Nature 437:376–380

Pedrosa A, Sandal N, Stougaard J, Schweizer D, Bachmair A (2002) Chromosomal map of the model legume *Lotus japonicus*. Genetics 161:1661–1672

Sato S, Gupta V, Malolepszy A, Hirakawa H, Urbanski D, Fukai E, Sandal N, Carretero-Paulet L, Hanzly L, Jin H et al (2014) A *Lotus japonicus* genomic platform for integrated analysis of plant-microbe interactions reveals predominant legume root defenses. submitted

Sato S, Nakamura Y, Kaneko T, Asamizu E, Kato T, Nakao M, Sasamoto S, Watanabe A, Ono A, Kawashima K et al (2008) Genome structure of the legume, *Lotus japonicus*. DNA Res Int J Rapid Publ Rep Genes Genomes 15:227–239

Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song Q, Thelen JJ, Cheng J et al (2010) Genome sequence of the palaeopolyploid soybean. Nature 463:178–183

Urbanski DF, Malolepszy A, Stougaard J, Andersen SU (2012) Genome-wide LORE1 retrotransposon mutagenesis and high-throughput insertion detection in *Lotus japonicus*. Plant J 69:731–741

Urbanski DF, Malolepszy A, Stougaard J, Andersen SU (2013) High-Throughput and Targeted Genotyping of *Lotus japonicus* LORE1 Insertion Mutants. In: Rose RJ (ed) Legume genomics: methods and protocols. Springer, New York

Young ND, Debellé F, Oldroyd GED, Geurts R, Cannon SB, Udvardi MK, Benedito VA, Mayer KFX, Gouzy J, Schoof H et al (2011) The Medicago genome provides insight into the evolution of rhizobial symbioses. Nature 480:520–524

Genome Sequence and Gene Functions in *Mesorhizobium loti* and Relatives

5

Kazuhiko Saeki and Clive W. Ronson

Abstract

Mesorhizobium loti is a collective name for mesorhizobial species that establish nitrogen-fixing symbiosis with *Lotus* species. Accumulating genetic and genomic data indicate that diverse strains of *M. loti* have been generated through lateral integration of symbiosis islands into core chromosomes of a range of bacteria. The *M. loti* symbiosis islands probably derived from a common ancestral island and are evolving by acquiring accessory genetic elements while maintaining gene sets essential for nodulation and nitrogen fixation together with genes for some supportive processes. This view was supported by preliminary mappings of next-generation sequencing data of three strains, R7A, NZP2037, and NZP2213, on the whole-genome sequence of the strain MAFF303099. Common properties of *M. loti* genes involved in symbiosis and their regulation are also described along with genetic resources to study *M. loti*.

5.1 Introduction

Rhizobia are a group of soil bacteria that establish mutualistic relationships with leguminous plants. With a compatible legume host, rhizobium cells are capable of inducing root nodules, within which they reside intracellularly and fix atmospheric nitrogen to ammonium for nutritional exchange (Perret et al. 2000).

This chapter deals with genomic and genetic analysis of rhizobia that establish symbiosis with *Lotus japonicus*. *Mesorhizobium loti* strains MAFF303099 and R7A are described mainly, since they had been used in landmark studies such as the first determination of the complete genome sequence of a nitrogen-fixing bacterium (Kaneko et al. 2000) and the discovery of a symbiosis island as an integrative and conjugative element (ICE) that harbors most genes involved in root nodule formation and nitrogen fixation (Sullivan and Ronson 1998). However, taxonomy of rhizobia including those associated with *L. japonicus* is still in a state of flux (Sawada et al. 2003). Hence, we firstly describe

K. Saeki (✉)
Department of Biological Science, Faculty of Science, Nara Women's University, Nara, Japan
e-mail: ksaeki@cc.nara-wu.ac.jp

C.W. Ronson
Department of Microbiology and Immunology, University of Otago, PO Box 56, Dunedin New Zealand

S. Tabata and J. Stougaard (eds.), *The Lotus japonicus Genome*, Compendium of Plant Genomes,
DOI 10.1007/978-3-662-44270-8_5,

what *M. loti* is and the range of bacterial species to be discussed. Secondly, we describe genome structures of *M. loti* strains with special emphasis on symbiosis islands. We then discuss genes involved in symbiosis and their regulation. Genetic resources to study *M. loti* are also described.

5.2 *Lotus* Rhizobia, Definition of *M. loti* and Their Life Cycle

As stated above, this chapter mainly describes the two *M. loti* strains, MAFF303099 and R7A. A third strain NZP2037 that has an expanded host range compared to the other strains is also discussed. The three strains were all originally isolated from nodules of field-grown *Lotus* species (Jarvis et al. 1982; Kaneko et al. 2000). The complete genome sequence of MAFF303099 (Kaneko et al. 2000) as well as the symbiosis island sequences of R7A and NZP2037 has been reported (Sullivan et al. 2002; Kasai-Maita et al. 2013), and the complete genomes of R7A and NZP2037 have recently become available through the DOE Joint Genome Institute (JGI) Web site. These strains all belong to the same *Mesorhizobium* species. Nevertheless, there are concerns regarding the taxonomical definition not only of this species but also of the genus *Mesorhizobium* (Turner et al. 2002).

5.2.1 Original Definition of *M. loti*

Historically, bacterial species that form nitrogen-fixing nodules on the roots of plants of the genus *Lotus* were called *Lotus* rhizobia. *Lotus* rhizobia were recognized to include fast- and slow-growing groups with distinguishable features such as protein and extracellular polysaccharide compositions (Pankhurst et al. 1979). At this early stage of bacterial taxonomy, rhizobial nomenclature was based on cross-inoculation grouping of rhizobia and host legumes, although it was also known that such groups were not mutually exclusive. In 1982, Jarvis and colleagues proposed *Rhizobium loti* as a new species for fast-growing *Lotus* rhizobia with a type strain NZP2213 (Jarvis et al. 1982). Later, a group of *Rhizobium* species including *R. loti* were transferred from the genus *Rhizobium* to a new genus *Mesorhizobium* (Jarvis et al. 1997). In the second edition of Bergey's Manual of Systematic Bacteriology (2005), the genus *Mesorhizobium* is described as a member constituting the family *Phyllobacteriaceae* of the order *Rhizobiales* of the class *Alphaproteobacteria* (Sawada et al. 2003). *M. loti* cells are Gram negative, aerobic, non-spore-forming rods, and motile. The cells might contain, often as bacteroids in nodules, poly-beta-hydroxybutyrate inclusion bodies. On yeast mannitol agar, they form colonies greater than 1 mm after seven days at 28 °C, but their growth is inhibited at temperatures higher than 42 °C. Their growth is also inhibited by 2 % NaCl. They can grow within the pH range between 4.0 and 10.0. In addition to mannitol, *M. loti* cells utilize a number of carbohydrates, including glucose, galactose, arabinose, and maltose and produce acidic end products (Jarvis et al. 1982, 1997).

5.2.2 Diverse Nature of *M. loti* Strains

At the time of definition of *Rhizobium loti*, it was known that some *M. loti* strains have a broader host range than the type strain NZP2213. They can form effective nodules on some of *Lotus* species and non-*Lotus* plants on which NZP2213 cannot; such hosts include *Lotus pedunculatus* (or *Lotus uliginosus*) and a tree legume *Leucaena leucocephala* (Jarvis et al. 1982). NZP2037, originally isolated from a nodule off *Lotus divaricatus,* is the representative of the broader-host-range strains (Chua et al. 1985; Pankhurst et al. 1987). This strain is also capable of forming effective nodules on some species of *Carmichaelia*, *Ornithopus*, *Clianthus*, and *Vigna* (Chua et al. 1985; Pankhurst et al. 1987; Pueppke and Broughton 1999).

The strain R7A is a field reisolate of strain ICMP3153 (also known as NZP2238) and was

isolated in field plant passage experiments that showed the horizontal transfer of a large gene cluster from ICMP3153 to indigenous non-symbiotic mesorhizobia (Sullivan et al. 1995, 1996). The strain MAFF303099 was originally obtained in Japan from wild *L. japonicus*. Shortly after the determination of the complete genome sequence of this strain, Turner and colleagues argued that MAFF303099 should be classified as *Mesorhizobium huakuii* rather than *M. loti*, because the nucleotide sequences of MAFF303099 for 16S rRNA, *recA*, and some other genes are more similar to those of *M. huakuii* type strain CCBAU2609 (synonymous to IFO15243 and IAM14158) than those of *M. loti* type strain NZP2213 (Turner et al. 2002). According to this claim, some authors use *M. huakuii* biovar *loti* MAFF303099 or *Mesorhizobium* sp. MAFF303099.

Complicating this analysis, heterogeneity was reported for the *M. loti* type strain NZP2213 stocked in a number of culture collections (Willems et al. 2001). Some of culture stock strains synonymous to NZP2213 (ATCC33669, LMG6124, LMG17826, and IAM13588) were not homogenous and possessed 16S rRNA sequences more similar to those of the *M. huakuii* type strain than the originally registered nucleotide sequence of 16S rRNA gene for NZP2213 (GenBank accession number D14514). It is also documented that some of the stocks contained strains with morphological and protein profile diversities. It is unknown how these manifold type strains have emerged or kept coexistence. However, it might be noted that 16S rRNA sequences of *M. loti* strains MAFF303099, R7A, and NZP2037 are more similar to those of *M. huakuii* type strain than the sequence of NZP2213 (D14514), and hence, that "*M. loti*" strains fall into at least two distinct clusters that could be considered separate species.

Despite the disputes on taxonomic nomenclature, the accumulating data suggest that most rhizobial strains nodulating *Lotus* species contain similar sets of symbiosis genes present as part of the accessory genome, a symbiosis island, on chromosomes belonging to a certain range of *Phyllobacterium* species (Sullivan et al. 1996). The *nodA* gene of MAFF303099 (and R7A and NZP2037) is more closely related to that from NZP2213 than to the *M. huakuii nodA* sequences, as reported by Turner et al. (2002). It is also clearly distinguishable from other rhizobia such as *Rhizobium etli* CE3 and *Rhizobium* sp. NGR234 (or *Sinorhizobium* sp. NGR234 or *Ensifer fredii* NGR234) (Banba et al. 2001; Hussain et al. 1999; Schumpp et al. 2009).

5.3 Genome Structure of *M. loti* Strains

The complete genome sequence of *M. loti* MAFF303099 was reported in 2000 as the first rhizobial genome sequenced (Kaneko et al. 2000). The MAFF303099 genome consists of three circular replicons, a chromosome (7,036,071 bp) and two plasmids designated as pMLa (351,911 bp) and pMLb (208,315 bp). The genome comprises 7,281 potential protein-coding genes, two identical and consecutive sets of rRNA genes, 50 tRNA genes representing 47 tRNA species together with an RNase P RNA gene. Of the protein-coding genes, 6,752 genes are on the chromosome. About 80 % of the chromosomal genes show homology to genes of known function, while the ratios are about 50 and 35 % for those located on pMLa and pMLb. Adjacent to the single phenylalanine tRNA (phe-tRNA) gene, a 610,975-bp DNA segment is inserted on the chromosome with a 17-bp duplication of the 3′-terminal part of the tRNA gene demarcating the other end of the inserted element and with a phage P4-type integrase gene (*mll6432*) as the first gene at the phe-tRNA end of the element. This inserted element corresponds to the symbiosis island as defined in *M. loti* R7A and contains a number of genes for nitrogen fixation and symbiosis (Sullivan et al. 2002). The MAFF303099 symbiosis island and the two plasmids have GC contents (59.8, 59.3, and 59.9 %) lower than those of the core (non-symbiosis island part) chromosome (63.0 %). The GC content difference reflects the difference in the third position of codons of protein-coding genes: 70.2, 69.6, 71.2, and 80.5 % for symbiosis island, pMLa, pMLb, and the core chromosome,

respectively. Such differences indicate that the island and plasmids have been acquired by horizontal gene transfer from other genetic systems. The presence of these accessory genomes contributes to the existence of a number of genes present in multiple copies in the genome. For example, there are two *rpoN* genes for the alternative sigma factor sigma-54 involved in transcriptional control: one on the core chromosome (*mll3196* or *rpoN1*) and the other on the symbiosis island (*mlr5872* or *rpoN2*). Also, five sets of *groES–groEL* genes are present: three sets on the core chromosome (*mll2233–mll2232*, *mlr2393–mlr2394*, and *mll8202–mll8201*), one in the island (*msl5812–mll5810*), and the other on pMLa (*msr9431–mlr9342*).

5.3.1 Ongoing Genome Sequence Projects

The genome sequences of NZP2037 and R7A available at JGI and preliminary analysis of a few other *M. loti* strains indicate that their total genome sizes are about 7 Mb (JGI Web site and unpublished data). Some strains possess plasmids, while others do not; however, the main chromosomes of all the strains are larger than 5.5 Mb and contain symbiosis islands. The possession of symbiosis islands is a characteristic also found in *Bradyrhizobium* species (Itakura et al. 2009).

5.3.2 A Preliminary Whole-Genome Comparison of *M. loti* Strains

A preliminary genome comparison of MAFF303099 with R7A, NZP2037, and NZP2213 using Illumina GAIIx data of the latter three strains (Sakamoto and Saeki, original analysis) aligned on MAFF303099 chromosome is shown in Fig. 5.1. (The NZP2213 strain examined here had 16S rRNA sequence identical to that of D14514.) Shaded regions denote genes similar to MAFF303099 genes. Red color shade indicates similarity within the symbiosis island, while gray denotes similarity in the core chromosome. The mapping suggests that the extent of similarity in the core chromosome is in accordance with that of 16S rRNA genes. More than 85 and 80 % of the genes on MAFF303099 chromosome have homologues in R7A and NZP2037, respectively, while only about 55 % have homologues in NZP2213. The similarity between MAFF303099 and R7A is more evident in the core chromosome region than in the symbiosis island: Nearly 95 % of MAFF303099 genes have homologues in R7A, and many such genes, probably about 85 %, are only polymorphic at the nucleotide level. Nearly 85 % of the MAFF303099 chromosomal genes seem to be shared by NZP2037 with more polymorphisms, including at the amino acid sequence level. Similarity to NZP2213 seems less significant. For the genes on the MAFF303099 symbiosis island, about 45 % are also present in R7A, NZP2037, and NZP2213. Similarity of the shared genes seems more evident in R7A and NZP2213 than in NZP2037. These observations support the hypothesis that the diversity of *M. loti* strains results from the transfer of evolving symbiosis islands to diverse core chromosomes.

5.3.3 Chromosomal Genes Involved in Symbiotic Interaction

Mutagenic analyses of genes for exopolysaccharide (EPS) synthesis in NZP2037 and R7A have revealed that the symbiotic phenotype of EPS mutants of *M. loti* is influenced by the particular mutant, host plant ecotype, and environmental conditions (Hotter and Scott 1991; Kelly et al. 2013). The EPS genes are located in the core chromosome in MAFF303099 and are syntenic in the three strains (Kaneko et al. 2000; Kelly et al. 2013). Mutants affected in predicted early biosynthetic steps (e.g., *exoB*) formed nitrogen-fixing nodules on *L. japonicus* Gifu, whereas mutants affected in mid- or late biosynthetic steps (e.g., *exoU)* induced uninfected nodule primordia and, occasionally, a few

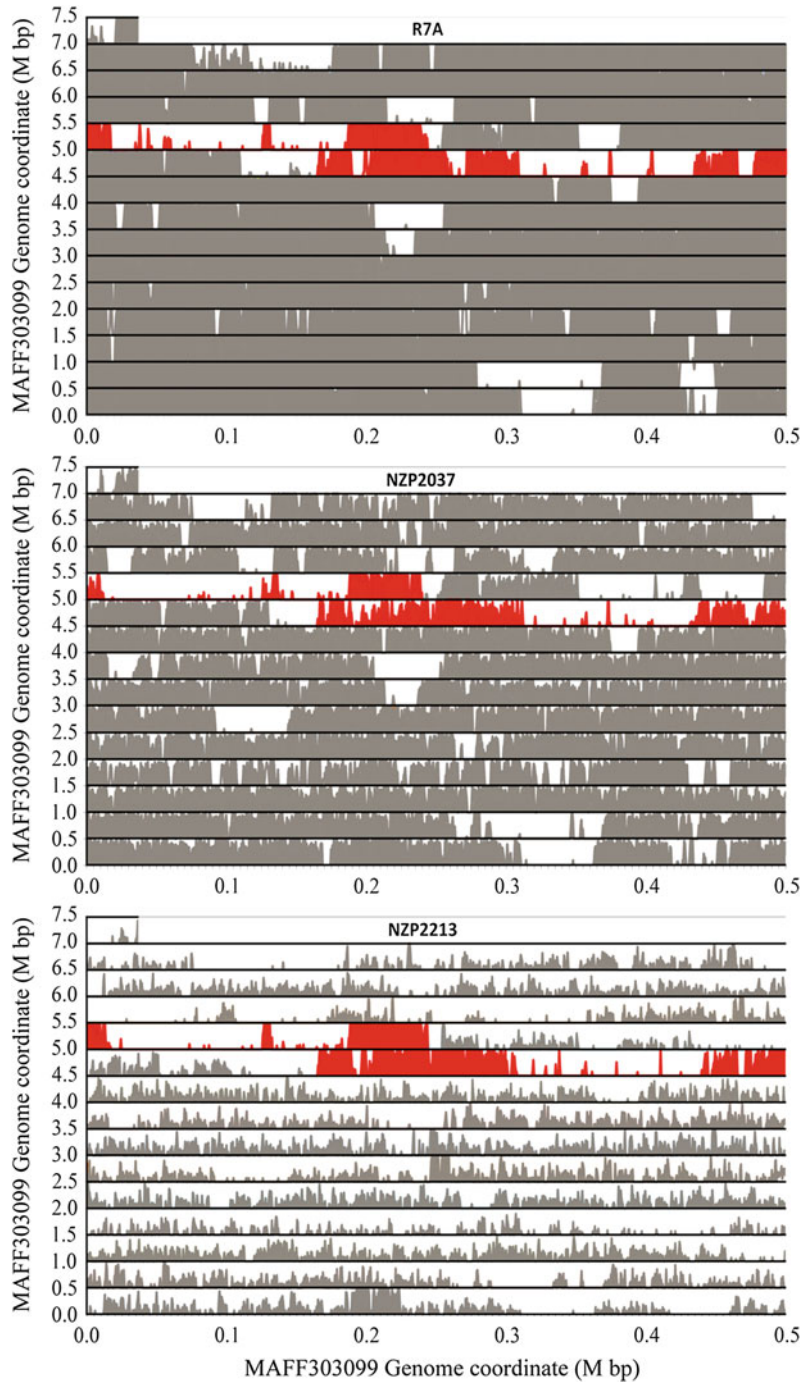


Fig. 5.1 Genome comparison of MAFF303099 with R7A, NZP2037, and NZP2213 by aligning Illumina GAIIx data on MAFF303099 chromosome

infected nodules following a lengthy delay. These mutants were disrupted at the stage of infection thread (IT) development. Symbiotically defective EPS and Nod factor mutants functionally complemented each other in coinoculation experiments. The majority of full-length IT observed harbored only the EPS mutant strain and did not show bacterial release, whereas the nitrogen-fixing nodules contained both mutants. Examination of the symbiotic proficiency of the

exoU mutant on various *L. japonicus* ecotypes revealed that both host and environmental factors were linked to the requirement for EPS. These results reveal a complex function for *M. loti* EPS in determinate nodule formation and suggest that EPS plays a signaling role at the stages of both IT initiation and bacterial release (Kelly et al. 2013).

Genes for LPS synthesis may also have a role in nodulation. At least some LPS mutants of *M. loti* form nodules impaired in nitrogen fixation (Turska-Szewczuk et al. 2009), while others form nitrogen-fixing nodules but are impaired in competitive ability (D'Antuono et al. 2005). The purified lipid A moiety of *M. loti* strain MAFF303099 LPS is able to induce nitrous oxide production in *L. japonicus* roots, suggesting that the plant is able to recognize rhizobial LPS (Hashimoto et al. 2012; Murakami et al. 2011). However, a precise role for rhizobial LPS in the nodulation of *Lotus* species is yet to be defined. *M. loti* mutants defective in cyclic ß-glucan synthesis are also defective in nodule invasion, with a possible role of the cyclic ß-glucan being in osmotic protection of the rhizobia (D'Antuono et al. 2005; Kawaharada et al. 2007, 2010).

5.4 Common Properties of *M. loti* Symbiosis Islands

To date, sequences of symbiosis islands from *M. loti* MAFF303099, R7A, and NZP2037 have been published. The three islands share highly collinear regions with multiple deletions and insertions (Fig. 5.2; see also Sullivan et al. 2002). The shared DNA regions contain genes for nodulation, nitrogen fixation, conjugal transfer of the island, and metabolic enzymes. The sizes of islands are approximately 502, 533, and 611 kb in R7A, NZP2037, and MAFF303099, respectively (Kaneko et al. 2000; Kasai-Maita et al. 2013; Sullivan et al. 2002). The number of genes they harbor is proportional to their size, about 1 gene/kbp: 414, 504, and 583 genes. As shown for R7A, the island in MAFF303099 is also inserted as an entity in the phe-tRNA gene, with a 17-bp duplication of the 3′-portion demarcating the other end of the island. The island of NZP2037 seems to have similarly inserted in a phe-tRNA gene; however, it has been split into two parts of 528 and 5 kb, through a recombination event likely mediated by a transposon. The collinear regions of three islands share 165 genes among a total of 1,078 non-redundant genes. These conserved genes, or core genes, on the symbiosis islands substantially cover the genes for nodulation, nitrogen fixation, and energy conversion in nodules as listed in Table 5.1. The numbers of genes shared by two of the three strains are 33, 25, and 35, respectively, by MAFF303099–R7A, MAFF303099–NZP2037, and R7A–NZP2037 pairs. The rest of the genes, 820 in total, are strain specific and mostly genes of unknown function and transposon-related genes such as transposases, integrases, and resolvases (Kasai-Maita et al. 2013). Nevertheless, some of the strain-specific genes include genes related to Nod factor synthesis and protein secretion systems that are involved in the interaction with host legumes.

5.4.1 Genes Involved in Nod Factor Synthesis and Export

Among the core genes on the symbiosis islands, there are 16 *nod*, *noe*, and *nol* genes involved in the synthesis, modification, and excretion of the lipochito-oligosaccharides (LCOs) or Nod factors that trigger the host developmental program to form nodules (Kaneko et al. 2000; Kasai-Maita et al. 2013; Sullivan et al. 2002). These include two copies of the regulatory gene *nodD,* the common *nod* genes *nodA, B,* and *C* (with *nodB* in a separate operon from *nodAC*, unlike other rhizobia) as well as other genes involved in chemical modifications of the core LCO. The onset of signal exchange generally takes place when flavonoids or isoflavones secreted by the host legume are bound by the rhizobial NodD protein that then activates transcription of other

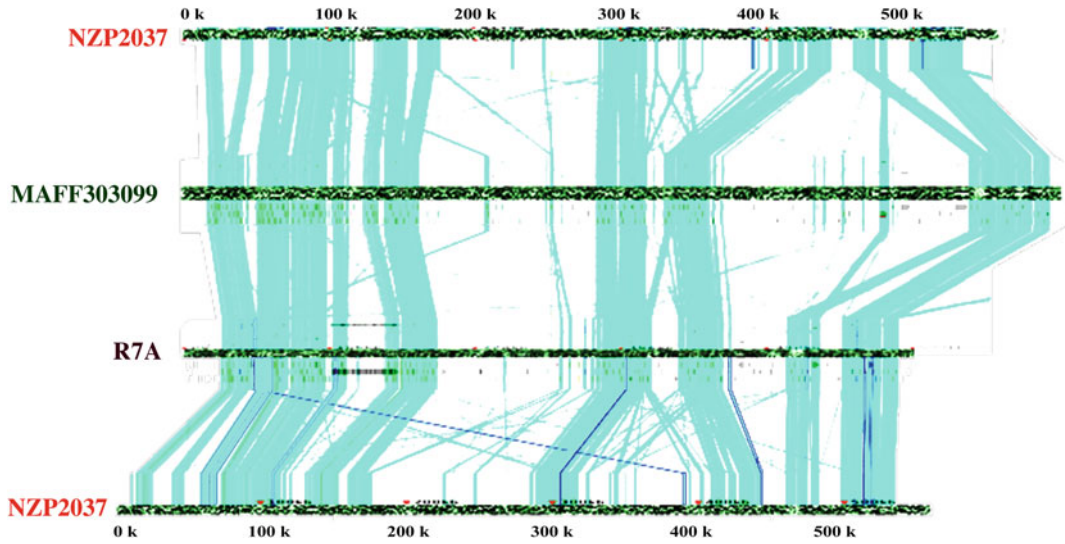


Fig. 5.2 Comparison of three *M. loti* symbiosis islands at nucleotide level by Murasaki (Popendorf et al. 2010)

nod genes to synthesize LCO. It might be noted that the *Lotus* compounds that activate NodD for induction of *nod* genes have not been determined yet, though aldonic acids have been reported to induce *M. loti nod* genes at high concentration (Gagnon and Ibrahim 1998).

The predominant LCO species produced by *M. loti* strains MAFF303099 and R7A is an N-acetylglucosamine pentasaccharide. The non-reducing residue is N-methylated and N-acylated with cis-vaccenic acid (C18:1) or stearic acid (C18:0) and carries a carbamoyl group, and the reducing terminal residue carries a 4-O-acetylfucosyl residue (Lopez-Lara et al. 1995; Niwa et al. 2001; Rodpothong et al. 2009). The core *N*-acetylglucosamine oligomer is synthesized by the *nodC* gene product, *N*-acetyl-glucosaminyltransferase, that extends the oligomer at the non-reducing terminus (Geremia et al. 1994). The *nodB* gene encodes a deacetylase that removes the *N*-acetyl moiety from the assembled *N*-acetylglucosamine oligomer at the non-reducing end (John et al. 1993). An acyl chain is then linked to the deacetylated oligomer by the *nodA* gene product (Rohrig et al. 1994). Although these assembly reactions are common and the genes shared in all rhizobia, the number of oligomeric repeats and the nature of added acyl chains are determined in part by the substrate specificity of the NodC and NodA proteins (Ritsema et al. 1996; Roche et al. 1996; Kamst et al. 1995).

Most other *nod* genes participate in specific modifications of the core structure. The *nodS* and *nolO* genes flanking *nodACIJ* encode methyltransferase and carbamoyltransferase, respectively. The *nolL* gene encodes O-acetyltransferase, while *nodM* is a glutamine-fructose-6-phosphate transaminase. Gene products of *noeK–noeJ* and *noeL–nolK* operons participate in the synthesis of GDP-D-mannose and GDP-L-fucose from GDP-D-mannose, respectively, while the *nodZ* gene encodes a fucosyltransferase. The assembled LCO is probably secreted by the function of NodI and NodJ proteins (Spaink 1995). Mutagenesis of *nodL* and *nodZ* genes in R7A revealed that the *N*-acetyl-fucose modification(s) of the LCO affects symbiotic performance depending on host *Lotus* species, whereas the methylation and carbamoylation apparently had no effect (Rodpothong et al. 2009).

Although not involved in Nod factor synthesis, the gene *acdS*, encoding 1-aminocyclopropane-1-carboxylic acid (ACC) deaminase, is conserved in the three symbiosis islands. The product possibly perturbs host ethylene synthesis in order to enhance nodulation capacity (Uchiumi et al. 2004; Nukui et al. 2006).

Table 5.1 Gene common among three symbiosis islands of *M. loti* strains

Function	Operon or gene	MAFF303099	Other position	R7A	NZP2037
		Symbiosis island		Symbiosis island	Symbiosis island
Nodulation gene					
	nodS–nodA–nodC–nodI–nodJ–nolO	mlr6161, mlr8755–mlr6171		ML0135, ML0133–ML0129	mln393–mln398
	nodB–nodD–nolL–nodD	mlr6175, mll6179, mlr8757, mlr6182		ML0126, ML0122, ML0120, ML0119	mln403, mln412, mln414, mln416
	nodM	mlr6386		ML0038	mln475
	noeK–noeJ	mlr5801, mlr5802		ML0396, ML0395	mln038–mln039
	nodZ–noeL–nolK	mlr5848, mlr5849, mlr8749		ML0366–ML0364	mln078–mln080
Nitrogen fixation genes					
	nifHDKEX	mlr5905–mlr5909		ML0303–ML0298	mln124–mln129
	nifB–fdxN–nifZ–fixU	mll5855–mll5852		ML0358–ML0361	mln085–mln082
	nifA	mll5837, mll5857		ML0372, ML0357	mln086, mln072
	nifSW	mll5865–mll5864		ML0351–ML0352	mln092–mln091
	fdxB–nifQ	mlr5869–mlr5871		ML0347–ML0346	mln096–mln097
	fixABCX	mll5862–msl5859		ML0358–ML0361	mln090–mln087
	fixNOPQ	mlr6411–mlr6414		ML0017–ML0014	mln497–mln500
	fixGHIS	mlr6415–msr6418		ML0013–ML0010	mln501–mln504
Conjugative transfer genes					
	trbBCDEJLFGI	mlr6397–mlr6405		ML0030–ML0022	mln484–mln492
	traG	mlr6395		ML0032	mln482
	rdfS–traF–rlxS	msr6186, mlr6188, mlr6189		ML0113–ML0111	mln422
	ardC	mlr6154		ML0157	mln380
	traI1	mlr6103, mlr6385		ML0178, ML0039	mln474
	traRI2, qseM,	mlr6102		ML0179	mln276, mln277

(continued)

Table 5.1 (continued)

Substrate transport					
	dctABD	mll5840, mlr5841–mlr5842		ML0370, ML0369–ML0368	mln074, mln075–mln077
			mll7237, mlr7238–mlr7239 (on chromosome)		
Perturbation of ethylene synthesis					
	acdS	mlr5932		ML0280	mln153
Cofactor synthesis					
Nicotinate biosynthesis	*nadC–nadB–nadA*	mll5833–mll5835		ML0376–ML0374	mln066–mln068
Biotin biosynthesis	*bioZ–bioA–bioD–bioF–bioB*	mll5827–mll5831		ML0382–ML0378	mln060–mln064
Thiamine biosynthesis	*thiD–thiE–thiG–thiS–thiO–thiC*	mll5788–mll5795		ML0406–ML0401	mln015–mln020
Aspartate 1-decarboxylase	*panD*	mll5826		ML0383	mln059
Carbon metabolism					
Putative phosphoenolpyruvate mutase	pepM	mlr5882		ML0337	mln102
Acetoacetate decarboxylase	*adc*	mll5917		ML0291	mln134
Fumarate hydratase	*fumA*	mlr6099		ML0182	mln273
A-type carbonic anhydrase	*cah*	mll6384		ML0040	mln473
Amino acid metabolism					
Sarcosine oxidase	*soxB1–soxD1—soxG1*	mll6240–mll6237		ML0245–ML0243	N. D.
	soxB2–soxD2–soxA2–soxG2	N. D.		ML0264–ML0450	N. D.
Serine hydroxymethyltransferase	*glyA*	mlr6114		ML0172	mln286
S-adenosylmethionine synthetase	*metK*	mlr6115		ML0171	mln287
5-methyltetrahydr opteroyltriglutamate					
-homocysteine methyltransferase	*metE*	mll6123		ML0165	mln294
L-threonine aldolase	*ltaM*	mlr6130		ML0161	mln300
Aspartate transaminase	*aatA*	mlr5883		ML0336	mln103
Asparagine synthetase	*asnB*	mlr5888		ML0333	mln106
Dihydrodipicolinate synthetase	*dapA*	mll6376		ML0043	mln470

(continued)

Table 5.1 (continued)

3-methylaspartate ammonia-lyase	*maaL*	mlr6095		ML0187	mln268
2-amino-3-ketobutyrate CoA ligase	*kbl*		mll9204 (on pMLa)	ML0254	N. D.
Cysteine synthase	*cycM*	N. D.	mll9227on (pMLa)	ML0212	N. D.
Phosphate metabolism					
Phosphonate degradation	*phnG–phnH–phnI–phnJ–phnK–phnM–phnL*	N. D.	mlr9277–mlr9286 (on pMLa)	ML0321–ML0315	N. D.

N. D.: Not Detected

5.4.2 Genes Involved in Nitrogen Fixation

There are 28 conserved genes probably involved in nitrogen fixation and related reactions. These include two copies of the regulatory gene *nifA*, genes encoding the nitrogenase enzyme and genes specifically involved in the synthesis of the complex metallo-clusters, electron transport, and energy conversion processes (Kaneko et al. 2000; Sullivan et al. 2002; Kasai-Maita et al. 2013). The regulatory hierarchy of the two NifA proteins is described below.

The first three genes of the *nifHDKENX* cluster encode structural proteins of nitrogenase enzyme complex that consists of two components. The terminal enzyme, dinitrogenase or MoFe protein, responsible for reduction of molecular nitrogen (N_2) is a hetero-tetramer ($\alpha_2\beta_2$) of NifD and NifK proteins that contains the metallo-clusters, iron–molybdenum cofactor (FeMo-co) and P cluster (Dixon and Kahn 2004; Rubio and Ludden 2008). Reducing equivalents and free energy to carry out the reaction are transferred to dinitrogenase from dinitrogenase reductase or Fe protein that is a homodimer of NifH protein containing a 4Fe–4S cluster and binding sites for Mg–ATP. The transfer of an electron is achieved by transient association between reduced ATP-bound Fe protein and MoFe protein. The hydrolysis of ATP by Fe protein causes a conformational change and is coupled to electron transfer from the 4Fe–4S cluster to P cluster in MoFe protein and dissociation. Electrons are then transferred from P cluster to FeMo-co where the catalytic reaction takes place.

FeMo-co is a unique cofactor only found in molybdenum nitrogenase. This cofactor consists of an inorganic moiety of Mo–Fe_7–S_9 and an organic moiety of *R*-homocitrate (Rubio and Ludden 2008). Products of *nifEN* are central scaffold of FeMo-co synthesis and transfer to *nifDK* product. NifH protein is also believed to function in FeMo-co synthesis with NifE and NifN protein complex. Most other *nif* genes have functions to supply or process constituents of FeMo-co. The *nifS* product is a cysteine desulfurase to supply sulfur to assemble transient Fe–S clusters on a scaffold protein generally encoded by *nifU*. No *nifU* homologue is located on the three *M. loti* symbiosis islands; however, there is one homologue, *mll0920*, on the core chromosome of strain MAFF303099, whose product might function as NifU scaffold. The Fe–S cluster on NifU is then transferred to another scaffold, a dimeric product of *nifB,* that utilizes *S*-adenosyl methionine to assemble the FeMo-co precursor. The precursor on NifB is then transferred to a NifEN scaffold with the help of the product of *nifX*. On the NifEN scaffold, molybdenum is donated by the product of *nifQ* (Rubio and Ludden 2008), while *R*-homocitrate is donated by the product of *nifV* (Hoover et al. 1987). Although *nifQ* is present on the *M. loti* symbiosis islands, no *nifV* that encodes the homocitrate synthase is found, not only on the three symbiosis islands but also on the core

chromosome (Kasai-Maita et al. 2013). It is also known that most rhizobia do not possess this gene to catalyze the condensation of acetyl coenzyme A and α-ketoglutarate to form homocitrate (Masson-Boivin et al. 2009). Intriguingly, this bacterial inability to synthesize homocitrate can be compensated by the host's capacity to supply the compound, as shown using a *L. japonicus* mutant lacking the *FEN1* gene encoding homocitrate synthase (Hakoyama et al. 2009).

The remaining *nif* genes, *nifW* and *nifZ*, may not be essential for nitrogenase function; however, their products can have roles in the maturation or stabilization of the MoFe protein (Rubio and Ludden 2008). In addition to the role in FeMo-co synthesis with NifEN, NifH is also required for maturation of the P clusters in the MoFe protein (Ribbe et al. 2002).

As described above, reducing equivalents and free energy are needed to carry out the nitrogenase reaction. For electron transport to nitrogenase, it is postulated that *fixABC* encodes an oxidoreductase complex composed of electron transfer flavoprotein (ETF) homologues and an ETF ubiquinone oxidoreductase homologue and that *fixX* encodes a ferredoxin-like protein to deliver electrons to nitrogenase. Nevertheless, experimental support for this postulation is still awaited. To generate ATP to be used by nitrogenase, rhizobial cells need to take up energy substrate(s) to generate membrane potential and ATP. The C4-dicarboxylate transport protein encoded by the conserved *dctA* gene is required for the uptake of the energy substrate, C4-dicarboxylates such as malate that are supplied by plant host, and its expression is regulated by the products of the adjacent *dctBD* genes. A set of genes homologous to *dctA–dctBD* is present on the MAFF303099 and R7A core chromosomes (*mll7237* and *mlr7238–7239* in MAFF303099), but these genes are non-functional because of a defect in the *dctB* gene (J. Weaver and C. Ronson, unpublished data). The products of *fixNOPQ* constitute an energy-transducing cytochrome cbb_3 oxidase that has high affinity for oxygen to meet the microaerobic conditions of the nodule environment. Adjacent to *fixNOPQ*, there are genes *fixGHIS* that encodes proteins essential for maturation of the oxidase. MAFF303099 has another set of *fixNOPQ* and *fixGHIS* genes on the core chromosome (Kaneko et al. 2000), and both sets are induced under free-living microaerobic conditions as well as under symbiotic conditions (Uchiumi et al. 2004).

5.4.3 Genes for Island Mobilization

There are more than 20 conserved genes that are probably involved in conjugative transfer of the symbiosis island. These include genes encoding proteins that constitute the transfer machinery as well as several regulatory genes (see below). Products of *rdfS* and *rlxS* function in excision and processing of DNA, while those of *traG* and the *trbBCDEJLFGI* operon function in mating-pore formation (Ramsay et al. 2006, 2009).

5.4.4 Genes for Metabolism of Vitamins and Other Compounds

There are conserved gene sets for biotin synthesis (Sullivan et al. 2001), nicotinic acid mononucleotide (NaMN) synthesis, and thiamine biosynthesis (Kaneko et al. 2000; Sullivan et al. 2002; Kasai-Maita et al. 2013). Non-symbiotic mesorhizobia are auxotrophic for these three vitamins (Sullivan et al. 1996), and these genes are absent from the MAFF303099 core chromosome. Therefore, symbiosis island transfer confers vitamin independence on the recipient mesorhizobia, as well as the ability to nodulate and fix nitrogen with *Lotus* species. Also conserved are 8 genes, *glyA, metK, metE, ltaM, aatA, asnB, dapA,* and *maaL*, encoding enzymes involved in amino acid metabolism. Most of these genes are highly expressed in bacteroids but are not essential for nitrogen fixation (Sullivan et al. 2013).

5.4.5 Genes for Protein Secretion into Host Cells

The three symbiosis islands contain gene clusters for secretion systems to deliver proteins into host cells, although the cluster in MAFF303099 encodes a type III secretion system (T3SS), whereas those in R7A and NZP2037 encode type IV secretion systems (T4SS). These secretion systems are not required for the nodulation of *L. japonicus* but do have negative and positive effects on the nodulation of some other host species (Hubber et al. 2004; Okazaki et al. 2010; Sanchez et al. 2012).

MAFF303099 possesses a number of *tts* genes or *rhc* (*rhizobium*-conserved) genes that encode the secretion machinery as well as *nop* (nodulation outer protein) genes that encode putative secreted effector proteins. The *rhc* genes together with the regulatory *ttsI* gene are also conserved in several other rhizobia including *Rhizobium* sp. NGR234 (Freiberg et al. 1997), *Sinorhizobium fredii* USDA257 (Krishnan et al. 2003), *S. fredii* HH103 (Marie et al. 2001; de Lyra et al. 2006), *B. japonicum* USDA110 (Krause et al. 2002), and *B. elkanii* USDA61 (Okazaki et al. 2009). The repertoire of Nops differs depending on the rhizobial species. Mutagenesis of *mlr6361* enables nodulation on *Lotus halophilus* which wild-type MAFF303099 is unable to nodulate, indicating that its product is a negative effector (Okazaki et al. 2010).

In R7A and NZP2037, the *virB1–virB11* and *virD4* genes encode the secretion machinery of T4SS (Hubber et al. 2004), while *virA* and *virG* encode a regulatory system that is itself under the control of the NodD regulatory protein that activates *nod* gene expression (Hubber et al. 2007). R7A and NZP2037 share some of T4SS effector proteins including Msi059 and Mln450 that are negative effectors that prevent nodulation of *L. leucocephala* (Hubber et al. 2004). Notably, Msi059 also shares strong similarity with the T3SS effector Mlr6316 in MAFF303099 that is also a negative effector for nodulation of *L. leucocephala* (Hubber et al. 2004). This commonality and the presence of *vir* gene remnants in MAFF303099 (Sullivan et al. 2002) suggest that the T3SS gene cluster in MAFF303099 replaced a T4SS gene cluster in a more ancient symbiosis island similar to that in R7A and NZP2037.

5.5 Regulation of Symbiotic Genes

5.5.1 Regulation of Genes Involved in Nod Factor Synthesis

In all the three *M. loti* symbiosis islands, there are two copies of *nodD* with syntenic positioning: *nodD1* is monocistronic and possesses no obvious upstream element, whereas *nodD2* is situated near *nodD1* but in the opposite direction and may be cotranscribed with the upstream *nolL* gene that is preceded by a *nod*-box. NodD is a transcriptional regulator that, when bound to a plant-derived flavonoid(s), binds to *nod*-boxes that are usually located upstream of other *nod* genes that participate in the synthesis of LCOs. The consensus sequence of the *nod*-box, 5′-YATCCAY-NNYRYRGATGNNNNYNATCNAAACAAT-CRATTTTACAATCY-3′, is conserved among rhizobia (Schlaman et al. 1998); hence, it can be possible to activate *nod* genes using heterologous *nodD* genes and flavonoid molecules. In fact, *M. loti* strains harboring *Rhizobium leguminosarum nodD* gene synthesize LCO when naringenin, the pea flavonoid that induces *nod* genes in *R. leguminosarum* bv. *viciae*, is added (Lopez-Lara et al. 1995; Niwa et al. 2001). In R7A, disruption of both *nodD1* and *nodD2* results in no nodulation (Nod^- phenotype), while disruption of *nodD1* gives delayed nodulation and that of *nodD2* has no apparent phenotype with *Lotus* species (Rodpothong et al. 2009). However, *nodD1* disruption alone abolishes nodulation capacity with *L. leucocephala* (Rodpothong et al. 2009). These observations indicate that *nodD1* and *nodD2* are functionally exchangeable, although *nodD1* is more important than *nodD2*.

5.5.2 Regulation of Genes Involved in Symbiotic Metabolism

In most rhizobia, cascading control is common to regulate gene expression for nitrogen fixation and symbiotic metabolism (Terpolilli et al. 2012). Such cascades are not fully conserved among rhizobial species. For example, in *S. meliloti*, the FixL kinase that senses oxygen and its cognate response regulator FixJ are required to express *nifA* and *fixK*, for transcription of other *nif* genes and *fixNOPQ* genes, respectively. In *B. japonicum*, phosphorelay from redox-sensing RegS to RegR is required to express *nifA*, since only *fixK2* but not *nifA* is under the direct control of FixL/FixJ. In addition, the alternative sigma factor RpoN is required for transcription of NifA-induced genes.

The *M. loti* symbiosis islands carry two *nifA* genes, *nifA1* located adjacent to *nifB* (the same genomic context as in *S. meliloti*) and *nifA2* situated away from known *nif* genes. There is also one *rpoN* gene, designated as *rpoN2*, in addition to *rpoN1* (*mll3196*) located on the core chromosome. Other genes in the common cascade, such as *fixL* (*mll6607*), *fixJ* (*mll6606*), *fixK* (*mll6578*), *regS* (*mlr5307*), and *regR* (*mlr5308*), are not located on symbiosis islands but are on the core chromosome. Mutational and expression analysis using both MAFF303099 and R7A indicates that *nifA2* but not *nifA1* is essential for symbiotic nitrogen fixation (Nukui et al. 2006; Sullivan et al. 2001, 2013). Only, *rpoN2* is required for nitrogen fixation (Sullivan et al. 2013). Further analysis indicated that FixL/FixJ and RegS/RegR are dispensable for *M. loti* to carry out symbiotic nitrogen fixation, unlike the other rhizobial counterparts (Sullivan et al. 2013). These analyses led Sullivan et al. to find a novel LacI/GalR family regulator FixV that is encoded adjacent to *nifA2* and activates *nifA2* expression, possibly in response to a host-derived inositol metabolite. NifA2 then activates expression of several genes including *rpoN2* and *nifA1* (Sullivan et al. 2013). Taken together, gene regulation of symbiotic nitrogen fixation in *M. loti* is peculiar to this species.

5.5.3 Regulation of Genes Involved in Protein Secretion

The T3SS and T4SS in *M. loti* strains have acquired the capacity to be induced by host flavonoids. These systems have specific transcriptional activators, TtsI and VirA/VirG, to express other genes for the secretion system. TtsI binds to *tts*-boxes and induces transcription of other *tts* or *rhc* genes (Krause et al. 2002; Zehner et al. 2008; Wassem et al. 2008). Likewise, VirG activated by VirA binds to *vir*-boxes and induces transcription of other *vir* genes (Leroux et al. 1987; Winans et al. 1986). In MAFF303099, there is a *nod*-box sequence upstream of *ttsI* that is responsible for transcriptional induction upon receiving plant flavonoid. TtsI-dependent *nopX* gene expression was evidenced by experiments with heterologous *R. leguminosarum* NodD and naringenin (Okazaki et al. 2010). In R7A, a *nod*-box sequence is found upstream of *virA*. The location of the *cis*-element enables induction of *virA* by NodD with host-derived flavonoid and subsequent activation of VirG to express T4SS machinery (Hubber et al. 2007).

5.5.4 Regulation of Genes Involved in Island Transfer

The regulation of symbiosis island transfer has been extensively studied in *M. loti* strain R7A, the only strain where the island has been shown to be mobile. The symbiosis island in MAFF303099 has a transposon insertion in the *oriT* site on the island, while the NZP2037 island has been disrupted by recombination. Hence, both of these islands have likely lost mobility. ICE*Ml*Sym^{R7A} chromosomal excision and conjugative transfer are stimulated by the quorum-sensing regulator TraR in complex with *N*-(3-oxohexanoyl)-*L*-homoserine lactone (3-oxo-C6-HSL). This quorum-sensing stimulation results in the expression of two conserved genes, *msi172* and *msi171*, that then control downstream excision and transfer steps. Quorum sensing and ICE*Ml*-Sym^{R7A} excision/transfer are inhibited in the

majority of the *M. loti* population by QseM, a recently identified protein antiactivator of TraR. QseM expression is in turn regulated by the DNA-binding protein QseC, which binds and represses the *qseM* promoter in a concentration-dependent manner. This complex regulation may form a molecular switch allowing activation of quorum sensing, excision, and transfer in a small proportion of cells in the population, rather than the population-wide activation typical of many quorum-sensing systems (Ramsay et al. 2009, 2013). This novel regulation mechanism with *qseM–qseC* seems to be conserved among laterally transferable integrative and conjugative elements that are found in various bacteria, although most of these elements do not contain TraR homologues (Ramsay et al. 2013).

5.6 Genetic Resources

A collection of transposon mutants of MAFF303099 was constructed by Shimoda and colleagues (Shimoda et al. 2008). Using the signature-tagged mutagenesis (STM) technique, more than 20,000 mutants were generated. Determination of insertion positions in more than 7,800 mutants identified insertions into approximately 50 % of MAFF303099 genes. These mutants with defined insertion points are deposited with the National BioResource Project (*L. japonicus* and *G. max*) in Japan and accessible from the Web site (http://www.legumebase.brc.miyazaki-u.ac.jp). Also available at the National BioResource Project (*L. japonicus* and *G. max*) in Japan are plasmid clones of defined inserts obtained during the sequencing project of the strain MAFF303099 (Kaneko et al. 2000).

5.7 Concluding Remark and Future Prospects

Here, we treated *M. loti* as a collective name for mesorhizobial species that nodulate *Lotus* species. Ever accumulating data on the species indicate that the various *M. loti* strains have been generated through lateral integration of symbiosis islands of probably a single origin into the core chromosome of a range of bacteria belonging to *Phyllobacteriaceae*. The *M. loti* symbiosis islands are evolving by acquiring (and losing) insertional elements while maintaining conserved gene sets for nodulation, nitrogen fixation, and other supportive processes. Acquisition of new elements on the islands does not necessarily enhance the symbiotic capacity, but rather often restrict the capacity as exemplified by the nodulation restrictions with *L. halophilus* and *L. leucocephala*, respectively, by possession of *mlr6361* in MAFF303099 and *msi059* in R7A (Hubber et al. 2004; Okazaki et al. 2010). Although it is still obscure how the combination of core chromosome and the island determines the nodulation and nitrogen-fixing efficiency and host range of a *Lotus* rhizobium, the peculiar gene regulation for nitrogen fixation, substrate metabolism as well as surface or extracellular components determined by the core chromosome might restrict the combinations. With the rapidly advancing sequencing technology and more traditional molecular genetic analysis, the genomic basis of the *Lotus–Mesorhizobium* symbiosis will be elucidated.

References

Banba M, Siddique AB, Kouchi H, Izui K, Hata S (2001) *Lotus japonicus* forms early senescent root nodules with *Rhizobium etli*. Mol Plant Microbe Interact 14 (2):173–180

Chua KY, Pankhurst CE, Macdonald PE, Hopcroft DH, Jarvis BD, Scott DB (1985) Isolation and characterization of transposon *Tn5*-induced symbiotic mutants of *Rhizobium loti*. J Bacteriol 162(1):335–343

D'Antuono AL, Casabuono A, Couto A, Ugalde RA, Lepek VC (2005) Nodule development induced by *Mesorhizobium loti* mutant strains affected in polysaccharide synthesis. Mol Plant Microbe Interact 18 (5):446–457

de Lyra Mdo C, Lopez-Baena FJ, Madinabeitia N, Vinardell JM, Espuny Mdel R, Cubo MT, Belloguin RA, Ruiz-Sainz JE, Ollero FJ (2006) Inactivation of the *Sinorhizobium fredii* HH103 *rhcJ* gene abolishes nodulation outer proteins (Nops) secretion and decreases the symbiotic capacity with soybean. Int Microbiol 9(2):125–133

Dixon R, Kahn D (2004) Genetic regulation of biological nitrogen fixation. Nat Rev Microbiol 2(8):621–631

Freiberg C, Fellay R, Bairoch A, Broughton WJ, Rosenthal A, Perret X (1997) Molecular basis of symbiosis between *Rhizobium* and legumes. Nature 387 (6631):394–401

Gagnon H, Ibrahim RK (1998) Aldonic Acids: A Novel Family of nod Gene Inducers of *Mesorhizobium loti*, *Rhizobium lupini*, and *Sinorhizobium meliloti*. Mol Plant Microbe Interact 11(10):988–998

Geremia RA, Mergaert P, Geelen D, Van Montagu M, Holsters M (1994) The NodC protein of *Azorhizobium caulinodans* is an N-acetylglucosaminyltransferase. Proc Natl Acad Sci USA 91(7):2669–2673

Hakoyama T, Niimi K, Watanabe H, Tabata R, Matsubara J, Sato S, Nakamura Y, Tabata S, Jichun L, Matsumoto T, Tatsumi K, Nomura M, Tajima S, Ishizaka M, Yano K, Imaizumi-Anraku H, Kawaguchi M, Kouchi H, Suganuma N (2009) Host plant genome overcomes the lack of a bacterial gene for symbiotic nitrogen fixation. Nature 462(7272):514–517

Hashimoto M, Tanishita Y, Suda Y, Murakami E, Nagata M, Kucho K-I, Abe M, Uchiumi T (2012) Characterization of nitric oxide-inducing lipid a derived from *Mesorhizobium loti* lipopolysaccharide. Microbes Environ 27(4):490–496

Hoover TR, Robertson AD, Cerny RL, Hayes RN, Imperial J, Shah VK, Ludden PW (1987) Identification of the V factor needed for synthesis of the iron-molybdenum cofactor of nitrogenase as homocitrate. Nature 329:855–857

Hotter GS, Scott DB (1991) Exopolysaccharide mutants of *Rhizobium loti* are fully effective on a determinate nodulating host but are ineffective on an indeterminate nodulating host. J Bacteriol 173(2):851–859

Hubber A, Vergunst AC, Sullivan JT, Hooykaas PJJ, Ronson CW (2004) Symbiotic phenotypes and translocated effector proteins of the *Mesorhizobium loti* strain R7A VirB/D4 type IV secretion system. Mol Microbiol 54(2):561–574

Hubber AM, Sullivan JT, Ronson CW (2007) Symbiosis-induced cascade regulation of the *Mesorhizobium loti* R7A VirB/D4 Type IV secretion system. Mol Plant Microbe Interact 20(3):255–261

Hussain AKMA, Jiang Q, Broughton WJ, Gresshoff PM (1999) *Lotus japonicus* nodulates and fixes nitrogen with the broad host range *Rhizobium* sp. NGR234. Plant Cell Physiol 40(8):894–899

Itakura M, Saeki K, Omori H, Yokoyama T, Kaneko T, Tabata S, Ohwada T, Tajima S, Uchiumi T, Honnma K, Fujita K, Iwata H, Saeki Y, Hara Y, Ikeda S, Eda S, Mitsui H, Minamisawa K (2009) Genomic comparison of *Bradyrhizobium japonicum* strains with different symbiotic nitrogen-fixing capabilities and other *Bradyrhizobiaceae* members. ISME J 3(3):326–339

Jarvis BDW, Pankhurst CE, Patel JJ (1982) *Rhizobium loti*, a New Species of Legume Root Nodule Bacteria. Int J Syst Bacteriol 32(3):378–380

Jarvis BDW, Van Berkum P, Chen WX, Nour SM, Fernandez MP, Cleyet-Marel JC, Gillis M (1997) Transfer of Rhizobium loti, Rhizobium huakuii, Rhizobium ciceri, Rhizobium mediterraneum, and Rhizobium tianshanense to Mesorhizobium gen. nov. Int J Syst Bacteriol 47(3):895–898

John M, Rohrig H, Schmidt J, Wieneke U, Schell J (1993) *Rhizobium* NodB protein involved in nodulation signal synthesis is a chitooligosaccharide deacetylase. Proc Natl Acad Sci USA 90(2):625–629

Kamst E, van der Drift KM, Thomas-Oates JE, Lugtenberg BJ, Spaink HP (1995) Mass spectrometric analysis of chitin oligosaccharides produced by *Rhizobium* NodC protein in *Escherichia coli*. J Bacteriol 177(21):6282–6285

Kaneko T, Nakamura Y, Sato S, Asamizu E, Kato T, Sasamoto S, Watanabe A, Idesawa K, Ishikawa A, Kawashima K, Kimura T, Kishida Y, Kiyokawa C, Kohara M, Matsumoto M, Matsuno A, Mochizuki Y, Nakayama S, Nakazaki N, Shimpo S, Sugimoto M, Takeuchi C, Yamada M, Tabata S (2000) Complete genome structure of the nitrogen-fixing symbiotic bacterium *Mesorhizobium loti*. DNA Res 7(6):331–338

Kasai-Maita H, Hirakawa H, Nakamura Y, Kaneko T, Miki K, Maruya J, Okazaki S, Tabata S, Saeki K, Sato S (2013) Commonalities and differences among symbiosis islands of three *Mesorhizobium loti* strains. Microbes Environ 28(2):275–278

Kawaharada Y, Eda S, Minamisawa K, Mitsui H (2007) A *Mesorhizobium loti* mutant with reduced glucan content shows defective invasion of its host plant *Lotus japonicus*. Microbiology 153(12):3983–3993

Kawaharada Y, Kiyota H, Eda S, Minamisawa K, Mitsui H (2010) Identification of the *Mesorhizobium loti* gene responsible for glycerophosphorylation of periplasmic cyclic β-1,2-glucans. FEMS Microbiol Lett 302 (2):131–137

Kelly SJ, Muszynski A, Kawaharada Y, Hubber AM, Sullivan JT, Sandal N, Carlson RW, Stougaard J, Ronson CW (2013) Conditional requirement for exopolysaccharide in the *Mesorhizobium-Lotus* symbiosis. Mol Plant-Microbe Interact MPMI 26 (3):319–329

Krause A, Doerfel A, Gottfert M (2002) Mutational and transcriptional analysis of the type III secretion system of *Bradyrhizobium japonicum*. Mol Plant-Microbe Interact MPMI 15(12):1228–1235

Krishnan HB, Lorio J, Kim WS, Jiang G, Kim KY, DeBoer M, Pueppke SG (2003) Extracellular proteins involved in soybean cultivar-specific nodulation are associated with pilus-like surface appendages and exported by a type III protein secretion system in *Sinorhizobium fredii* USDA257. Mol Plant Microbe Interact 16(7):617–625

Leroux B, Yanofsky MF, Winans SC, Ward JE, Ziegler SF, Nester EW (1987) Characterization of the *virA* locus of *Agrobacterium tumefaciens*: a transcriptional regulator and host range determinant. EMBO J 6 (4):849–856

Lopez-Lara IM, van den Berg JD, Thomas-Oates JE, Glushka J, Lugtenberg BJ, Spaink HP (1995) Structural identification of the lipo-chitin oligosaccharide nodulation signals of *Rhizobium loti*. Mol Microbiol 15(4):627–638

Marie C, Broughton WJ, Deakin WJ (2001) Rhizobium type III secretion systems: legume charmers or alarmers? Curr Opin Plant Biol 4(4):336–342

Masson-Boivin C, Giraud E, Perret X, Batut J (2009) Establishing nitrogen-fixing symbiosis with legumes: how many *rhizobium* recipes? Trends Microbiol 17 (10):458–466

Murakami E, Nagata M, Shimoda Y, Kucho K-I, Higashi S, Abe M, Hashimoto M, Uchiumi T (2011) Nitric Oxide Production Induced in Roots of *Lotus japonicus* by Lipopolysaccharide from *Mesorhizobium loti*. Plant Cell Physiol 52(4):610–617

Niwa S, Kawaguchi M, Imazumi-Anraku H, Chechetka SA, Ishizaka M, Ikuta A, Kouchi H (2001) Responses of a model legume *Lotus japonicus* to lipochitin oligosaccharide nodulation factors purified from *Mesorhizobium loti* JRL501. Mol Plant Microbe Interact 14(7):848–856

Nukui N, Minamisawa K, Ayabe S, Aoki T (2006) Expression of the 1-aminocyclopropane-1-carboxylic acid deaminase gene requires symbiotic nitrogen-fixing regulator gene *nifA2* in *Mesorhizobium loti* MAFF303099. Appl Environ Microbiol 72 (7):4964–4969

Okazaki S, Zehner S, Hempel J, Lang K, Göttfert M (2009) Genetic organization and functional analysis of the type III secretion system of *Bradyrhizobium elkanii*. FEMS Microbiol Lett 295(1):88–95

Okazaki S, Okabe S, Higashi M, Shimoda Y, Sato S, Tabata S, Hashiguchi M, Akashi R, Göttfert M, Saeki K (2010) Identification and functional analysis of type III effector proteins in *Mesorhizobium loti*. Mol Plant Microbe Interact 23(2):223–234

Pankhurst CE, Craig AS, Jones WT (1979) Effectiveness of *Lotus* root nodules: I. Morphology and flavolan content of nodules formed on *Lotus pedunculatus* by fast growing *Lotus* rhizobia. J Exp Bot 30 (6):1085–1093

Pankhurst CE, Hopcroft DH, Jones WT (1987) Comparative morphology and flavolan content of *Rhizobium loti* induced effective and ineffective root nodules on *Lotus* species, *Leuceana leucocephala*, *Carmichaelia flagelliformis*, *Ornithopus sativus*, and *Clianthus puniceus*. Can J Bot 65:2676–2685

Perret X, Staehelin C, Broughton WJ (2000) Molecular basis of symbiotic promiscuity. Microbiol Mol Biol Rev 64(1):180–201

Popendorf K, Tsuyoshi H, Osana Y, Sakakibara Y (2010) Murasaki: a fast, parallelizable algorithm to find anchors from multiple genomes. PLoS One. 5(9):e12651

Pueppke SG, Broughton WJ (1999) *Rhizobium* sp. Strain NGR234 and *R. fredii* USDA257 Share Exceptionally Broad, Nested Host Ranges. Mol Plant Microbe Interact 12(4):293–318

Ramsay JP, Sullivan JT, Stuart GS, Lamont IL, Ronson CW (2006) Excision and transfer of the *Mesorhizobium loti* R7A symbiosis island requires an integrase IntS, a novel recombination directionality factor RdfS, and a putative relaxase RlxS. Mol Microbiol 62(3):723–734

Ramsay JP, Sullivan JT, Jambari N, Ortori CA, Heeb S, Williams P, Barrett DA, Lamont IL, Ronson CW (2009) A LuxRI-family regulatory system controls excision and transfer of the *Mesorhizobium loti* strain R7A symbiosis island by activating expression of two conserved hypothetical genes. Mol Microbiol 73 (6):1141–1155

Ramsay JP, Major AS, Komarovsky VM, Sullivan JT, Dy RL, Hynes MF, Salmond GP, Ronson CW (2013) A widely conserved molecular switch controls quorum sensing and symbiosis island transfer in *Mesorhizobium loti* through expression of a novel antiactivator. Mol Microbiol 87(1):1–13

Ribbe MW, Hu Y, Guo M, Schmid B, Burgess BK (2002) The FeMoco-deficient MoFe protein produced by a *nifH* deletion strain of *Azotobacter vinelandii* shows unusual P-cluster features. J Biol Chem 277 (26):23469–23476

Ritsema T, Wijfjes AH, Lugtenberg BJ, Spaink HP (1996) *Rhizobium* nodulation protein NodA is a host-specific determinant of the transfer of fatty acids in Nod factor biosynthesis. Mol Gen Genet 251(1):44–51

Roche P, Maillet F, Plazanet C, Debelle F, Ferro M, Truchet G, Prome JC, Denarie J (1996) The common *nodABC* genes of *Rhizobium meliloti* are host-range determinants. Proc Natl Acad Sci USA 93(26): 15305–15310

Rodpothong P, Sullivan JT, Songsrirote K, Sumpton D, Cheung KWJT, Thomas-Oates J, Radutoiu S, Stougaard J, Ronson CW (2009) Nodulation Gene Mutants of *Mesorhizobium loti* R7A—*nodZ* and *nolL* Mutants Have Host-Specific Phenotypes on *Lotus* spp. Mol Plant Microbe Interact 22(12):1546–1554

Rohrig H, Schmidt J, Wieneke U, Kondorosi E, Barlier I, Schell J, John M (1994) Biosynthesis of lipooligosaccharide nodulation factors: *Rhizobium* NodA protein is involved in N-acylation of the chitooligosaccharide backbone. Proc Natl Acad Sci USA 91(8):3122–3126

Rubio LM, Ludden PW (2008) Biosynthesis of the iron-molybdenum cofactor of nitrogenase. Annu Rev Microbiol 62:93–111

Sanchez C, Mercante V, Babuin MF, Lepek VC (2012) Dual effect of *Mesorhizobium loti* T3SS functionality on the symbiotic process. FEMS Microbiol Lett 330 (2):148–156

Sawada H, Kuykendall LD, Young JM (2003) Changing concepts in the systematics of bacterial nitrogen-fixing legume symbionts. J General Appl Microbiol 49 (3):155–179

Schlaman HR, Phillips DA, Kondorosi E (1998) Genetic organization and transcriptional regulation of rhizobial nodulation genes. In: Spaink HP, Kondorosi A, Hooykaas PJJ (eds) The Rhizobiaceae: Molecular Biology of Model Plant-Associated Bacteria. Kluwer Academic Publishers, Dordrecht, pp 361–386

Schumpp O, Crevecoeur M, Broughton WJ, Deakin WJ (2009) Delayed maturation of nodules reduces symbiotic effectiveness of the *Lotus japonicus-Rhizobium* sp. NGR234 interaction. J Exp Bot 60(2):581–590

Shimoda Y, Mitsui H, Kamimatsuse H, Minamisawa K, Nishiyama E, Ohtsubo Y, Nagata Y, Tsuda M, Shinpo S, Watanabe A, Kohara M, Yamada M, Nakamura Y, Tabata S, Sato S (2008) Construction of signature-tagged mutant library in *Mesorhizobium loti* as a powerful tool for functional genomics. DNA Res 15 (5):297–308

Spaink HP (1995) The molecular basis of infection and nodulation by rhizobia: the ins and outs of sympathogenesis. Annu Rev Phytopathol 33(1):345–368

Sullivan JT, Ronson CW (1998) Evolution of rhizobia by acquisition of a 500-kb symbiosis island that integrates into a phe-tRNA gene. Proc Natl Acad Sci USA 95(9):5145–5149

Sullivan JT, Patrick HN, Lowther WL, Scott DB, Ronson CW (1995) Nodulating strains of *Rhizobium loti* arise through chromosomal symbiotic gene transfer in the environment. Proc Natl Acad Sci 92(19):8985–8989

Sullivan JT, Eardly BD, van Berkum P, Ronson CW (1996) Four unnamed species of nonsymbiotic rhizobia isolated from the rhizosphere of *Lotus corniculatus*. Appl Environ Microbiol 62(8):2818–2825

Sullivan JT, Brown SD, Yocum RR, Ronson CW (2001) The *bio* operon on the acquired symbiosis island of *Mesorhizobium* sp. strain R7A includes a novel gene involved in pimeloyl-CoA synthesis. Microbiology 147(5):1315–1322

Sullivan JT, Trzebiatowski JR, Cruickshank RW, Gouzy J, Brown SD, Elliot RM, Fleetwood DJ, McCallum NG, Rossbach U, Stuart GS, Weaver JE, Webby RJ, De Bruijn FJ, Ronson CW (2002) Comparative sequence analysis of the symbiosis island of *Mesorhizobium loti* strain R7A. J Bacteriol 184(11): 3086–3095

Sullivan JT, Brown SD, Ronson CW (2013) The NifA-RpoN regulon of *Mesorhizobium loti* strain R7A and its symbiotic activation by a novel LacI/GalR-family regulator. PLoS ONE 8(1):e53762

Terpolilli JJ, Hood GA, Poole PS (2012) What Determines the Efficiency of N2-Fixing Rhizobium-Legume Symbioses ? Adv Microb Physiol 60:325–389

Turner SL, Zhang X-X, Li F-D, Young JPW (2002) What does a bacterial genome sequence represent? Misassignment of MAFF 303099 to the genospecies *Mesorhizobium loti*. Microbiology 148(11):3330–3331

Turska-Szewczuk A, otocka B, Kutkowska J, Król J, Urbanik-Sypniewska T, Russa R (2009) The incomplete substitution of lipopolysaccharide with O-chain prevents the establishment of effective symbiosis between *Mesorhizobium loti* NZP2213.1 and Lotus corniculatus. Microbiol Res 164(2):163–173

Uchiumi T, Ohwada T, Itakura M, Mitsui H, Nukui N, Dawadi P, Kaneko T, Tabata S, Yokoyama T, Tejima K, Saeki K, Omori H, Hayashi M, Maekawa T, Sriprang R, Murooka Y, Tajima S, Simomura K, Nomura M, Suzuki A, Shimoda Y, Sioya K, Abe M, Minamisawa K (2004) Expression islands clustered on the symbiosis island of the *Mesorhizobium loti* genome. J Bacteriol 186(8):2439–2448

Wassem R, Kobayashi H, Kambara K, Le Quéré A, Walker GC, Broughton WJ, Deakin WJ (2008) TtsI regulates symbiotic genes in *Rhizobium* species NGR234 by binding to *tts* boxes. Mol Microbiol 68(3):736–748

Willems A, Hoste B, Tang J, Janssens D, Gillis M (2001) Differences between subcultures of the *Mesorhizobium loti* type strain from different culture collections. Syst Appl Microbiol 24(4):549–553

Winans SC, Ebert PR, Stachel SE, Gordon MP, Nester EW (1986) A gene essential for *Agrobacterium virulence* is homologous to a family of positive regulatory loci. Proc Natl Acad Sci USA 83(21):8278–8282

Zehner S, Schober G, Wenzel M, Lang K, Göttfert M (2008) Expression of the *Bradyrhizobium japonicum* Type III Secretion System in Legume Nodules and Analysis of the Associated *tts* box Promoter. Mol Plant Microbe Interact 21(8):1087–1093

Plant Genes Involved in Symbiotic Signal Perception/Signal Transduction

6

A. Binder, T. Soyano, M. Hayashi, M. Parniske, and S. Radutoiu

Abstract

A host genetic programme that is initiated upon recognition of specific rhizobial Nod factors governs the symbiosis of legumes with nitrogen-fixing bacteria. This programme coordinates two major developmental processes that run in parallel in legume roots: *de novo* cortical cell division leading to nodule primordia formation, and the infection thread initiation in the root hairs guiding bacteria towards dividing cortical cells. This chapter focuses on the plant genes involved in the recognition of the symbiotic signal produced by rhizobia, and the downstream genes, which are part of a complex symbiotic signalling pathway that leads to the generation of calcium spiking in the nuclear regions and activation of transcription factors controlling symbiotic genes induction.

6.1 Perception of Symbiotic Signals at the Plasma Membrane

Genetic and molecular studies of host–microbe interaction identified a two-way signal exchange as a central mechanism for partner recognition during the symbiotic establishment. Plant-produced strigolactones are recognized and perceived by the obligate arbuscular mycorrhiza (AM) fungi, while specific flavonoids or isoflavonoids secreted from the host roots (Peters et al. 1986; Spaink et al. 1989) are recognized by the symbiotic rhizobia. It has been well established that in rhizobia, this recognition has a direct consequence, as the flavonoid-activated rhizobial NodD proteins promote transcription of bacterial *nod* genes involved in synthesis and secretion of lipochitin oligosaccharides (Nod factors). These molecules serve as major bacterial signals detected by the legume host (Mulligan and Long 1985; Spaink et al. 1991; Truchet et al. 1991). Chitin-derived signals, both in the form of GlcNac tetra/pentamers and acylated GlcNac oligomers, are also produced by the AM fungi and are perceived by the plants as symbiotic signals (Maillet et al. 2011; Genre et al. 2013), but the AM genes

A. Binder · M. Parniske
Faculty of Biology, Ludwig-Maximilians University, Munich, Germany

T. Soyano · M. Hayashi
National Institute of Agrobiological Sciences, Tsukuba, Ibaraki, Japan

S. Radutoiu (✉)
Centre for Carbohydrate Recognition and Signalling, Department of Molecular Biology and Genetics, Aarhus University, Gustav Wieds Vej 10, 8000 Aarhus C, Denmark
e-mail: sir@mb.au.dk

S. Tabata and J. Stougaard (eds.), *The Lotus japonicus Genome*, Compendium of Plant Genomes,
DOI 10.1007/978-3-662-44270-8_6,

involved in synthesis of these molecules await identification.

Rhizobial Nod factors trigger physiological responses, gene expression, and cell division in susceptible legume roots during establishment of root nodule symbiosis (RNS). As a consequence, a new plant organ, the root nodule, hosting the symbiont is formed (reviewed in Oldroyd et al. 2011). The rhizobial symbiont of *Lotus japonicus*, *Mesorhizobium loti* strain R7A, produces a pentameric Nod factor with a 4-*O*-fucose at the reducing end, which either has an acetyl group or a proton in position 3 or 4. The non-reducing end is N-methylated and N-acylated (*cis*-vaccenic acid or stearic acid) and has a carbamoyl group in position 3 (Lopez-Lara et al. 1995; Bek et al. 2010). The acetylated fucosyl group is important for effective Nod factor signalling and the absence of this decoration leads to host-dependent nodulation phenotypes among different *Lotus* species: *L. japonicus*, *L. filicaulis*, *L. corniculatus* and *L. burttii* (Rodpothong et al. 2009).

Perception of Nod factors in legumes is mediated by receptor kinases containing LysM modules in their extracellular domains. In *Lotus*, the two receptor kinases perceiving the Nod factor signal, NFR1 and NFR5, are predicted to have a topology where single-pass transmembrane domains anchor the proteins to the plasma membrane exposing the LysM domains to the extracytoplasmic space, and the serine/threonine kinase to the cytoplasm (Madsen et al. 2003; Radutoiu et al. 2003; Madsen et al. 2011). Based on their domain similarity, and the fact that mutations in both *Nfr1* and *Nfr5* genes equally abolish Nod factor or rhizobia recognition, a heteromeric receptor complex (NFR1–NFR5) was proposed to initiate signal transduction following perception of a correctly decorated Nod factor (Radutoiu et al. 2003). Follow-up studies using bimolecular fluorescence complementation (BiFC) in *Nicotiana benthamiana* and leek cells revealed NFR1 and NFR5 localization at the plasma membrane and their interaction upon co-expression, supporting the originally suggested model (Madsen et al. 2011). The intracellular region of the NFR1 protein has all the subdomains of a typical kinase, and based on in vitro analyses, it has been shown to have the capacity for autophosphorylation and for NFR5 phosphorylation. On the other hand, NFR5, which lacks important kinase subdomains, and has therefore been considered a pseudokinase, failed to display kinase activity in a similar in vitro assay with myelin basic protein as a catalytic substrate (Madsen et al. 2011). The observation that a deletion of nine amino acids in the NFR5 kinase domain of the *nfr5-1* mutant abolishes the symbiotic interaction (Madsen et al. 2003) indicates that the intracellular region of NFR5 is also crucial, possibly serving as an interacting domain for downstream signalling components. A plasma membrane-associated remorin, SYMREM1, which is specifically induced during RNS, has been identified both in *Lotus japonicus* and in *Medicago truncatula*, and it has been shown to interact with Nod factor receptors in both model legumes (Levebvre et al. 2010; Tóth et al. 2012). Recently, using an Y2H approach, a ROP GTPase has been identified as an NFR5 kinase interactor, and transcriptional deregulation of this gene by RNAi led to a symbiotic defective phenotype (Ke et al. 2012). With the exception of *Lj*SYMREM1, no other interaction partners of NFR1 have been identified so far. However, the fact that particular amino acids and defined regions of the NFR1 kinase domain play a major role in downstream symbiotic signalling has been clearly shown by using point mutation constructs (Madsen et al. 2011) and domain swaps between *Lotus* NFR1 and *Arabidopsis* CERK1 (Nakagawa et al. 2011). CERK1 recognizes chitin (Miya et al. 2007; Wan et al. 2008), a microbial signal that is biochemically similar to rhizobial Nod factor. The chitin octamer is a PAMP produced by fungal pathogens, and CERK1 provides resistance in plants by activating a specific signalling cascade, leading to activation of defence genes (reviewed by Gust et al. 2012). NFR1 and CERK1 proteins show high sequence and structural similarity, especially in the intracellular region (Radutoiu et al. 2003; Miya et al. 2007; Wan et al. 2008). It is therefore unclear how specificity is achieved in the downstream signalling upon binding of symbiotic Nod factor or pathogen-derived chitin

to the ectodomain. It has been shown that a chimeric NFR1-CERK1 construct containing the YAQ amino acid sequence from NFR1 in the αEF helix of CERK1 kinase allows reconstitution of symbiotic signalling in the *nfr1* mutant (Nakagawa et al. 2011). This poses an important question: Does the YAQ sequence allow for the recruitment of specific downstream symbiotic partners or prevent the interaction with CERK1 partners acting in the defence pathway?

The presence of GlcNac-binding LysM modules in the extracellular regions of NFR1 and NFR5 represents a strong indication for the receptors' ability to bind rhizobial Nod factors. However, biochemical studies showing direct binding proved to be challenging due to the recalcitrant nature of NFR proteins and the amphiphilic properties of Nod factor ligands. Recently, a breakthrough in the biochemical analysis of NFRs-Nod factor binding ability has been achieved, and detailed studies using plant-produced proteins showed the ability of both *Lotus* receptors to directly bind, with high affinity, the Nod factors produced by *M. loti* (Broghammer et al. 2012). Using two different techniques, it has been shown that full length NFR1 and NFR5 proteins have K_d values for Nod factor binding in the nanomolar range. These are comparable with the ligand concentrations inducing membrane depolarization and calcium spiking in legume roots (Radutoiu et al. 2003; Miwa et al. 2006). These recent biochemical evidence complemented previous molecular studies that demonstrated the ability of NFR1 and NFR5 to mediate Nod factor perception and to ensure the specificity in the legume–rhizobia interaction. By transforming *M. truncatula* with the *L. japonicus Nfr1* and *Nfr5* genes, it was shown that the two receptors act in concert as host determinants, allowing *M. truncatula*, the non-host, to recognize and be infected by *M. loti,* the symbiont of *Lotus* (Radutoiu et al. 2007). Recognition of *M. loti* triggers initiation of nodule organogenesis in the root cortex as well as infection thread formation in the root hairs. This extended NFR1- and NFR5-mediated signal cascade is dependent on both Nod factor synthesis and structure. By using domain swap experiments and amino acid substitutions between NFRs of related *Lotus* species, it has been shown that a single amino acid variation, L118 to F, in the LysM2 domain of NFR5 plays a major role in discriminating *M. loti* and *R. leguminosarum* DZL Nod factors in *L. filicaulis* (Radutoiu et al. 2007). Interestingly, the same domain has been found in CERK1 to bind chitin (Liu et al. 2012), and homology modelling of the NFR5 LysM2 domain onto the CERK1 structure identified a possible binding groove, indicating a direct interaction with the ligand (M. Blaise, pers. communication). However, the presence of three LysM domains in the NFR5 and NFR1 receptors (Madsen et al. 2003; Radutoiu et al. 2003) suggests the involvement of more than one LysM domain in Nod factor perception. Three lines of evidence support this notion: (i) the non-nodulation phenotype caused by an amino acid substitution in the LysM1 domain of the *M. truncatula* homologue of NFR5 called NFP (Arrighi et al. 2006), (ii) the involvement of LysM1 of the pea SYM37 NFR1-like receptor in distinguishing 'European' and 'Middle East' *Rhizobium leguminosarum* bv. *viciae* strains (Zhukov et al. 2008) and (iii) the reported binding affinity of the CERK1 LysM2 domain to chitin (Liu et al. 2012) is significantly lower ($K_d = 44, 8$ μM) compared to the one found for Nod factor (Broghammer et al. 2012) in the case of NFR1 (K_d = 4.9 nM) and NFR5 (K_d = 10 nM). Surprisingly, a very high-affinity chitin-binding site ($K_d = 280$ pM), formed intramolecularly by the LysM1 and LysM3 domains, has been identified in the crystal structure of the *C. fulvum* fungal effector *Ecp6* (Sanchez-Vallet et al. 2013). By comparison with the CERK1 structure, the possibility of a similar LysM1–LysM3 arrangement in the ectodomain of the receptor has been excluded (Sanchez-Vallet et al. 2013). However, functionally CERK1 acts as a dimer, and NFR1–NFR5 forms a heterodimer, therefore the possibility of a corresponding high-affinity binding groove, formed in these cases intermolecularly, represents a very attractive hypothesis.

Protein–carbohydrate recognition events are central to cell–cell communication, cellular defence mechanisms, protein trafficking, and host–microbe recognition (Sacchettini et al. 2001).

In the past couple of years, an important role for chitin oligosaccharides (COs) or their derivatives, as signal molecules in plant and animal developmental processes and defence mechanisms, has emerged. In addition, carbohydrate-based microbe-associated molecular patterns (MAMPs), which are partly unidentified, are thought to be produced by microbes or plants during microbial infection (Boller and Felix 2009; Gimenez-Ibanez et al. 2009). Compared to other plant species, such as *Arabidopsis* or rice, the family of LysM receptor kinase in legumes, has expanded, but so far only *Nfr1, Nfr5* and *Lys3* (Kawaharada et al. manuscript in preparation) have an assigned function in rhizobial infection. In total, 17 loci encoding LysM proteins were identified in *Lotus* as having a typical receptor kinase structure (Lohmann et al. 2010). Analysis of the *Lys* genes location in the *L. japonicus* genome revealed the contribution of both tandem and segmental duplications for the expansion of this gene family. Seven *Lys* genes are arranged in tandem repeats (*Nfr1-Lys1-Lys2*, *Nfr5-Lys12*, *Lys13-Lys14*), while the occurrence of segmental duplication is supported by the syntenic regions on: (i) chromosome 4 containing *Lys11*, the closest paralog of *Nfr5*, and chromosome 2 where *Nfr5-Lys12* resides and (ii) chromosome 6 where *Lys6* is located and the *Nfr1-Lys1-Lys2* region. A tempting hypothesis for the evolutionary diversification of this receptor family in legumes is their unique capacity to decipher various structures of microbe-derived molecules produced by an extended spectrum of interacting organisms: associative, symbiotic, and parasitic bacteria or fungi. Gene expression studies performed on *Lotus* identified 13 *Lys* genes expressed in roots and nodules, which represent attractive candidates for the suggested cortical receptor of rhizobial Nod factors (Madsen et al. 2010) and for Myc signal(s) perception. However, the identification of receptors proves to be challenging due to the complex nature or the unknown structure of these microbial signals, and the possible redundancy among receptor genes. Rhizobial genes involved in Nod factor synthesis are transcriptionally active in the cortical infection threads (Timmers et al. 1998; Schlaman et al. 1991), but neither the structure, nor the composition of these signals is currently known. In the case of AM, both chitin oligomers and acylated forms of chitin with different decorations are able to induce calcium spiking or to activate nodulin genes (Maillet et al. 2011; Genre et al. 2013), and as a consequence of such complex signalling from the microbe, one could expect a complex deciphering mechanism in the host as well.

In both AM and RNS, the symbiotic signal perception at the plasma membrane leads to the activation of calcium spiking and ultimately to mycorrhiza- and nodulation-associated gene expression. Both symbioses share a common genetic programme, which was co-opted from a pre-existing AM pathway during the evolution of RNS (Kistner and Parniske 2002). In *Lotus japonicus,* at least eight genes of this common pathway have been identified so far: *SymRK*, *Nup85, Nup133, Nena, Castor, Pollux, CCaMK and Cyclops* (Stracke et al. 2002; Mitra et al. 2004; Kistner et al. 2005; Imaizumi-Anraku et al. 2005; Kanamori et al. 2006; Saito et al. 2007; Groth et al. 2010). This poses an interesting question: How are distinct rhizobial and AM fungal stimuli integrated by one common pathway to activate specific downstream signalling events for either symbiosis?

6.2 Signalling from Plasma Membrane to Nucleus

The receptor-like kinase (RLK) SYMRK (NORK/DMI2) mediates symbiotic signal transduction following Nod factor perception, and *symrk* mutants are deficient in both RNS and AM (Endre et al. 2002; Stracke et al. 2002). SYMRK is composed of an intracellular kinase, a transmembrane domain and an extracytoplasmic region consisting of leucine-rich repeats (LRRs) and a malectin-like domain (MLD) (Antolìn-Llovera et al. 2014). At least three distinct SYMRK variants with different sizes of the extracellular region exist in angiosperms. The longest version, containing three LRRs, is needed for RNS, while shorter versions are sufficient for

AM (Markmann et al. 2008). SYMRK interacts with both SYMREM1 and NFR5, potentially acting as a coreceptor of NFR5 in symbiotic signalling (Tóth et al. 2012; Antolìn-Llovera et al. 2014). The alteration of a conserved extracellular 'GDPC' sequence in the *symrk-14* mutant affects symbiotic development in the epidermis, but not in the cortex (Kosuta et al. 2011). Dependent on the 'GDPC' sequence, full length SYMRK is cleaved *in planta*, resulting in the release of the extracellular MLD. The cleavage product lacking the MLD (SYMRK-ΔMLD) outcompetes full length SYMRK for NFR5 interaction, suggesting that the MLD interferes with NFR5 binding (Antolìn-Llovera et al. 2014). Moreover, SYMRK-ΔMLD is rapidly degraded if the LRR region is present. Removal of the LRRs stabilizes the truncated SYMRK and results in an increased formation of infection threads (Antolìn-Llovera et al. 2014). The degradation of SYMRK-ΔMLD is potentially mediated by the E3 ligase SINA, an interactor of SYMRK, whose ectopic expression was correlated with reduced SYMRK protein levels and impaired infection thread development (Den Herder et al. 2012). Taken together, these results emphasize the importance of intricate SYMRK regulation in symbiotic signalling.

The components involved in signal transduction from the PM receptors to the nuclear calcium spiking machinery have not been determined; however, screens for interaction partners of known signalling components identified candidates for missing pieces. Among the interactors of SYMRK, a 3-hydroxy-3-methylglutaryl coenzyme A reductase1 (HMGR1) coenzyme (Kevei et al. 2007) as well as the MAP kinase kinase (MAPKK) SIP2 (Chen et al. 2012) has been identified, and silencing of either of these causes nodulation defects. MAP kinase or HMGR signalling might be involved in downstream signal transduction. HMGRs are the rate controlling enzymes of the mevalonate pathway that produces sterols, isoprenoids and in particular cytokinins, which are necessary for the induction of nodule morphogenesis (reviewed in Oldroyd et al. 2011). Metabolites of the HMGR pathway could act as symbiotic secondary messengers (Kevei et al. 2007). Although MtHMGR1 is upregulated during initial stages of mycorrhizal symbiosis (Liu et al. 2003), neither HMGR1 nor SIP2 was implicated with a function in mycorrhiza colonization, indicating that additional pathways downstream of SYMRK might be involved in AM signal transduction.

6.3 Generation of Nuclear Calcium Spiking

Nuclear and perinuclear calcium oscillations are initiated after contact with both rhizobia and AM fungi (Kosuta et al. 2008; Chabaud et al. 2011; Sieberer et al. 2012) and can be triggered directly by the addition of rhizobial lipo-chitooligosaccharide nodulation factors (NF) (Ehrhardt et al. 1996; Miwa et al. 2006; Sieberer et al. 2009) and short-chain chito-oligomers (COs), which are present in AM fungal exudates (Genre et al. 2013). Forward and reverse genetic screens in legumes have identified proteins involved in the generation of the calcium response. The closely related ion channels CASTOR and POLLUX (*M. truncatula* DMI1) are required for calcium spiking, and mutant alleles are deficient for RNS and AM colonization (Ané et al. 2004; Imaizumi-Anraku et al. 2005; Charpentier et al. 2008). The channels are located in the nuclear envelope (Riely et al. 2007; Charpentier et al. 2008), and in the case of DMI1, preferentially targeted to the inner side of the nuclear membrane (Capoen et al. 2011). Electrophysiological and functional analyses revealed that the proteins are cation channels. CASTOR showed a preference for K^+, and POLLUX could complement a yeast K^+ import and export mutant (Charpentier et al. 2008). Symbiotic signalling in *L. japonicus* requires both CASTOR and POLLUX. In *M. truncatula,* DMI1 alone is sufficient and expression of *Dmi1* in *Lotus* was able to complement a *castor/pollux* double mutant (Venkateshwaran et al. 2012). This functional difference between POLLUX and DMI1 was pinpointed to a single amino acid exchange in the putative selectivity filter region of the channels (alanine in POLLUX, serine in DMI1), which resulted in an increased mean channel opening time of DMI1 compared

with CASTOR. Exchanging the filter region of POLLUX to that of DMI1 allowed for the complementation of *dmi1* and *castor/pollux* double mutants; however, the same was not true when the change was introduced into CASTOR, which failed to rescue either *dmi1* or *castor/pollux* but was still able to rescue a *castor* single mutant (Venkateshwaran et al. 2012). This finding indicates a functional or regulatory difference between CASTOR and POLLUX/DMI1 that goes beyond their K^+ conductivity.

A mathematical model predicted that calcium-dependent activation of DMI1 and voltage-dependent opening of calcium channels in addition to the presence of a calcium pump are sufficient for sustained calcium oscillations (Granqvist et al. 2012). A SERCA type calcium pump, *M. truncatula* Calcium ATPase8 (MCA8), was localized to both the inner and the outer nuclear membrane. Silencing of *Mca8* perturbed spiking and resulted in reduced mycorrhization (Capoen et al. 2011).

Several hypotheses regarding the function of CASTOR and POLLUX/DMI1 in the calcium spiking machinery have been put forward. One model assumes that CASTOR and POLLUX/DMI1 are activated by secondary messengers, causing K^+ to flow into the perinuclear space. This would cause hyperpolarization of the nuclear membranes and in turn could lead to the opening of voltage-gated calcium channels (Venkateshwaran et al. 2012). In a slightly different model, it was suggested that for continued calcium spiking both DMI1 and calcium channels would need to be simultaneously activated by the binding of second messenger molecules. In this case, DMI1 (as well as CASTOR and POLLUX) would predominantly act as a counterion channel, but also initially contribute to the activation of the calcium channels by hyperpolarization of the nuclear membrane (Charpentier et al. 2013).

The role of the nuclear pore complex in symbiotic signalling remains poorly understood. Mutations in three nucleoporin genes, *Nup85, Nup133* and *Nena* (*Seh1*), abolish calcium spiking and cause defects in both RNS and AM symbiosis, similar to the phenotypes observed in *castor* and *pollux* mutants (Kistner et al. 2005; Kanamori et al. 2006; Saito et al. 2007; Groth et al. 2010). Yeast and vertebrate homologues of NUP85, NUP133 and NENA are part of the nuclear pore NUP107-160 subcomplex, which is an essential component of the NPC scaffold and required for NPC assembly (Walther et al. 2003; Doucet et al. 2010). Given the apparent lack of broad pleiotropic defects in the mutants, the subcomplex likely remains at least partially intact but can no longer fulfil certain functions required in symbiotic signal transduction. Aberrations in structure or distribution of the NPCs could prevent ongoing calcium oscillations by affecting the electrophysiological properties of the nucleus. Alternatively, changes in the NPC scaffold could also interfere with nucleo-cytoplasmatic transport of symbiotic proteins or messengers (reviewed in Binder and Parniske 2013). While import and export of macromolecules through the central NPC channel does not depend on the NUP107-160 subcomplex, larger membrane proteins ($>\sim$25 kDa), which are imported from the outer to the inner nuclear membrane, have to pass through both the central channel as well as the NPC scaffold (Meinema et al. 2011). As this implies a remodelling of nucleoporin connections in order to create on opening, it is conceivable that structural defects in the *nup* mutants can impair proper localization of nuclear envelope membrane proteins such as CASTOR, POLLUX, MCA8, or the calcium channels and thus affect calcium spiking.

6.4 Decoding the Calcium Signal

As the likely primary decoder of symbiotic calcium signatures, a nuclear calcium- and calmodulin-dependent kinase (CCaMK) plays a central role in symbiotic signal transduction (reviewed in Singh and Parniske 2012). *ccamk* mutants do not form infection threads, nodules and arbuscules when inoculated with rhizobia or AM fungi (Levy et al. 2004; Mitra et al. 2004). A calmodulin (CaM)-binding domain and three calcium-binding EF-hands mediate CCaMKs regulation during calcium spiking (Swainsbury

et al. 2012). Calcium binding induces a conformational change in the protein (Swainsbury et al. 2012) and promotes its autophosphorylation (Takezawa et al. 1996; Sathyanarayanan et al. 2000). Constitutive activation of CCaMK, caused by mutations in an autophosphorylation site (T265D, T265I), leads to spontaneous nodule development in the absence of rhizobia (Gleason et al. 2006; Tirichine et al. 2006). Negative regulation of CCaMK as a result of autophosphorylation within the calcium/CaM-binding domain (Liao et al. 2012; Routray et al. 2013) is also required for normal cortical infection and AM development (Liao et al. 2012), demonstrating an intricate modulation of CCaMK activity during symbiotic signalling. Autoactive CCaMK mutants are able to restore nodulation, and AM symbiosis in the mutants *symrk, castor, pollux, nup85* and *nup133* (Hayashi et al. 2010; Madsen et al. 2010). This indicates that the primary function of these genes is the activation of calcium spiking and highlights the importance of calcium signalling in symbiotic signal transduction.

It emerges that the calcium oscillations themselves may carry cell type and stage-specific information. Live cell imaging demonstrated a transition from low- to high-frequency spiking during apoplastic cell entry that was very similar for both mycorrhizal and rhizobial symbionts (Sieberer et al. 2012). Previously, calcium oscillations induced by AM fungal hyphae were described to be less regular than Nod factor-induced spiking (Kosuta et al. 2008; Chabaud et al. 2011); however, so far there has been no proof of differential decoding of AM and Nod factor-induced calcium signals by CCaMK.

CCaMK forms a complex with the nuclear protein CYCLOPS (*M. truncatula IPD3*) (Singh et al. 2014), which is essential for microbial infection (Messinese et al. 2007; Yano et al. 2008). CYCLOPS was revealed to be a novel type of transcriptional activator, which upon phosphorylation by CCaMK binds to a CYCLOPS responsive *cis* element in the *Nin* promoter and actives *Nin* gene expression (Singh et al. 2014). Phosphorylation of CYCLOPS S50 and S154 is critical for promoter binding and symbiotic development. An autoactive phosphomimetic mutant version of CYCLOPS (S50D/S154D) triggers spontaneous nodule formation independent of CCaMK, indicating that CYCLOPS acts as a master regulator of root nodule organogenesis (Singh et al. 2014).

6.5 Common Symbiosis Genes Involved in Microbial Accommodation

Several AM genes with putative functions in membrane trafficking are also involved in RNS. *M. truncatula Vapyrin* (*Petunia Pam1*) is required for arbuscule formation and efficient fungal entry of the root (Reddy et al. 2007; Feddermann et al. 2010; Pumplin et al. 2010; Murray et al. 2011) and deletion of the gene prevents rhizobial infection threads from reaching the cortical cell layer, resulting in an increased number of uninfected nodule primordia (Murray et al. 2011). *Vapyrin* encodes a protein with an N-terminal VAMP-associated protein (VAP)/major sperm protein (MSP) domain and a C-terminal ankyrin-repeat domain. Based on the domain structure and observed localization in the nucleus, cytosol and in distinct puncta in colonized cells, VAPYRIN was proposed to be involved in membrane trafficking and cellular rearrangement during symbiotic accommodation; however, this has not been experimentally verified (Pumplin et al. 2010; Murray et al. 2011). Two closely related *M. truncatula* vesicle-associated membrane proteins (VAMP) are involved in rhizobium–legume symbiosis and AM (Ivanov et al. 2012). Silencing of both *Vamp721d* and *Vamp721e* inhibits arbuscule and symbiosome formation and blocks bacterial release from the infection thread. Both gene products localize to small vesicles, which accumulate at bacterial release sites near symbiosome membranes and VAMP721e also accumulates at the tips of arbuscule branches, potentially at the periarbuscular membrane (Ivanov et al. 2012). While other VAMP72 proteins in *A. thaliana* are recruited during the interaction with biotrophic fungi (Kwon et al. 2008), VAMP721d/e are not present in the *Arabidopsis* genome. Considering

that *Arabidopsis* has lost several symbiosis genes (Zhu et al. 2006), it is possible that the exocytotic pathway involving VAMP721d/e is specific to perimicrobial membrane synthesis.

6.6 Transcription Factors Involved in Early Symbiotic Responses

Transcriptional regulation is important to integrate signalling pathways and to coordinately regulate molecular networks. The calcium signal that is decoded by CCaMK leads to the expression of subsets of genes. Autoactive CCaMK substitutes for Nod factors to activate expression of *ENOD11*, which is induced by both rhizobial and mycorrhizal infection (Gleason et al. 2006; Journet et al. 2001). CCaMK activation seems to be an intracellular switch that activates transcriptional networks.

Several transcription factors that regulate symbiotic processes have been identified so far. Among them, the GRAS family transcription factors, NSP1 and NSP2 (Nodulation Signalling Pathway 1 and 2), act furthest upstream of the CCaMK-mediated pathway. They were initially identified as factors specific to RNS (Catoira et al. 2000; Oldroyd and Long 2003; Kaló et al. 2005; Smit et al. 2005; Heckmann et al. 2006; Murakami et al. 2006), and then involvement in mycorrhizal colonization was recently identified (Maillet et al. 2011; Delaux et al. 2013). *nsp1* and *nsp2* mutations downregulate expression of *Vapyrin* and *Enod11* (Oldroyd and Long 2003; Hirsch et al. 2009; Murray et al. 2011). This is consistent with the idea that *Nsp1* and *Nsp2* are common to both AM and RNS. In nodulation processes, *nsp1* and *nsp2* mutants exhibit phenotypes similar to loss-of-function *ccamk* mutants in regard to symbiotic root hair responses as well as nodule formation. Root hairs of these mutants are deformed in response to Nod factors, but do not display root hair curling caused by rhizobial infection (Catoira et al. 2000; Oldroyd and Long 2003; Heckmann et al. 2006; Murakami et al. 2006). Autoactive CCaMK mutants do not rescue phenotypes of *nsp1* and *nsp2,* unlike those of common *SYM* mutants defective in the nuclear calcium spiking (Hayashi et al. 2010; Madsen et al. 2010). These results suggest that a site in the nodulation processes where NSP1 and NSP2 act is close to that of CCaMK and that the GRAS family proteins are required for the CCaMK-mediated pathway. How activities of the GRAS proteins are regulated is an important issue for understanding symbiotic signal transduction.

NSP1 and NSP2 form a heterodimer and bind to the promoters of the transcription factors *M. truncatula Ern1* (*ERF Required for Nodulation1*) and *Nin* (*NODULE INCEPTION*) in vitro as well as that of *Enod11* in vitro and in vivo (Hirsch et al. 2009). Expression of *Ern1* and *Nin* is induced by rhizobial infection depending on NSP1 and NSP2 (Murakami et al. 2006; Marsh et al. 2007; Hirsch et al. 2009; Cerri et al. 2012). NSP2 also interacts with an AM-specific GRAS family protein, RAM1 (Required for Arbuscular Mycorrhization1) (Gobbato et al. 2012), which directly targets *Ram2* expression. Both *Ram* genes were identified during a forward genetic screen that aimed to identify loci specifically involved in mycorrhizal signalling (Gobbato et al. 2012; Wang et al. 2012). Multiple dimerization of the GRAS family transcription factors is involved in the regulation of symbiotic processes. The heterodimerization between NSP1 and NSP2 seems to be important for NSP2 function, because an NSP2 derivative with an amino acid substitution in the domain responsible for binding with NSP1 resulted in the reduction of nodulation efficiency (Hirsch et al. 2009).

ERN1 coordinately regulates *Enod11* expression with the NSP1–NSP2 complex by targeting a *cis*-acting element different from those for NSP1–NSP2 complex. ERN1 is required to activate Nod factor-elicited *Enod11* expression during early pre-infection, while NSP1–NSP2 mediates *Enod11* expression during subsequent rhizobial infection (Cerri et al. 2012). NIN is a RWP-RK domain-containing transcription factor specific to and essential for RNS (Schauser et al. 1999). Root hairs of *nin* mutants are deformed and excessively curled in response to rhizobial infection, and failed to initiate infection thread

development. This indicates that NIN acts downstream of CCaMK and the GRAS transcription factors to regulate symbiotic root hair responses (Marsh et al. 2007). *Nin* is activated by cytokinin through a *L. japonicus* cytokinin receptor, LHK1 (Tirichine et al. 2007). Gain-of-function LHK1 spontaneously induces nodules without rhizobial infection (Tirichine et al. 2007). Ectopic expression of *Nin* also induces cortical cell divisions in the absence of rhizobia (Soyano et al. 2013). NIN regulates cortical cell divisions downstream of the cytokinin signalling.

LjNF-YA1 and *LjNF-YB1* have been identified as direct targets of NIN (Soyano et al. 2013). They are involved in stimulation of cell division. They encode different subunits of a CCAAT-box binding heterotrimeric complex. Knock-down of *LjNF-YA1* prevents the nodule formation. Co-overexpression of the two *Lotus NF-Y* genes stimulates cells division in lateral root primordia as well as cortical cell division. NF-Y regulates expression of its target genes by influencing histone modification and requires an additional transcriptional activator to efficiently activate transcription. NIN may also regulate other transcription factors that act together with the NF-Y to induce cortical cell division. *Arabidopsis* NIN-like proteins (NLPs) play a central role in the transcriptional regulation of nitrate-responsive genes and target nitrate-responsive elements (Konishi and Yanagisawa 2013), which are almost identical to NIN-binding nucleotide sequences. Nitrate is known as an inhibitor of nodulation. There may be a linkage between nodulation control and nitrate-response pathways.

Acknowledgments We thank Sylvia Singh for carefully reading the manuscript.

References

Ané JM, Kiss GB, Riely BK, Penmetsa RV, Oldroyd GED, Ayax C, Levy J, Debelle F, Baek JM, Kalo P, Rosenberg C, Roe BA, Long SR, Denarie J, Cook DR (2004) *Medicago truncatula* DMI1 required for bacterial and fungal symbioses in legumes. Science 303:1364–1367

Antolín-Llovera M, Ried MK, Parniske M (2014) Cleavage of the SYMBIOSIS RECEPTOR-LIKE KINASE ectodomain promotes complex formation with nod factor receptor 5. Curr Biol 24(4):422–427. doi:10.1016/j.cub.2013.12.053

Arrighi JF, Barre A, Ben Amor B, Bersoult A, Soriano LC, Mirabella R, de Carvalho-Niebel F, Journet EP, Gherardi M, Huguet T, Geurts R, Denarie J, Rouge P, Gough C (2006) The *Medicago truncatula* lysin [corrected] motif-receptor-like kinase gene family includes NFP and new nodule-expressed genes. Plant Physiol 142:265–279

Bek AS, Sauer J, Thygesen MB, Duus JO, Petersen BO, Thirup S, James E, Jensen KJ, Stougaard J, Radutoiu S (2010) Improved characterization of nod factors and genetically based variation in LysM receptor domains identify amino acids expendable for nod factor recognition in *Lotus* spp. Mol Plant Microbe Interact 23:58–66

Binder A, Parniske M (2013) The Nuclear pore complex in symbiosis and pathogen defense. In: Evans ER, Graumann K, Bryant JA (eds) Annual plant reviews: plant nuclear structure, genome architecture and gene regulation, vol 46. Wiley, Oxford, pp 229–254

Boller T, Felix G (2009) A renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. Annu Rev Plant Biol 60:379–406

Broghammer A, Krusell L, Blaise M, Sauer J, Sullivan JT, Maolanon N, Vinther M, Lorentzen A, Madsen EB, Jensen KJ, Roepstorff P, Thirup S, Ronson CW, Thygesen MB, Stougaard J (2012) Legume receptors perceive the rhizobial lipochitin oligosaccharide signal molecules by direct binding. Proc Natl Acad Sci USA 109:13859–13864

Capoen W, Sun J, Wysham D, Otegui MS, Venkateshwaran M, Hirsch S, Miwa H, Downie JA, Morris RJ, Ane JM, Oldroyd GE (2011) Nuclear membranes control symbiotic calcium signaling of legumes. Proc Natl Acad Sci USA 108:14348–14353

Catoira R, Galera C, de Billy F, Penmetsa RV, Journet EP, Maillet F, Rosenberg C, Cook D, Gough C, Denarie J (2000) Four genes of Medicago truncatula controlling components of a nod factor transduction pathway. Plant Cell 12:1647–1666

Cerri MR, Frances L, Laloum T, Auriac MC, Niebel A, Oldroyd GE, Barker DG, Fournier J, de Carvalho-Niebel F (2012) *Medicago truncatula* ERN transcription factors: regulatory interplay with NSP1/NSP2 GRAS factors and expression dynamics throughout rhizobial infection. Plant Physiol 160:2155–2172

Chabaud M, Genre A, Sieberer BJ, Faccio A, Fournier J, Novero M, Barker DG, Bonfante P (2011) Arbuscular mycorrhizal hyphopodia and germinated spore exudates trigger Ca^{2+} spiking in the legume and nonlegume root epidermis. New Phytol 189:347–355

Charpentier M, Bredemeier R, Wanner G, Takeda N, Schleiff E, Parniske M (2008) *Lotus japonicus* CASTOR and POLLUX are ion channels essential for perinuclear calcium spiking in legume root endosymbiosis. Plant cell 20:3467–3479

Charpentier M, Vaz Martins T, Granqvist E, Oldroyd GED, Morris RJ (2013) The role of DMI1 in establishing Ca^{2+} oscillations in legume symbioses. Plant Signal Behav 8:e22894

Chen T, Zhu H, Ke D, Cai K, Wang C, Gou H, Hong Z, Zhang Z (2012) A MAP kinase kinase interacts with SymRK and regulates nodule organogenesis in *Lotus japonicus*. Plant Cell 24:823–838

Delaux PM, Becard G, Combier JP (2013) NSP1 is a component of the Myc signaling pathway. New Phytol 199:59–65

Den Herder G, Yoshida S, Antolin-Llovera M, Ried MK, Parniske M (2012) *Lotus japonicus* E3 ligase SEVEN IN ABSENTIA4 destabilizes the symbiosis receptor-like kinase SYMRK and negatively regulates rhizobial infection. Plant Cell 24:1691–1707

Doucet CM, Talamas JA, Hetzer MW (2010) Cell cycle-dependent differences in nuclear pore complex assembly in metazoa. Cell 141:1030–1041

Ehrhardt DW, Wais R, Long SR (1996) Calcium spiking in plant root hairs responding to Rhizobium nodulation signals. Cell 85:673–681

Endre G, Kereszt A, Kevei Z, Mihacea S, Kalo P, Kiss GB (2002) A receptor kinase gene regulating symbiotic nodule development. Nature 417:962–966

Feddermann N, Muni RR, Zeier T, Stuurman J, Ercolin F, Schorderet M, Reinhardt D (2010) The PAM1 gene of petunia required for intracellular accommodation and morphogenesis of arbuscular mycorrhizal fungi encodes a homologue of VAPYRIN. Plant J 64:470–481

Genre A, Chabaud M, Balzergue C, Puech-Pages V, Novero M, Rey T, Fournier J, Rochange S, Becard G, Bonfante P, Barker DG (2013) Short-chain chitin oligomers from arbuscular mycorrhizal fungi trigger nuclear Ca^{2+} spiking in *Medicago truncatula* roots and their production is enhanced by strigolactone. New Phytol 198:190–202

Gimenez-Ibanez S, Ntoukakis V, Rathjen JP (2009) The LysM receptor kinase CERK1 mediates bacterial perception in *Arabidopsis*. Plant Signal Behav 4:539–541

Gleason C, Chaudhuri S, Yang T, Munoz A, Poovaiah BW, Oldroyd GE (2006) Nodulation independent of rhizobia induced by a calcium-activated kinase lacking autoinhibition. Nature 441:1149–1152

Gobbato E, Marsh JF, Vernie T, Wang E, Maillet F, Kim J, Miller JB, Sun J, Bano SA, Ratet P, Mysore KS, Denarie J, Schultze M, Oldroyd GE (2012) A GRAS-type transcription factor with a specific function in mycorrhizal signaling. Curr Biol 22:2236–2241

Granqvist E, Wysham D, Hazledine S, Kozlowski W, Sun J, Charpentier M, Martins TV, Haleux P, Tsaneva-Atanasova K, Downie JA, Oldroyd GE, Morris RJ (2012) Buffering capacity explains signal variation in symbiotic calcium oscillations. Plant Physiol 160:2300–2310

Groth M, Takeda N, Perry J, Uchida H, Draxl S, Brachmann A, Sato S, Tabata S, Kawaguchi M, Wang TL, Parniske M (2010) NENA, a *Lotus japonicus* homolog of Sec13, is required for rhizodermal infection by arbuscular mycorrhiza fungi and rhizobia but dispensable for cortical endosymbiotic development. Plant Cell 22:2509–2526

Gust AA, Willmann R, Desaki Y, Grabherr HM, Nurnberger T (2012) Plant LysM proteins: modules mediating symbiosis and immunity. Trends Plant Sci 17:495–502

Hayashi T, Banba M, Shimoda Y, Kouchi H, Hayashi M, Imaizumi-Anraku H (2010) A dominant function of CCaMK in intracellular accommodation of bacterial and fungal endosymbionts. Plant J 63:141–154

Heckmann AB, Lombardo F, Miwa H, Perry JA, Bunnewell S, Parniske M, Wang TL, Downie JA (2006) *Lotus japonicus* nodulation requires two GRAS domain regulators one of which is functionally conserved in a non-legume. Plant Physiol 142:1739–1750

Hirsch S, Kim J, Munoz A, Heckmann AB, Downie JA, Oldroyd GED (2009) GRAS proteins form a DNA binding complex to induce gene expression during nodulation signaling in Medicago truncatula. Plant Cell 21:545–557

Imaizumi-Anraku H, Takeda N, Charpentier M, Perry J, Miwa H, Umehara Y, Kouchi H, Murakami Y, Mulder L, Vickers K, Pike J, Downie JA, Wang T, Sato S, Asamizu E, Tabata S, Yoshikawa M, Murooka Y, Wu GJ, Kawaguchi M, Kawasaki S, Parniske M, Hayashi M (2005) Plastid proteins crucial for symbiotic fungal and bacterial entry into plant roots. Nature 433:527–531

Ivanov S, Fedorova EE, Limpens E, De Mita S, Genre A, Bonfante P, Bisseling T (2012) Rhizobium-legume symbiosis shares an exocytotic pathway required for arbuscule formation. Proc Natl Acad Sci USA 109:8316–8321

Journet EP, El-Gachtouli N, Vernoud V, de Billy F, Pichon M, Dedieu A, Arnould C, Morandi D, Barker DG, Gianinazzi-Pearson V (2001) *Medicago truncatula* ENOD11: a novel RPRP-encoding early nodulin gene expressed during mycorrhization in arbuscule-containing cells. Mol Plant Microbe Interact 14:737–748

Kalo P, Gleason C, Edwards A, Marsh J, Mitra RM, Hirsch S, Jakab J, Sims S, Long SR, Rogers J, Kiss GB, Downie JA, Oldroyd GED (2005) Nodulation signaling in legumes requires NSP2, a member of the GRAS family of transcriptional regulators. Science 308:1786–1789

Kanamori N, Madsen LH, Radutoiu S, Frantescu M, Quistgaard EM, Miwa H, Downie JA, James EK, Felle HH, Haaning LL, Jensen TH, Sato S, Nakamura Y, Tabata S, Sandal N, Stougaard J (2006) A nucleoporin is required for induction of Ca^{2+} spiking in legume nodule development and essential for rhizobial and fungal symbiosis. Proc Natl Acad Sci USA 103:359–364

Ke D, Fang Q, Chen C, Zhu H, Chen T, Chang X, Yuan S, Kang H, Ma L, Hong Z, Zhang Z (2012) The small GTPase ROP6 interacts with NFR5 and is involved in nodule formation in *Lotus japonicus*. Plant Physiol 159:131–143

Kevei Z, Lougnon G, Mergaert P, Horvath GV, Kereszt A, Jayaraman D, Zaman N, Marcel F, Regulski K, Kiss

GB, Kondorosi A, Endre G, Kondorosi E, Ane JM (2007) 3-hydroxy-3-methylglutaryl coenzyme a reductase 1 interacts with NORK and is crucial for nodulation in *Medicago truncatula*. Plant Cell 19:3974–3989

Kistner C, Parniske M (2002) Evolution of signal transduction in intracellular symbiosis. Trends Plant Sci 7:511–518

Kistner C, Winzer T, Pitzschke A, Mulder L, Sato S, Kaneko T, Tabata S, Sandal N, Stougaard J, Webb KJ, Szczyglowski K, Parniske M (2005) Seven *Lotus japonicus* genes required for transcriptional reprogramming of the root during fungal and bacterial symbiosis. Plant Cell 17:2217–2229

Konishi M, Yanagisawa S (2013) *Arabidopsis* NIN-like transcription factors have a central role in nitrate signalling. Nature Commun 4:1617

Kosuta S, Hazledine S, Sun J, Miwa H, Morris RJ, Downie JA, Oldroyd GE (2008) Differential and chaotic calcium signatures in the symbiosis signaling pathway of legumes. Proc Natl Acad Sci USA 105:9823–9828

Kosuta S, Held M, Hossain MS, Morieri G, Macgillivary A, Johansen C, Antolin-Llovera M, Parniske M, Oldroyd GE, Downie AJ, Karas B, Szczyglowski K (2011) *Lotus japonicus* symRK-14 uncouples the cortical and epidermal symbiotic program. Plant J 67:929–940

Kwon C, Neu C, Pajonk S, Yun HS, Lipka U, Humphry M, Bau S, Straus M, Kwaaitaal M, Rampelt H, El Kasmi F, Jurgens G, Parker J, Panstruga R, Lipka V, Schulze-Lefert P (2008) Co-option of a default secretory pathway for plant immune responses. Nature 451:835–840

Lefebvre B, Timmers T, Mbengue M, Moreau S, Herve C, Toth K, Bittencourt-Silvestre J, Klaus D, Deslandes L, Godiard L, Murray JD, Udvardi MK, Raffaele S, Mongrand S, Cullimore J, Gamas P, Niebel A, Ott T (2010) A remorin protein interacts with symbiotic receptors and regulates bacterial infection. Proc Natl Acad Sci USA 107:2343–2348

Levy J, Bres C, Geurts R, Chalhoub B, Kulikova O, Duc G, Journet EP, Ane JM, Lauber E, Bisseling T, Denarie J, Rosenberg C, Debelle F (2004) A putative Ca^{2+} and calmodulin-dependent protein kinase required for bacterial and fungal symbioses. Science 303:1361–1364

Liao J, Singh S, Hossain MS, Andersen SU, Ross L, Bonetta D, Zhou Y, Sato S, Tabata S, Stougaard J, Szczyglowski K, Parniske M (2012) Negative regulation of CCaMK is essential for symbiotic infection. Plant J 72:572–584

Liu J, Blaylock LA, Endre G, Cho J, Town CD, VandenBosch KA, Harrison MJ (2003) Transcript profiling coupled with spatial expression analyses reveals genes involved in distinct developmental stages of an arbuscular mycorrhizal symbiosis. Plant Cell 15:2106–2123

Liu TT, Liu ZX, Song CJ, Hu YF, Han ZF, She J, Fan FF, Wang JW, Jin CW, Chang JB, Zhou JM, Chai JJ (2012) Chitin-induced dimerization activates a plant immune receptor. Science 336:1160–1164

Lohmann GV, Shimoda Y, Nielsen MW, Jorgensen FG, Grossmann C, Sandal N, Sorensen K, Thirup S, Madsen LH, Tabata S, Sato S, Stougaard J, Radutoiu S (2010) Evolution and regulation of the *Lotus japonicus* LysM receptor gene family. Mol Plant Microbe Interact 23:510–521

Lopez-Lara IM, van den Berg JD, Thomas-Oates JE, Glushka J, Lugtenberg BJ, Spaink HP (1995) Structural identification of the lipo-chitin oligosaccharide nodulation signals of Rhizobium loti. Mol Microbiol 15:627–638

Madsen EB, Antolin-Llovera M, Grossmann C, Ye JY, Vieweg S, Broghammer A, Krusell L, Radutoiu S, Jensen ON, Stougaard J, Parniske M (2011) Autophosphorylation is essential for the in vivo function of the *Lotus japonicus* nod factor receptor 1 and receptor-mediated signalling in cooperation with nod factor receptor 5. Plant J 65:404–417

Madsen EB, Madsen LH, Radutoiu S, Olbryt M, Rakwalska M, Szczyglowski K, Sato S, Kaneko T, Tabata S, Sandal N, Stougaard J (2003) A receptor kinase gene of the LysM type is involved in legume perception of rhizobial signals. Nature 425:637–640

Madsen LH, Tirichine L, Jurkiewicz A, Sullivan JT, Heckmann AB, Bek AS, Ronson CW, James EK, Stougaard J (2010) The molecular network governing nodule organogenesis and infection in the model legume *Lotus japonicus*. Nat Commun 1:1–12

Maillet F, Poinsot V, Andre O, Puech-Pages V, Haouy A, Gueunier M, Cromer L, Giraudet D, Formey D, Niebel A, Martinez EA, Driguez H, Becard G, Denarie J (2011) Fungal lipochitooligosaccharide symbiotic signals in arbuscular mycorrhiza. Nature 469:58–63

Markmann K, Giczey G, Parniske M (2008) Functional adaptation of a plant receptor-kinase paved the way for the evolution of intracellular root symbioses with bacteria. PLoS Biol 6(3):e68. doi:10.1371/journal.pbio.0060068

Marsh JF, Rakocevic A, Mitra RM, Brocard L, Sun J, Eschstruth A, Long SR, Schultze M, Ratet P, Oldroyd GE (2007) *Medicago truncatula* NIN is essential for rhizobial-independent nodule organogenesis induced by autoactive calcium/calmodulin-dependent protein kinase. Plant Physiol 144:324–335

Meinema AC, Laba JK, Hapsari RA, Otten R, Mulder FA, Kralt A, van den Bogaart G, Lusk CP, Poolman B, Veenhoff LM (2011) Long unfolded linkers facilitate membrane protein import through the nuclear pore complex. Science 333:90–93

Messinese E, Mun JH, Yeun LH, Jayaraman D, Rouge P, Barre A, Lougnon G, Schornack S, Bono JJ, Cook DR, Ane JM (2007) A novel nuclear protein interacts with the symbiotic DMI3 calcium- and calmodulin-dependent protein kinase of Medicago truncatula. Mol Plant Microbe Interact 20:912–921

Mitra RM, Gleason CA, Edwards A, Hadfield J, Downie JA, Oldroyd GE, Long SR (2004) A Ca^{2+}/calmodulin-dependent protein kinase required for symbiotic nodule development: gene identification by transcript-based cloning. Proc Natl Acad Sci USA 101:4701–4705

Miwa H, Sun J, Oldroyd GED, Downie JA (2006) Analysis of nod-factor-induced calcium signaling in root hairs of

symbiotically defective mutants of *Lotus japonicus*. Mol Plant Microbe Interact 19:914–923

Miya A, Albert P, Shinya T, Desaki Y, Ichimura K, Shirasu K, Narusaka Y, Kawakami N, Kaku H, Shibuya N (2007) CERK1, a LysM receptor kinase is essential for chitin elicitor signaling in *Arabidopsis*. Proc Natl Acad Sci USA 104:19613–19618

Mulligan JT, Long SR (1985) Induction of rhizobium meliloti nodC expression by plant exudate requires nodD. Proc Natl Acad Sci USA 82:6609–6613

Murakami Y, Miwa H, Imaizumi-Anraku H, Kouchi H, Downie JA, Kawaguchi M, Kawasaki S (2006) Positional cloning identifies *Lotus japonicus* NSP2, a putative transcription factor of the GRAS family required for NIN and ENOD40 gene expression in nodule initiation. DNA Res 13:255–265

Murray JD, Muni RR, Torres-Jerez I, Tang Y, Allen S, Andriankaja M, Li G, Laxmi A, Cheng X, Wen J, Vaughan D, Schultze M, Sun J, Charpentier M, Oldroyd G, Tadege M, Ratet P, Mysore KS, Chen R, Udvardi MK (2011) Vapyrin a gene essential for intracellular progression of arbuscular mycorrhizal symbiosis is also essential for infection by rhizobia in the nodule symbiosis of *Medicago truncatula*. Plant J 65:244–252

Nakagawa T, Kaku H, Shimoda Y, Sugiyama A, Shimamura M, Takanashi K, Yazaki K, Aoki T, Shibuya N, Kouchi H (2011) From defense to symbiosis: limited alterations in the kinase domain of LysM receptor-like kinases are crucial for evolution of legume-rhizobium symbiosis. Plant J 65:169–180

Oldroyd GE, Long SR (2003) Identification and characterization of nodulation-signaling pathway 2, a gene of *Medicago truncatula* involved in nod actor signaling. Plant Physiol 131:1027–1032

Oldroyd GE, Murray JD, Poole PS, Downie JA (2011) The rules of engagement in the legume-rhizobial symbiosis. Annu Rev Genet 45:119–144

Peters NK, Frost JW, Long SR (1986) A plant flavone luteolin induces expression of rhizobium-meliloti nodulation genes. Science 233:977–980

Pumplin N, Mondo SJ, Topp S, Starker CG, Gantt JS, Harrison MJ (2010) *Medicago truncatula* vapyrin is a novel protein required for arbuscular mycorrhizal symbiosis. Plant J 61:482–494

Radutoiu S, Madsen LH, Madsen EB, Felle HH, Umehara Y, Gronlund M, Sato S, Nakamura Y, Tabata S, Sandal N, Stougaard J (2003) Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases. Nature 425:585–592

Radutoiu S, Madsen LH, Madsen EB, Jurkiewicz A, Fukai E, Quistgaard EM, Albrektsen AS, James EK, Thirup S, Stougaard J (2007) LysM domains mediate lipochitin-oligosaccharide recognition and Nfr genes extend the symbiotic host range. EMBO J 26:3923–3935

Reddy DMR, Schorderet S, Feller U, Reinhardt D (2007) A petunia mutant affected in intracellular accommodation and morphogenesis of arbuscular mycorrhizal fungi. Plant J 51:739–750

Riely BK, Lougnon G, Ane JM, Cook DR (2007) The symbiotic ion channel homolog DMI1 is localized in the nuclear membrane of *Medicago truncatula* roots. Plant J 49:208–216

Rodpothong P, Sullivan JT, Songsrirote K, Sumpton D, Cheung KW, Thomas-Oates J, Radutoiu S, Stougaard J, Ronson CW (2009) Nodulation gene mutants of *Mesorhizobium loti* R7A-nodZ and nolL mutants have host-specific phenotypes on *Lotus* spp. Mol Plant Microbe Interact 22:1546–1554

Routray P, Miller JB, Du L, Oldroyd G, Poovaiah BW (2013) Phosphorylation of S344 in the calmodulin-binding domain negatively affects CCaMK function during bacterial and fungal symbioses. Plant J 76 (2):287–296

Sacchettini JC, Baum LG, Brewer CF (2001) Multivalent protein-carbohydrate interactions. A new paradigm for supermolecular assembly and signal transduction. Biochemistry 40:3009–3015

Saito K, Yoshikawa M, Yano K, Miwa H, Uchida H, Asamizu E, Sato S, Tabata S, Imaizumi-Anraku H, Umehara Y, Kouchi H, Murooka Y, Szczyglowski K, Downie JA, Parniske M, Hayashi M, Kawaguchi M (2007) NUCLEOPORIN85 is required for calcium spiking fungal and bacterial symbioses and seed production in *Lotus japonicus*. Plant Cell 19:610–624

Sanchez-Vallet A, Saleem-Batcha R, Kombrink A, Hansen G, Valkenburg DJ, Thomma BP, Mesters JR (2013) Fungal effector Ecp6 outcompetes host immune receptor for chitin binding through intrachain LysM dimerization. Elife 2:e00790

Sathyanarayanan PV, Cremo CR, Poovaiah BW (2000) Plant Chimeric Ca^{2+}/Calmodulin-dependent Protein Kinase: role of the neural visinin-like domain in regulating autophosphorylation and calmodulin affinity. J Biol Chem 275:30417–30422

Schauser L, Roussis A, Stiller J, Stougaard J (1999) A plant regulator controlling development of symbiotic root nodules. Nature 402:191–195

Schlaman HR, Horvath B, Vijgenboom E, Okker RJ, Lugtenberg BJ (1991) Suppression of nodulation gene expression in bacteroids of rhizobium leguminosarum biovar viciae. J Bacteriol 173:4277–4287

Sieberer BJ, Chabaud M, Timmers AC, Monin A, Fournier J, Barker DG (2009) A nuclear-targeted cameleon demonstrates intranuclear Ca^{2+} spiking in *Medicago truncatula* root hairs in response to rhizobial nodulation factors. Plant Physiol 151:1197–1206

Sieberer BJ, Chabaud M, Fournier J, Timmers AC, Barker DG (2012) A switch in Ca^{2+} spiking signature is concomitant with endosymbiotic microbe entry into cortical root cells of *Medicago truncatula*. Plant J 69:822–830

Singh S, Parniske M (2012) Activation of calcium- and calmodulin-dependent protein kinase (CCaMK), the central regulator of plant root endosymbiosis. Curr Opin Plant Biol 15:444–453

Singh S, Katzer K, Lambert J, Cerri M, Parniske M (2014) CYCLOPS, a DNA binding transcriptional

activator orchestrates symbiotic root nodule development. Cell Host Microbe 15(2):139–152

Smit P, Raedts J, Portyanko V, Debelle F, Gough C, Bisseling T, Geurts R (2005) NSP1 of the GRAS protein family is essential for rhizobial nod factor-induced transcription. Science 308:1789–1791

Soyano T, Kouchi H, Hirota A, Hayashi M (2013) Nodule inception directly targets NF-Y subunit genes to regulate essential processes of root nodule development in *Lotus japonicus*. PLoS Genet 9:e1003352. doi:10.1371/journal.pgen.1003352

Spaink HP, Okker RJ, Wijffelman CA, Tak T, Goosen-de Roo L, Pees E, van Brussel AA, Lugtenberg BJ (1989) Symbiotic properties of rhizobia containing a flavonoid-independent hybrid nodD product. J Bacteriol 171:4045–4053

Spaink HP, Sheeley DM, van Brussel AA, Glushka J, York WS, Tak T, Geiger O, Kennedy EP, Reinhold VN, Lugtenberg BJ (1991) A novel highly unsaturated fatty acid moiety of lipo-oligosaccharide signals determines host specificity of Rhizobium. Nature 354:125–130

Stracke S, Kistner C, Yoshida S, Mulder L, Sato S, Kaneko T, Tabata S, Sandal N, Stougaard J, Szczyglowski K, Parniske M (2002) A plant receptor-like kinase required for both bacterial and fungal symbiosis. Nature 417:959–962

Swainsbury DJ, Zhou L, Oldroyd GE, Bornemann S (2012) Calcium ion binding properties of *Medicago truncatula* calcium/calmodulin-dependent protein kinase. Biochemistry 51:6895–6907

Takezawa D, Ramachandiran S, Paranjape V, Poovaiah BW (1996) Dual regulation of a chimeric plant serine/threonine kinase by calcium and calcium/calmodulin. J Biol Chem 271:8126–8132

Timmers ACJ, Auriac MC, de Billy F, Truchet G (1998) Nod factor internalization and microtubular cytoskeleton changes occur concomitantly during nodule differentiation in alfalfa. Development 125:339–349

Tirichine L, Imaizumi-Anraku H, Yoshida S, Murakami Y, Madsen LH, Miwa H, Nakagawa T, Sandal N, Albrektsen AS, Kawaguchi M, Downie A, Sato S, Tabata S, Kouchi H, Parniske M, Kawasaki S, Stougaard J (2006) Deregulation of a Ca^{2+}/calmodulin-dependent kinase leads to spontaneous nodule development. Nature 441:1153–1156

Tirichine L, Sandal N, Madsen LH, Radutoiu S, Albrektsen AS, Sato S, Asamizu E, Tabata S, Stougaard J (2007) A gain-of-function mutation in a cytokinin receptor triggers spontaneous root nodule organogenesis. Science 315:104–107

Tóth K, Stratil TF, Madsen EB, Ye J, Popp C, Antolin-Llovera M, Grossmann C, Jensen ON, Schussler A, Parniske M, Ott T (2012) Functional domain analysis of the remorin protein LjSYMREM1 in *Lotus japonicus*. PLoS ONE 7:e30817. doi:10.1371/journal.pone.0030817

Truchet G, Roche P, Lerouge P, Vasse J, Camut S, de Billy F, Prome J-C, Denarie J (1991) Sulphated lipo-oligosaccharide signals of Rhizobium meliloti elicit root nodule organogenesis in alfalfa. Nature 351:670–673

Venkateshwaran M, Cosme A, Han L, Banba M, Satyshur KA, Schleiff E, Parniske M, Imaizumi-Anraku H, Ane JM (2012) The recent evolution of a symbiotic ion channel in the legume family altered ion conductance and improved functionality in calcium signalling. Plant Cell 24:2528–2545

Walther TC, Alves A, Pickersgill H, Loiodice I, Hetzer M, Galy V, Hulsmann BB, Kocher T, Wilm M, Allen T, Mattaj IW, Doye V (2003) The conserved Nup107-160 complex is critical for nuclear pore complex assembly. Cell 113:195–206

Wan J, Zhang XC, Neece D, Ramonell KM, Clough S, Kim SY, Stacey MG, Stacey G (2008) A LysM receptor-like kinase plays a critical role in chitin signaling and fungal resistance in *Arabidopsis*. Plant Cell 20:471–481

Wang E, Schornack S, Marsh JF, Gobbato E, Schwessinger B, Eastmond P, Schultze M, Kamoun S, Oldroyd GE (2012) A common signaling process that promotes mycorrhizal and oomycete colonization of plants. Curr Biol 22:2242–2246

Yano K, Yoshida S, Muller J, Singh S, Banba M, Vickers K, Markmann K, White C, Schuller B, Sato S, Asamizu E, Tabata S, Murooka Y, Perry J, Wang TL, Kawaguchi M, Imaizumi-Anraku H, Hayashi M, Parniske M (2008) CYCLOPS, a mediator of symbiotic intracellular accommodation. Proc Natl Acad Sci USA 105:20540–20545

Zhu H, Riely BK, Burns NJ, Ane JM (2006) Tracing nonlegume orthologs of legume genes required for nodulation and arbuscular mycorrhizal symbioses. Genetics 172:2491–2499

Zhukov V, Radutoiu S, Madsen LH, Rychagova T, Ovchinnikova E, Borisov A, Tikhonovich I, Stougaard J (2008) The pea Sym37 receptor kinase gene controls infection-thread initiation and nodule development. Mol Plant Microbe Interact 21:1600–1608

Genes for Autoregulation of Nodulation

7

Masayoshi Kawaguchi

Abstract

The phenomenon in which developed nodules or nodule primordia suppress the emergence of further nodules in legumes is termed autoregulation of nodulation (AON) (Nutman in Ann Bot 16:79–101, 1952). AON consists of two presumptive long-distance signal molecules involving roots and shoots (Caetano-Anollés and Gresshoff in Annu Rev Microbiol 45:345–382, 1991; Oka-Kira and Kawaguchi in Curr Opin Plant Biol 9:496–502, 2006) (Fig. 7.1) and is also related to the repression of nodules by nitrogen compounds such as nitrate. AON is of great interest with respect to morphological plasticity of plant organogenesis as well as long-distance signalling. This chapter is devoted to the current knowledge of AON mainly in *Lotus japonicus*..

7.1 Hypernodulating Mutants and HAR1

Mutants defective in AON are expected to lose the negative feedback control and consequently display a hypernodulating phenotype with a wider nodulation zone in their roots. Supernodulating or hypernodulating mutants such as *Glycine max nts/nark*, *Lotus japonicus har1*, *Medicago truncatula sunn* and *Pisum sativum sym29* allow nodules to develop in almost entire root regions (Carroll et al. 1985; Sagan and Duc 1996; Schauser et al. 1998; Szczyglowski et al. 1998; Wopereis et al. 2000; Kawaguchi et al. 2002; Oka-Kira et al. 2005; Schnabel et al. 2005). Interestingly, reciprocal grafting experiments using hypernodulating mutants and wild-type plants have shown that the nodulation phenotype in roots is determined by the shoot genotype (Delves et al. 1986). Very similar results were obtained in other leguminous species. These findings indicate that the mutated genes products function in the shoot and that the feedback regulation of nodulation requires long-distance signalling between the shoots and the roots. Positional cloning has revealed that *HAR1* and its orthologs encode an LRR receptor-like kinase and their functions are conserved among other legume species (Krusell et al. 2002; Nishimura et al. 2002; Schnabel et al. 2005; Searle et al. 2003). Surprisingly, among all receptor-like kinases in *Arabidopsis,* the *HAR1* gene is most similar to *CLAVATA1* (*CLV1*), which is a key component required for shoot apical meristem

M. Kawaguchi (✉)
Division of Symbiotic Systems, National Institute for Basic Biology, Okazaki, Aichi 444-8585, Japan
e-mail: masayosi@nibb.ac.jp

S. Tabata and J. Stougaard (eds.), *The Lotus japonicus Genome*, Compendium of Plant Genomes,
DOI 10.1007/978-3-662-44270-8_7,

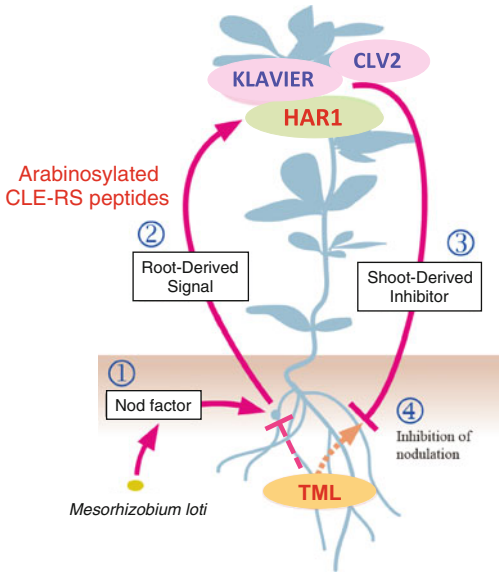


Fig. 1 Schematic illustration of a model for HAR1-, KLV, CLV2 and TML-mediated autoregulation of nodulation (AON). *1* Perception of the rhizobial Nod factor initiates the production of a long-distance inhibitor termed the root-derived signal. *2* Arabinosylated CLE-RS1 and CLE-RS2 peptides are transported to the shoot, and *3* activate the production of the shoot-derived signal. HAR1, KLV and CLV2 mediate this process. *4* The shoot-derived signal(s) is translocated to the root and negatively regulates nodulation via TML. TML F-box protein is a root factor acting at the final stage of AON

(SAM) maintenance via cell–cell communication (Clark et al. 1997). *CLV1* is specifically expressed in the centre of SAM, whereas *HAR1* and legume orthologs are expressed in various organs, such as leaves, stems and roots, but significantly suppressed in the shoot apex (Krusell et al. 2002; Nishimura et al. 2002; Schnabel et al. 2005; Searle et al. 2003). Further expression analysis using a GUS reporter gene driven by the *HAR1* native promoter has elucidated that *HAR1* is expressed predominantly in the phloem tissues of leaves and stems (Nontachaiyapoom et al. 2007). This phloem-specific expression of *HAR1* seems to make sense because phloem tissues function as the conduit for long-distance communication between distantly located organs. As the HAR1 ligand is presumed to act as a root-derived long-distance mobile signal, the next important question is 'What is a ligand for HAR1?'

7.2 Root-Derived Signal and Arabinosylated CLE Peptide

The existence of the root-derived signal, also known as 'Q', was first proposed by grafting and split root experiments using soybean *nts* mutants (Caetano-Anollés and Gresshoff 1990). It is thought that the root-derived signal is generated in roots via early symbiotic signalling activated by Nod factor secreted from rhizobia and then translocated to the shoot. However, since the early 1990s, the chemical nature of the root-derived signal had remained unknown. In 2009, Okamoto et al. (2009) first reported a candidate through an *in silico* search of the *L. japonicus* genome database. They found that *L. japonicus CLE* genes (*CLE-RS1* and *RS2*) are specifically and rapidly induced in the roots in response to its symbiotic bacteria *Mesorhizobium loti* and that the overexpression of *CLE-RS1/2* genes drastically reduced or abolished nodulation in a HAR1-dependent manner. Of particular importance is that this inhibitory effect travels systemically from transformed roots to untransformed roots (Okamoto et al. 2009). Similar *CLE* peptide genes showing local and systemic suppression of nodulation have been reported in *Medicago* and *Glycine* (Mortier et al. 2010, 2011; Reid et al. 2011). However, application of synthesized CLE peptides deduced from *CLE-RS1/2* gene structures did not suppress nodulation even at a micromolar concentration, implying that some specific posttranslational modification is required for the biological activity of CLE-RS peptides.

Recently, Okamoto et al. determined the mature structure of CLE-RS2 peptide by the overexpression of *CLE-RS2* genes using Arabidopsis submerged culture and *L. japonicus* hairy root culture systems (Okamoto et al. 2013). Through nano-liquid chromatography–mass spectrometry (nano-LC-MS) and nano-liquid chromatography–tandem mass spectrometry (nano-LC-MS/MS) analyses of peptides diffused into culture media, they identified a 13-amino-acid CLE-RS2 peptide modified with three or more residues of arabinose. Chemically synthesized arabinosylated CLE-RS peptides bind

directly to an LRR receptor domain of HAR1 and significantly suppress nodulation at nano-molar concentrations when applied from the cut surface of the cotyledon. Furthermore, the arabinosylated CLE-RS2 peptide can be detected in xylem sap collected from the soybean shoots when CLE-RS2 is specifically expressed in the soybean hairy roots (Okamoto et al. 2013). These results suggest that arabinosylated CLE-RS2 peptide is the long sought after root-derived signal for the onset of AON. On the other hand, the mature structure of CLE-RS1 peptide remains unknown.

7.3 KLAVIER and CLV2 as Shoot Factors

L. japonicus mutant *har1* and other legume mutants carrying mutations in the *HAR1* orthologs do not exhibit any *clv1*-like shoot phenotypes such as fasciation by enlargement of SAM. This finding indicates that in legumes, *CLV1* orthologs play a specific role in the systemic regulation of nodulation but not in the feedback control of SAM development. Although the genes responsible for the regulation of SAM size were not identified in legumes until relatively recently, several hypernodulating mutants were known to exhibit *clv1*-like phenotypes. In *L. japonicus*, *klavier* (*klv*) exhibits not only a typical hypernodulation phenotype, but also *clv*-like phenotypes such as fasciated stems, an increased number of flowers per peduncle and bifurcated pistils (Oka-Kira et al. 2005; Miyazawa et al. 2010). Grafting experiments using *klv* shoots and wild-type roots have demonstrated that KLV functions in the shoots to control nodule numbers, as is the case for HAR1 (Oka-kira et al. 2005). Similarly, the *pea sym28* mutant also exhibits shoot-regulated hypernodulation, fasciated stems and an increased number of flowers (Sagan and Duc 1996). These pleiotropic phenotypes suggest a potential evolutionary link between SAM maintenance and AON in legumes.

Positional cloning identified *KLV*, which encodes an LRR receptor-like kinase (Miyazawa et al. 2010). More importantly, *KLV* is most closely related to Arabidopsis *RPK2*, which is essential for SAM maintenance (Kinoshita et al. 2010). A double-mutant analysis indicates that *KLV* and *HAR1* act in the same genetic pathway that governs the long-distance control of nodulation. *KLV* is predominantly expressed in the vascular tissues of mature leaves, as is *HAR1*. The biochemical analyses actually demonstrated that KLV physically interacts with HAR1 in *Nicotiana benthamiana* (Miyazawa et al. 2010), suggesting that the potential KLV–HAR1 receptor complex systemically regulates nodulation by receiving the root-derived arabinosylated CLE peptides.

In the reproductive phase, pea *sym28* shoots develop additional flowers, fascinated stems and abnormal phyllotaxis as well as hypernodulation (Sagan and Duc 1996). Recently, molecular genetic approaches identified the causal gene *Sym28* (Krusell et al. 2011). *Sym28* encodes an LRR receptor-like protein and is closely related to *Arabidopsis CLAVATA2* (CLV2) that lacks a kinase domain and forms a complex with CORYNE/SOR2, a membrane-associated kinase that regulates the SAM maintenance in Arabidopsis (Bleckmann et al. 2010; Guo et al. 2010; Zhu et al. 2010). On the other hand, downregulation of the *L. japonicus Clv2* gene by RNAi resulted in enhanced nodulation (Krusell et al. 2011). Thus, as with KLV, legume CLV2 receptor-like proteins appear to be involved in AON via long-distance signalling as well as the maintenance of SAM development.

7.4 TOO MUCH LOVE as a Root Factor

In contrast to shoot factors, *Pisum sativum nod3*, *M. truncatula sickle* and *rdn1* mutants belong to the category of the root genotype-determined hypernodulating mutants (Postoma et al. 1998; Penmetsa and Cook 1997; Schnabel et al. 2011). Furthermore, *L. japonicus rdh1*, *too much love* (*tml*), and *plenty,* have been also isolated as root-determined hypernodulating mutants (Ishikawa et al. 2008; Magori et al. 2009; Yoshida et al. 2010). As *rdh1* turned out to be allelic to *tml* it has been designated *tml-4* (Takahara et al. 2013). Some of these mutants are thought to be deficient

in the generation of root-derived signals or perception of shoot-derived signals if they are responsible for shoot factor-mediated long-distance control of nodulation. Double-mutant analyses suggested that TML acts in the same genetic pathway as HAR1, while PLENTY is likely to act in a different pathway (Yoro et al. unpublished data). Inverted-Y grafting experiments suggested that TML is likely to function downstream of HAR1, possibly as a receptor or a mediator of the shoot-derived inhibitor (Magori et al. 2009; Magori and Kawaguchi 2009). Thus, it is expected that TML could provide not only a cue to unveil the as yet unidentified shoot-derived inhibitor but also a molecular link with components of the Nod factor signalling pathway.

Most recently, the map-based cloning and deep sequencing by a next generation sequencer identified a candidate of the *TML* gene (Takahara et al. 2013). The knock-down of the candidate in hairy roots of *L. japonicus* resulted in a dramatic increase in nodulation. The putative *TML* gene encodes a Kelch repeat-containing F-box protein with two nuclear localizing signals (Takahara et al. 2013). In *A. thaliana,* there are more than 100 Kelch repeat-containing F-box proteins but their function remains largely unknown, with the exception of ZTL, FKF1 and LKP2 that act as blue light receptors which are critical for light-controlled plant growth and development (Ito et al. 2012; Schumann et al. 2011). FKF1 physically interacts with a Dof transcription factor (CDF1) and mediates proteasome degradation of a CDF1 protein that directly represses *CONSTANS* expression (Imaizumi et al. 2005). These findings led us to speculate that TML functions as a receptor of shoot-derived inhibitors and represses nodule development by the degradation of a transcription factor that constitutes the Nod factor signalling pathway. The elucidation of the molecular functions of TML is now required.

7.5 Conclusion

In *L. japonicus*, HAR1, KLV and CLV2 act in the shoot, whereas TML acts in the root in the context of AON (Fig. 7.1). Arabinosylated CLE-RS peptides are expected to link both organs via long-distance communication. On the other hand, a second long-distance signal, termed the shoot-derived inhibitor, is essential for AON. The shoot-derived inhibitor is synthesized in the shoots and is translocated to the roots where it inhibits further nodule development. The perception of the root-derived signal by NARK/HAR1/SUNN/SYM28, KLAVIER and CLV2 is then thought to activate the production of the shoot-derived inhibitor, but the chemical nature of the long-distance signal of the shoot-derived inhibitor is currently unknown. To characterize the shoot-derived inhibitor, Gresshoff and associates have developed a novel feeding bioassay which involves feeding (or introduces) aqueous leaf extracts directly into the petiole of hypernodulating and supernodulating mutant plants of *Glycine max* (soybean) (Lin et al. 2010). They have found that suppression activity is inoculation dependent and Nod factor dependent, required GmNARK activity, and was heat-, proteinase K- and ribo-nuclease A-resistant. On the other hand, Yamaya and Arima (2010) also succeeded in detecting nodulation suppression activity, but their results differ partly. They showed that the suppressive activity of nodulation is constantly detected irrespective of *B. japonicum* inoculation. Although there is some information about the shoot-derived inhibitor in *L. japonicus*, further studies such as metabolome and transcriptome analyses will be needed to discover the signal molecule(s).

References

Bleckmann A, Weidtkamp-Peter S, Seidel CA, Simon R (2010) Stem cell signaling in *Arabidopsis* requires CRN to localize CLV2 to the plasma membrane. Plant Physiol 152:166–176

Caetano-Anollés G, Gresshoff PM (1990) Early induction of feedback regulatory responses governing nodulation in soybean. Plant Sci 71:69–81

Caetano-Anollés G, Gresshoff PM (1991) Plant genetic control of nodulation. Annu Rev Microbiol 45:345–382

Carroll BJ, McNeil DL, Gresshoff PM (1985) Isolation and properties of soybean [*Glycine max* (L.) Merr.] mutants that nodulate in the presence of high nitrate

concentrations. Proc Natl Acad Sci USA 82:4162–4166
Clark SE, Williams RW, Meyerowitz EM (1997) The CLAVATA1 gene encodes a putative receptor kinase that controls shoot and floral meristem size in *Arabidopsis*. Cell 89:575–585
Delves AC, Mathews A, Day DA et al (1986) Regulation of the soybean-*Rhizobium* nodule symbiosis by shoot and root factors. Plant Physiol 82:588–590
Guo Y, Han L, Hymes M et al (2010) CLAVATA2 forms a distinct CLE-binding receptor complex regulating *Arabidopsis* stem cell specification. Plant J 63:889–900
Imaizumi T, Schultz TF, Harmon FG, Ho LA, Kay SA (2005) FKF1 F-box protein mediates cyclic degradation of a repressor of CONSTANS in *Arabidopsis*. Science 309:293–297
Ishikawa K, Yokota K, Li YY et al (2008) Isolation of a novel root-determined hypernodulation mutant *rdh1* of *Lotus japonicus*. Soil Sci Plant Nutr 54:259–263
Ito S, Song YH, Imaizumi T (2012) LOV domain-containing F-box proteins: light-dependent protein degradation modules in *Arabidopsis*. Mol Plant 5:573–582
Kawaguchi M, Imaizumi-Anraku H, Koiwa H et al (2002) Root, root hair, and symbiotic mutants of the model legume *Lotus japonicus*. Mol Plant Microbe Interact 15:17–26
Kinoshita A, Betsuyaku S, Osakabe Y et al (2010) RPK2 is an essential receptor-like kinase that transmits the CLV3 signal in *Arabidopsis*. Development 137:3911–3920
Krusell L, Madsen LH, Sato S et al (2002) Shoot control of root development and nodulation is mediated by a receptor-like kinase. Nature 420:422–426
Krusell L, Sato N, Fukuhara I et al (2011) The *Clavata2* genes of pea and *Lotus japonicus* affect autoregulation of nodulation. Plant J 65:861–871
Lin YH, Ferguson BJ, Kereszt A, Gresshoff PM (2010) Suppression of supernodulation in soybean by a leaf-extracted, NARK- and inoculation- dependent small molecular fraction. New Phytol 185:1074–1086
Magori S, Kawaguchi M (2009) Long-distance control of nodulation: molecules and models. Mol Cells 28:29–34
Magori S, Oka-kira E, Shibata S et al (2009) Too much love, a root regulator associated with the long-distance control of nodulation in *Lotus japonicus*. Mol Plant-Microbe Interact 22:259–268
Miyazawa H, Oka-kira E, Sato N (2010) The receptor-like kinase KLAVIER mediates systemic regulation of nodulation and non-symbiotic shoot development in *Lotus japonicus*. Development 137:4317–4325
Mortier V, den Herder G, Whitford et al (2010) CLE peptides control *Medicago truncatula* nodulation locally and systemically. Plant Physiol 153:222–237
Mortier V, Fenta BA, Martens C et al (2011) Search for nodulation-related CLE genes in the genome of *Glycine max*. J Exp Bot 62:2571–2583
Nishimura R, Hayashi M, Wu GJ et al (2002) HAR1 mediates systemic regulation of symbiotic organ development. Nature 420:426–429
Nontachaiyapoom S, Scott PT, Men AE et al (2007) Promoters of orthologous *Glycine max* and *Lotus japonicus* nodulation autoregulation genes interchangeably drive phloem-specific expression in transgenic plants. Mol Plant-Microbe Interact 20:769–780
Nutman PS (1952) Studies on the physiology of nodule formation. Ann Bot 16:79–101
Oka-Kira E, Kawaguchi M (2006) Long-distance signaling to control root nodule number. Curr Opin Plant Biol 9:496–502
Oka-Kira E, Tateno K, Miura K et al (2005) *klavier (klv)*, a novel hypernodulation mutant of *Lotus japonicus* affected in vascular tissue organization and floral induction. Plant J 44:505–515
Okamoto S, Ohnishi E, Sato S et al (2009) Nod factor/nitrate-induced *CLE* genes that drive HAR1-mediated systemic regulation of nodulation. Plant Cell Physiol 50:67–77
Okamoto S, Shinohara H, Mori T et al (2013) Root-derived CLE glycopeptides control nodulation by direct binding to HAR1 receptor kinase. Nature Commun 4:2191
Penmetsa RV, Cook DR (1997) A legume ethylene-insensitive mutant hyperinfected by its rhizobial symbiont. Science 275:527–530
Postoma JG, Jacobsen E, Feenstra WJ (1988) Three pea mutants with an altered nodulation studied by genetic analysis and grafting. J Plant Physiol 132:424–430
Reid DE, Ferguson BJ, Gresshoff PM (2011) Inoculation- and nitrate-induced CLE peptides of soybean control NARK-dependent nodule formation. Mol Plant-Microbe Interact 24:606–618
Sagan M, Duc G (1996) Sym28 and Sym29, two new genes involved in regulation of nodulation in pea (*Pisum sativum* L.). Symbiosis 20:229–245
Schauser K, Handberg N, Sandal N et al (1998) Symbiotic mutants deficient in nodule establishment identified after T-DNA transformation of *Lotus japonicus*. Mol Gen Genet 259:414–423
Schnabel E, Journet EP, de Carvalho-Niebel et al (2005) The *Medicago truncatula SUNN* gene encodes a CLV1-like leucine-rich repeat receptor kinase that regulates nodule number and root length. Plant Mol Biol 58:809–822
Schnabel EL, Kassaw TK, Smith LS et al (2011) The *ROOT DETERMINED NODULATION1* gene regulates nodule number in roots of *Medicago truncatula* and defines a highly conserved, uncharacterized plant gene family. Plant Physiol 157:328–340
Schumann N, Navarro-Quezada A, Ullrich K et al (2011) Molecular evolution and selection patterns of plant F-box proteins with C-terminal kelch repeats. Plant Physiol 155:835–850
Searle IR, Men AE, Laniya TS et al (2003) Long-distance signaling in nodulation directed by a CLAVATA1-like receptor kinase. Science 299:109–112

Szczyglowski K, Shaw RS, Wopereis J et al (1998) Nodule organogenesis and symbiotic mutants of the model legume *Lotus japonicus*. Mol Plant Microbe Interact 11:684–697

Takahara M, Magori S, Soyano T et al (2013) TOO MUCH LOVE, a novel Kelch repeat-containing F-box protein, functions in the long-distance regulation of the legume-*Rhizobium* symbiosis. Plant Cell Physiol 54:433–447

Wopereis J, Pajuelo E, Dazzo FB et al (2000) Short root mutant of *Lotus japonicus* with a dramatically altered symbiotic phenotype. Plant J 23:97–114

Yamaya H, Arima Y (2010) Shoot-synthesized nodulation-restricting substances are present in the medium-polarity fraction of shoot extracts from wild-type soybean plants. Soil Sci Plant Nutr 56:418–421

Yoshida C, Funayama-Noguchi S, Kawaguchi M (2010) *plenty*, a novel hypernodulation mutant in *Lotus japonicus*. Plant Cell Physiol 51:1425–1435

Zhu Y, Wang Y, Li R et al (2010) Analysis of interactions among the CLAVATA3 receptors reveals a direct interaction between CLAVATA2 and CORYNE in *Arabidopsis*. Plant J 61:223–233

Lotus Genes Involved in Nodule Function and Nitrogen Fixation

8

Norio Suganuma

Abstract

The existence of host plant genes essential for symbiotic nitrogen fixation has been suggested by the isolation of legume Fix^- mutants, the nodules of which are normally endocytosed by rhizobia but exhibit little or no nitrogen-fixing activity. However, it has been difficult to identify the genes responsible for the Fix^- phenotype because of the large genome size of crop legumes. Genome sequencing of the model legume *Lotus japonicus* allowed us to identify the causal genes in Fix^- mutants by map-based cloning. This chapter describes the *Lotus* genes involved in nodule function and nitrogen fixation as identified by forward genetics.

8.1 Introduction

Inside infected host cells, the tip of infection threads releases rhizobia within symbiosomes. In *Lotus japonicus* nodules, these organelle-like units usually contain several rhizobia, which subsequently differentiate into bacteroides and fix atmospheric nitrogen gas. *Mesorhizobium loti* bacteroides, which are compatible with *L. japonicus*, are slightly larger than free-living rhizobia (Suganuma et al. 2003), despite having similar morphology and DNA content, and are able to form colonies on agar plates (Mergaert et al. 2006). This is in contrast to bacteroides within the nodules of galegoid legumes, which are terminally differentiated.

Nitrogen fixation is catalyzed by nitrogenase complex, the components of which are encoded in the rhizobial genome. This enzyme is oxygen-sensitive and has extremely high-energy requirements. Efficient nitrogen fixation is then achieved with the help of an oxygen-binding protein, leghemoglobin, which support the aerobic respiration of rhizobia and protects nitrogenase against inactivation by free oxygen. Fixed nitrogen is assimilated in the host cells of nodules and translocated to shoots for optimal plant growth. In general, nitrogen fixation occurs only when rhizobia are endocytosed in nodule cells, with exception of *Azorhizobium* and the photosynthetic *Bradyrhizobium* (Hakoyama et al. 2009). Most rhizobia do not fix nitrogen under free-living conditions. This suggests that nitrogen fixation by symbiotic rhizobia is strictly controlled by the host plants, although it remains unclear how they regulate the process.

N. Suganuma (✉)
Department of Life Science, Aichi University of Education, Kariya, Aichi 448-8542, Japan
e-mail: nsuganum@auecc.aichi-edu.ac.jp

S. Tabata and J. Stougaard (eds.), *The Lotus japonicus Genome*, Compendium of Plant Genomes,
DOI 10.1007/978-3-662-44270-8_8,

Large-scale gene-expression analysis has revealed that numerous *L. japonicus* genes are expressed differentially during nodule development (Colebatch et al. 2002, 2004; Kouchi et al. 2004). Some of these genes have been put forward as good candidates for the regulation of symbiotic nitrogen fixation, as they are specifically expressed in nodule cells and closely associated with the onset of the process. For instance, the expression of leghemoglobin genes is induced exclusively in infected nodule cells and concurrently with the onset of nitrogen-fixation activity. When leghemoglobin gene expression is suppressed in nodule cells, rhizobial symbiotic nitrogen fixation is lost (Ott et al. 2005), indicating that leghemoglobin genes are essential for symbiotic nitrogen fixation. However, until now, only a few genes have been proven by reverse genetics to be indispensable to nitrogen fixation.

Fix^- mutants are useful for identifying host plant genes that regulate rhizobial symbiotic nitrogen fixation. Fix^- mutants with nodules that are normally endocytosed by rhizobia, but that exhibit little or no nitrogen-fixing activity, have been isolated in *Pisum sativum* and *Medicago sativa* (Vance and Johnson 1983; Kneen et al. 1990; Tsyganov et al. 1998). However, it has been difficult to identify the causal genes in Fix^- mutants because the genomes of these crop legumes are large. Isolation of such Fix^- mutants has been succeeded in the model legume *L. japonicus* (Schauser et al. 1998; Szczyglowski et al. 1998; Kawaguchi et al. 2002), and progress in genome sequencing of *L. japonicus* (Sato et al. 2008) has greatly facilitated the identification of genes responsible for the Fix^- phenotype with the forward genetic approach. To date, five such genes have been identified by map-based cloning in *L. japonicus* (Fig. 8.1). This chapter describes the possible roles and implications of these genes in symbiotic nitrogen fixation. The reader is also referred to recent related reviews (Kouchi et al. 2010; Kouchi 2011; Udvardi and Poole 2013).

8.2 Host Plant Genes Critical to Nitrogen Fixation

8.2.1 Transport Proteins Located on the Symbiosome Membrane: SST1 and SEN1

Rhizobia endocytosed in nodules are surrounded by a symbiosome membrane, which is thought to be derived from the plasma membrane of infected cells. The exchange of metabolites mediated by the symbiosome membrane is essential for the maintenance of rhizobial nitrogen fixation. The carbon sources required by bacteroides for the nitrogenase reaction are provided by host plant cells, and fixed nitrogen is transported from bacteroides to host plant cells through the symbiosome membrane (Udvardi and Day 1997; White et al. 2007). It therefore acts not only as a physical barrier between the host plant cell and the bacteroides, but also as a critical player in nitrogen fixation.

Map-based cloning of the gene responsible for the *L. japonicus* Fix^- mutant has identified a sulfate transporter located in the symbiosome membrane (Wienkoop and Saalbach 2003) that is indispensable for nitrogen-fixation activity (Schauser et al. 1998; Kawaguchi et al. 2002; Krusell et al. 2005). A mutant deficient in the symbiotic sulfate transporter SST1 formed nodules that were normally endocytosed with rhizobia, but that exhibited lower nitrogen-fixation activity than the wild type. Nitrogenase is made up of NifHDK proteins containing three metal clusters, 4S-4Fe cluster, P-cluster, and FeMo-cofactor, all of which contain sulfur. The SST1 protein plays a role in transporting sulfate to the bacteroides from the host plant. Sulfate transporters are also required for plant growth, because sulfur is found in a few amino acids. However, the *SST1* gene is expressed exclusively in nodules (Krusell et al. 2005). Furthermore, the *sst1* mutant, which has nodules that do not express *SST1*, can grow when combined nitrogen is supplied

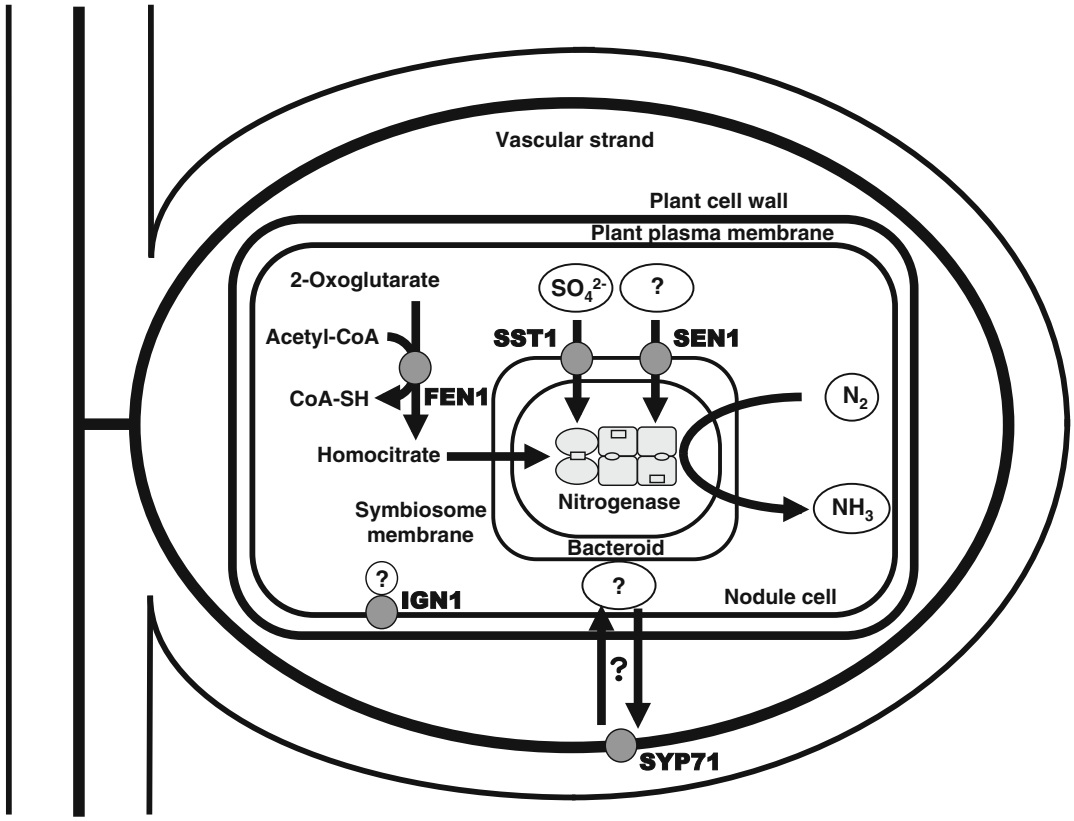


Fig. 8.1 Location and function of SST1, SEN1, FEN1, SYP71, and IGN1 for rhizobial symbiotic nitrogen fixation in *L. japonicus* nodules. SST1 located in the symbiosome membrane transports SO_4^{2-} from the plant cytosol to bacteroides, as sulfur is required for components of the nitrogenase complex. SEN1 is predicted to be located in the symbiosome membrane and to supply an essential element for nitrogenase. FEN1 located in the cytoplasm of infected cells catalyzes the formation of homocitrate, which is supplied to bacteroides for synthesis of the FeMo-cofactor, one component of the nitrogenase complex. SYP71 is expressed in the vascular tissues and might translocate an unknown substance crucial for nitrogen fixation. IGN1 localized in the plasma membrane of infected cells is hypothesized to interact with other proteins and to function in symbiosome and/or bacteroid differentiation and maintenance

(Krusell et al. 2005). These findings suggest that *SST1* might have been recruited from a preexisting sulfate transporter gene in *L. japonicus.*

The *SEN1* gene encodes an integral membrane protein homologous to nodulin-21 of *Glycine max* (Kawaguchi et al. 2002; Suganuma et al. 2003; Hakoyama et al. 2012a). The location and function of the SEN1 protein remain to be demonstrated. However, it is predicted to occur in the symbiosome membrane and to transport an essential element required for nitrogen fixation, because the *SEN1* gene is expressed exclusively in infected nodule cells. The nodules of the *sen1* mutant completely lack nitrogen-fixation activity, even though the proteins making up the nitrogenase complex are expressed (Suganuma et al. 2003; Hakoyama et al. 2012a); this suggests that SEN1 is indispensable for the induction of nitrogen-fixation activity. In addition, symbiosome and/or bacteroid differentiation are impaired even at an early stage of nodule development in *sen1* mutants (Hakoyama et al. 2012a). This suggests that SEN1 is more directly involved in symbiosome and/or bacteroid differentiation than in nitrogen-fixation activity, in contrast to SST1. The function of SEN1 needs to be clarified to elucidate the host regulation of rhizobial symbiotic nitrogen fixation.

8.2.2 Enzymes Involved in Carbon Metabolism: FEN1

The nitrogenase reaction requires 16 molecules of ATP to reduce one molecule of dinitrogen gas. This energy is produced by aerobic respiration of bacteroides, using the carbon source supplied by host plants. Carbon metabolism in host cells is therefore critical to rhizobial symbiotic nitrogen fixation. Indeed, the expression of genes encoding enzymes involved in carbon catabolism is greatly enhanced in *L. japonicus* nodules (Colebatch et al. 2004). One such gene, which encodes nodule-enhanced phosphoenolpyruvate carboxylase, was shown by RNA interference to be critical to rhizobial symbiotic nitrogen fixation in *L. japonicus* (Nomura et al. 2006). However, no *L. japonicus* Fix^- mutants defective in enzymes involved in carbon metabolism have been isolated so far.

The analysis of another *L. japonicus* Fix^- mutant, *fen1*, revealed a novel metabolic reaction in host plant cells crucial for rhizobial symbiotic nitrogen fixation. The defective gene in the *fen1* mutant was shown to encode homocitrate synthase, which catalyzes the conversion of 2-oxoglutarate to homocitrate with acetyl-coenzyme A (Imaizumi-Anraku et al. 1997; Kawaguchi et al. 2002; Hakoyama et al. 2009). Homocitrate is a component of the FeMo-cofactor required for the nitrogenase complex (Hoover et al. 1987, 1989). *M. loti* lacks the *NifV* gene encoding homocitrate synthase (Hakoyama et al. 2009). This indicates that host plant cells provide bacteroides with homocitrate to support efficient nitrogen-fixation activity. This hypothesis was supported by results showing that the mutant *fen1* phenotype was rescued either by expressing the wild-type *FEN1* gene or the *Azotobacter NifV* gene in *M. loti* rhizobia, or by supplying synthetic homocitrate. In *Saccharomyces cerevisiae*, homocitrate is a precursor to the biosynthesis of lysine. However, higher plants, including legumes, are able to synthesize lysine from a distinct pathway without homocitrate synthase. Legumes are therefore likely to have acquired the homocitrate synthase gene to overcome a lack of the *NifV* gene in rhizobia. The identification of the *FEN1* gene reveals a novel aspect of the interrelationship between legumes and rhizobia symbiosis, and further exploration could shed light on the evolution of symbiosis.

8.2.3 Proteins Expressed in Vascular Tissues: SYP71

Photosynthates assimilated by the host plant are translocated from the shoots to the nodules, and nitrogen fixed by rhizobia is transported from the nodules to the shoots. The plant vascular system, which carries both photosynthates and nitrogenous compounds, contributes significantly to effective symbiotic nitrogen fixation (as reviewed by Guinel 2009). Several genes required for nitrogen fixation are expressed specifically in the vascular tissues of nodules. Recently, a Fix^- mutant defective in the *LjSYP71* gene was identified (Hakoyama et al. 2012b). *LjSYP71* encodes a Qc-SNARE (soluble N-ethylmaleimide sensitive factor attachment protein receptor) homologous to *Arabidopsis thaliana* SYP71. SNARE proteins are involved in vesicle trafficking. *LjSYP71* was expressed in the whole plant, and transcripts were detected in the vascular tissues. The discovery of the LjSYP71-defective Fix^- mutant suggests the existence of a long-distance signal transported as cargo in the vesicle, which regulates nitrogen-fixing activity. It is well known that the number of nodules that form on legume roots is regulated by a long-distance signal derived from the shoot (Oka-Kira and Kawaguchi 2006). Further efforts are required to clarify whether the phenotype of the *Ljsyp71* mutant is recovered by shoot grafting and to determine what is transported in vesicle trafficking involving LjSYP71.

8.2.4 Proteins Required for Maintenance of Symbiosis: IGN1

The identification of the *L. japonicus* Fix^- mutant *ign1* suggested that the host plant gene was required for the maintenance of compatible rhizobia in nodule cells (Kumagai et al. 2007).

Even at early stages of nodule development, *ign1* nodules contain irregularly shaped and enlarged symbiosomes similar to lytic vacuoles with multiple bacteroides. Furthermore, as the nodules develop, infected *ign1* mutant cells disintegrate rapidly and the bacteroides appear to aggregate. The premature senescence observed in *ign1* nodules is more rapid than that in other Fix^- mutants, which generally show a phenotype of premature senescence, possibly because they lack nitrogen-fixation activity. The rapid disintegration of the symbiosomes might therefore not be due simply to a lack of nitrogen-fixation activity. The responsible gene, *IGN1*, encodes a novel protein containing ankyrin repeats that serve as a domain for protein–protein interactions. In addition, the IGN1 protein is targeted to the plasma membrane of infected cells. IGN1 might function as a membrane-anchorage protein that regulates the subcellular localization of other proteins, or interacts with and regulates other membrane proteins or transporters. The *IGN1* gene is expressed constitutively in all organs of *L. japonicus* plants, but the *ign1* mutant shows no growth abnormalities other than the symbiotic defect. IGN1 might thus be required to prevent host plant cells from inappropriately invoking premature senescence or as a kind of defense system against microsymbionts, thereby playing a critical role in the differentiation and/or persistence of bacteroides and symbiosomes (Kumagai et al. 2007). Elucidating the exact biochemical function of IGN1 could provide new insights into plant–rhizobium interactions.

8.3 Future Perspectives

Symbiotic nitrogen fixation in nodules is thought to be supported by host plants at various levels, including structure, metabolism, transport, and recognition. The identification of additional Fix^- mutants could shed light on the host plant regulation of rhizobial nitrogen fixation. Some such mutants are currently under investigation (Hossain et al. 2006; Sandal et al. 2006), and our knowledge of the host plant genes that are essential for symbiotic nitrogen fixation is predicted to improve substantially in the near future. Analysis of the *fen1* mutant revealed that its phenotype was recovered by inoculation with *M. loti* transformed with the *FEN1* or *Azotobacter NifV* gene. However, the transformed *M. loti* still lacked nitrogen-fixation activity under free-living conditions. This implies that other essential factors for rhizobial symbiotic nitrogen fixation are supplied by the host plant. The identification of the genes critical to symbiotic nitrogen fixation also might help to clarify the co-evolution of legumes and rhizobia.

Acknowledgments The author thanks Dr. F.C. Guinel (Wilfrid Laurier University, Waterloo, Canada) for critical reading of the manuscript.

References

Colebatch G, Kloska S, Trevaskis B et al (2002) Novel aspects of symbiotic nitrogen fixation uncovered by transcript profiling with cDNA arrays. Mol Plant-Microbe Interact 15:411–420

Colebatch G, Desbrosses G, Ott T et al (2004) Global changes in transcription orchestrate metabolic differentiation during symbiotic nitrogen fixation in *Lotus japonicus*. Plant J 39:487–512

Guinel FC (2009) Getting around the legume nodule: II. Molecular biology of its peripheral zone and approaches to study its vasculature. Botany 87:1139–1166

Hakoyama T, Niimi K, Watanabe H et al (2009) Host plant genome overcomes the lack of a bacterial gene for symbiotic nitrogen fixation. Nature 462:514–517

Hakoyama T, Niimi K, Yamamoto T et al (2012a) The integral membrane protein SEN1 is required for symbiotic nitrogen fixation in *Lotus japonicus* nodules. Plant Cell Physiol 53:225–236

Hakoyama T, Oi R, Hazuma K et al (2012b) The SNARE protein SYP71 expressed in vascular tissues is involved in symbiotic nitrogen fixation in *Lotus japonicus* nodules. Plant Physiol 160:897–905

Hoover T, Robertson AD, Cerny RL et al (1987) Identification of the V factor needed for synthesis of the iron-molybdenum cofactor of nitrogenase as homocitrate. Nature 329:855–857

Hoover TR, Imperial J, Ludden PW et al (1989) Homocitrate is a component of the iron-molybdenum cofactor of nitrogenase. Biochemistry 28:2768–2771

Hossain MS, Umehara Y, Kouchi H (2006) A novel Fix^- symbiotic mutant of *Lotus japonicus*, *Ljsym105*, shows impaired development and premature deterioration of nodule infected cells and symbiosomes. Mol Plant-Microbe Interact 19:780–788

Imaizumi-Anraku H, Kawaguchi M, Koiwa H et al (1997) Two ineffective-nodulating mutants of *Lotus japonicus* — Different phenotypes caused by the blockage of endocytotic bacterial release and nodule maturation —. Plant Cell Physiol 38:871–881

Kawaguchi M, Imaizumi-Anraku H, Koiwa H et al (2002) Root, root hair, and symbiotic mutants of the model legume *Lotus japonicus*. Mol Plant-Microbe Interact 15:17–26

Kneen BE, LaRue TA, Hirsch AM et al (1990) *sym 13*— A gene conditioning ineffective nodulation in *Pisum sativum*. Plant Physiol 94:899–905

Kouchi H, Shimomura K, Hata S et al (2004) Large-scale analysis of gene expression profiles during early stages of root nodule formation in a model legume, *Lotus japonicus*. DNA Res 11:263–274

Kouchi H, Imaizumi-Anraku H, Hayashi M et al (2010) How many peas in a pod? Legume genes responsible for mutualistic symbioses underground. Plant Cell Physiol 51:1381–1397

Kouchi H (2011) Symbiotic nitrogen fixation. In: Ashihara H, Crozier A, Komamine A (eds) Plant Metabolism and Biotechnology. Wiley, Chichester, pp 67–102

Krusell L, Krause K, Ott T et al (2005) The sulfate transporter SST1 is crucial for symbiotic nitrogen fixation in *Lotus japonicus* root nodules. Plant Cell 17:1625–1636

Kumagai H, Hakoyama T, Umehara Y et al (2007) A novel ankyrin-repeat membrane protein, IGN1, is required for persistence of nitrogen-fixing symbiosis in root nodules of *Lotus japonicus*. Plant Physiol 143:1293–1305

Mergaert P, Uchiumi T, Alunni B et al (2006) Eukaryotic control on bacterial cell cycle and differentiation in the *Rhizobium*–legume symbiosis. Proc Natl Acad Sci USA 103:5230–5235

Nomura M, Mai HT, Fujii M et al (2006) Phosphoenolpyruvate carboxylase plays a crucial role in limiting nitrogen fixation in *Lotus japonicus* nodules. Plant Cell Physiol 47:613–621

Oka-Kira E, Kawaguchi M (2006) Long-distance signaling to control root nodule number. Curr Opin Plant Biol 9:496–502

Ott T, van Dongen JT, Günther C et al (2005) Symbiotic leghemoglobins are crucial for nitrogen fixation in legume root nodules but not for general plant growth and development. Curr Biol 15:531–535

Sandal N, Petersen TR, Umehara Y et al (2006) Genetics of symbiosis in *Lotus japonicus*: recombinant inbred lines, comparative genetic maps, and map position of 35 symbiotic loci. Mol Plant-Microbe Interact 19:80–91

Sato S, Nakamura Y, Kaneko T et al (2008) Genome structure of the legume, *Lotus japonicus*. DNA Res 15:227–239

Schauser L, Handberg K, Sandal N et al (1998) Symbiotic mutants deficient in nodule establishment identified after T-DNA transformation of *Lotus japonicus*. Mol Gen Genet 259:414–423

Suganuma N, Nakamura Y, Yamamoto M et al (2003) The *Lotus japonicus Sen1* gene controls rhizobial differentiation into nitrogen-fixing bacteroides in nodules. Mol Gen Genomics 269:312–320

Szczyglowski K, Shaw RS, Wopereis J et al (1998) Nodule organogenesis and symbiotic mutants of the model legume *Lotus japonicus*. Mol Plant-Microbe Interact 7:684–697

Tsyganov VE, Morzhina EV, Stefanov AY et al (1998) The pea (*Pisum sativum* L.) genes *sym33* and *sym40* control infection thread formation and root nodule function. Mol Gen Genet 259:491–503

Udvardi MK, Day DA (1997) Metabolite transport across symbiotic membranes of legume nodules. Annu Rev Plant Physiol Plant Mol Biol 48:493–523

Udvardi MK, Poole PS (2013) Transport and metabolism in legume-rihzobai symbioses. Annu Rev Plant Biol 64:781–805

Vance CP, Johnson LEB (1983) Plant determined ineffective nodules in alfalfa (*Medicago sativa*): structural and biochemical comparisons. Can J Bot 61:93–106

White J, Prell J, James EK et al (2007) Nutrient sharing between symbionts. Plant Physiol 144:604–614

Wienkoop S, Saalbach G (2003) Proteome analysis: novel proteins identified at the peribacteroid membrane from *Lotus japonicus* root nodules. Plant Physiol 131: 1080–1090

Hormone Regulation of Root Nodule Formation in Lotus

9

Akihiro Suzuki

Abstract

Several phytohormones have been reported to positively or negatively regulate the formation of nitrogen-fixing nodules in *Lotus japonicus* and other legumes. Auxin is important for root nodulation and required for cortical cell division. Because auxin accumulation was observed in the root of the *L. japonicus* mutant *spontaneous nodule formation 2* (*snf2*), which has a gain-of-function mutation in a putative cytokinin receptor, it appears that auxin acts downstream of cytokinin signaling. Activation of cytokinin signaling is involved in the induction of root nodule formation. Ethylene, gibberellin (GA), and abscisic acid (ABA) inhibit the cortical cell divisions induced by cytokinin. ABA regulates nitrogen fixation activity through the control of nitric oxide levels. Though jasmonic acid (JA) is known as a negative regulator of nodulation, recent data suggest that it functions as a positive regulator over a certain range of concentrations. The increase in salicylic acid (SA) levels normally triggered as a defense response does not occur upon infection with compatible symbionts. *LjCCD7*-silenced *L. japonicus* plants, which were expected to have reduced concentrations of strigolactone, produced fewer nodules than the controls, suggesting that strigolactone promotes nodule formation in *L. japonicus*.

9.1 Introduction

The root nodule, an organ where atmospheric nitrogen is fixed, is formed by rhizobial infection. Because phytohormones are signal molecules involved in plant development, morphogenesis, and responses to the environment, it is expected that most of them are directly or indirectly involved in plant interaction with rhizobia and in root nodule organogenesis. The roles of phytohormones in root nodule formation are gradually being discovered as molecular genetics information from model legumes accumulates. In this chapter, the phytohormones that appear to play roles in root nodule formation are reviewed.

A. Suzuki (✉)
Department of Environmental Sciences, Faculty of Agriculture, Saga University, 1 Honjyo-machi, Saga 840-8502, Japan
e-mail: azuki@cc.saga-u.ac.jp

S. Tabata and J. Stougaard (eds.), *The Lotus japonicus Genome*, Compendium of Plant Genomes,
DOI 10.1007/978-3-662-44270-8_9,

9.2 Auxin

Auxin functions as a signal molecule that controls germination and growth, flower bud formation, flowering, and embryo formation and development. In addition, auxin is involved in responses to environmental stimuli such as light and gravity (i.e., tropisms). The roles of auxin in root nodule formation have been studied for many years in leguminous plants that form indeterminate-type nodules (those that continue to produce new cells after initiation), such as *Pisum sativum*, *Medicago sativa*, and *Medicago truncatula.*

Thimann (1936) reported that *P. sativum* root nodules contained auxin and that the auxin content increased during root nodule development. The *Medicago truncatula-like AUX1* (*MtLUX*) gene, which is similar to the *AUX1* gene of *Arabidopsis thaliana*, was expressed in the region of each organ where vasculature arises (i.e., in the center of lateral roots and the peripheral region of nodules). These results suggest that auxin is required at the stage of primordia development and differentiation of the vasculature within the nodule (de Billy et al. 2001). The effects of treatment with auxin polar transport inhibitors 1-naphthylphthalamic acid (NPT) and 2,3,5-triindobenzoic acid (TIBA) and auxin antagonist ∝-(phenylethyl-2-one)-indole-3-acetic acid (PEO-IAA) on root nodule formation were investigated using *Lotus japonicus*, which produces determinate-type nodules (nodules that stop producing new cells shortly after initiation). Both nodule number and nodule development were reduced, and formation of the lenticel, which normally develops on the root surface and originates from the root outer cortex, was also inhibited by the treatment (Takanashi et al. 2011). The expression of a transgene containing the auxin-responsive promoter from *Glycine max GH3* fused to a *GUS* reporter gene was analyzed in *L. japonicus* roots, and GUS activity was observed in the vascular tissues of nodules. These results suggest that auxin plays an important role in the development of nodule vasculature regardless of nodule type (Pacios-Bras et al. 2003; Takanashi et al. 2011). Though an auxin signal was detected in the dividing outer cortical cells during the first nodule cell divisions in *L. japonicus*, GUS expression was not detected in those same cells in white clover (*Trifolium repens*), which produces an indeterminate-type nodule. Flavonoids inhibit auxin transport, and formation of indeterminate-type nodules is affected by flavonoids. For example, root nodule development was inhibited on hairy roots of *M. truncatula* in which *chalcone synthase* (*CHS*) gene expression was suppressed; these roots had a lower concentration of flavonoids, so auxin transport was not inhibited. In contrast, flavonoid-regulated auxin transport inhibition is not crucial during root nodule formation in *G. max*, which produces determinate nodules (Subramanian et al. 2006). In *L. japonicus* root nodules, no inhibition of auxin transport was observed (Pacios-Bras et al. 2003). These differences in auxin distribution and transport inhibition between determinate and indeterminate nodules have been attributed to the difference in developmental pattern of the two nodule types.

Recently, Suzaki et al. (2012) investigated auxin distribution during root nodule development by using transformed *L. japonicus* carrying the auxin-responsive DR5 promoter fused to a *GFP* gene with a nuclear localization signal. The accumulation of auxin in the dividing cortical cells was positively regulated by NIN (nodule inception, a key transcription factor in nodule development) and was inhibited by a systemic negative regulatory mechanism called autoregulation of nodulation (AON). Moreover, auxin accumulation was observed in uninoculated roots of the *L. japonicus* mutant *spontaneous nodule formation 2* (*snf2*), which has a gain-of-function mutation in LHK1 (the putative cytokinin receptor lotus histidine kinase 1). Therefore, it appears that auxin is involved in the division of cortical cells and acts downstream of cytokinin signaling (Fig. 9.1).

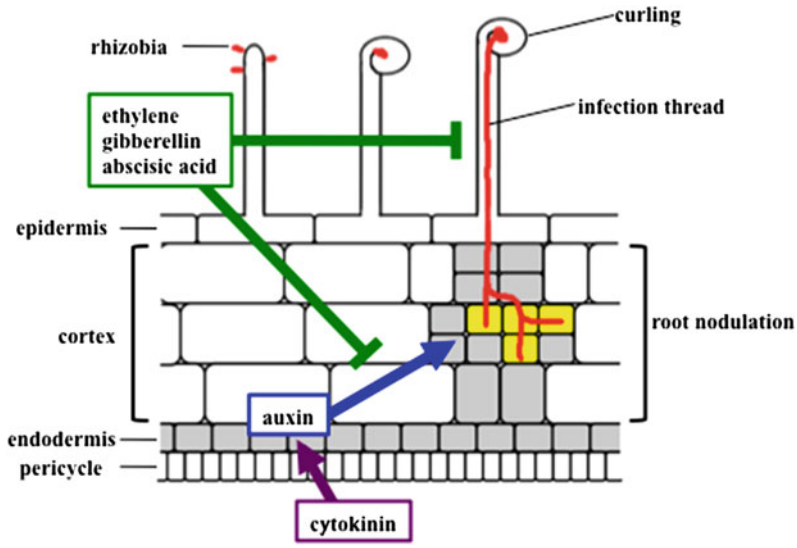


Fig. 9.1 Effects of phytohormones on root nodule formation. Auxin and cytokinin are required for cortical cell division and auxin acts downstream of cytokinin signaling. Ethylene, gibberellin, and abscisic acid inhibit the cortical cell divisions and rhizobial infection

9.3 Cytokinin

Cytokinins regulate cell division, induction of shoot formation from callus, activation of lateral bud growth, suppression of senescence, and movement of nutrients. Lohar et al. (2004) investigated the cytokinin distribution in transgenic *L. japonicus* containing a chimeric gene consisting of the promoter from the *Arabidopsis* response regulator gene *ARR5* driving the *GUS* reporter gene. Lohar et al. (2004) found that GUS distribution in uninoculated *L. japonicus* was similar to that in *A. thaliana*. In plants inoculated with *Mesorhizobium loti*, ARR5 expression was observed in curled/deformed root hairs and in nodule primordia, indicating that cytokinin had accumulated in those cells. Several reports have indicated that cytokinin promotes the root nodulation process. For example, *Sinorhizobium meliloti* strains that lacked the ability to form root nodules because of a mutation in the *nodA* or *nodB* genes recovered this ability when transformed with a plasmid containing the *IPT* gene, which encodes cytokinin synthetase. This result was very interesting because the phenotype was complemented not by an intact version of the mutated gene (i.e., *nodA* or *nodB*), but by a cytokinin synthetase gene (Cooper and Long 1994). Cytokinin treatment induced root nodule–like structures in uninoculated wild-type *L. japonicus* and also induced *NIN* expression (Heckmann et al. 2011). Root nodule formation was suppressed in *L. japonicus* transformed with a chimeric gene consisting of the *CaMV35S* promoter fused to either *Arabidopsis CKX3* or maize *CKX1*, both of which are cytokinin degradation genes (Lohar et al. 2004). Several years later, Murray et al. (2007) and Tirichine et al. (2007) reported that a cytokinin signal was necessary for root nodule formation in *L. japonicus*. Genetic integration of the cytokinin phosphorelay pathway in root nodule formation has been demonstrated using gain- and loss-of-function mutants of *LjLHK1*. In the *hit1* mutant (loss-of-function), infection frequency was increased but root nodules did not form (Murray et al. 2007). Meanwhile, spontaneous root nodule formation was observed in the *snf2* (gain-of-function) mutant, in which cytokinin signaling is constitutively "on" (Tirichine et al. 2007). These facts suggest that activation of cytokinin signaling is involved in the induction of root nodule formation (Fig. 9.1). In *M. truncatula*, root nodule primordium formation was inhibited by RNAi-mediated down-regulation of the gene for type A cytokinin response regulator MtRR9 and enhanced by its overexpression. Similar results were observed in *L. japonicus* transformed with *MtRR9*. Furthermore, the expression of *L.*

japonicus LjRR6, which showed high similarity to *MtRR9*, was increased by treatment with Nod factor. These results further illustrate the involvement of cytokinin signaling in root nodule formation (Op den Camp et al. 2011) (Fig. 9.1).

9.4 Ethylene

Ethylene is a gaseous phytohormone involved in fruit ripening, leaf and fruit abscission, germination, seedling morphogenesis, root emergence, root hair elongation, promotion of flowering, senescence, and stress response. Many reports have described the inhibitory effects of ethylene in the root nodule formation process. For example, ethylene inhibited the elongation of infection threads into the inner cortex of *P. sativum* roots (Lee and LaRue 1992), and calcium spiking (a response to bacterial nodulation signals) did not occur in ethylene-treated *M. truncatula*. The causative gene for the ethylene-insensitive *M. truncatula sickle* mutant is *EIN2*, which is involved in ethylene signaling. This mutant showed a hyperinfection phenotype following rhizobial inoculation (Penmetsa and Cook 1997; Oldroyd et al. 2001; Penmetsa et al. 2003, 2008). Ethylene also shows inhibitory effects in *L. japonicus*. Root nodule formation was suppressed by treatment with 1-aminocyclopropane-1-carboxylic acid (ACC), a precursor of ethylene, but enhanced by treatment with aminoethoxyvinyl glycine (AVG), which inhibits the activity of ACC synthetase (ACS), and silver thiosulfate (STS), which inhibits ethylene perception (Nukui et al. 2000). Nukui et al. (2004) produced transgenic *L. japonicus* carrying the mutated melon ethylene receptor gene *Cm-ERS1/H70A*, which confers ethylene insensitivity. When inoculated with *M. loti*, the transgenic plants showed markedly higher numbers of infection threads and nodule primordia on their roots than control plants did. In addition, *NIN* transcript levels increased in the inoculated transgenic plants as compared to wild-type plants. Similar results were obtained using *L. japonicus* transformed with a dominant-negative ethylene receptor mutant gene from *Arabidopsis* (Lohar et al. 2009). Interestingly, the number of root nodules was not increased in an ethylene-insensitive mutant of *G. max* that had a mutation in *ethylene-resistance 1* (*ETR1*) (Schmidt et al. 1999). Consistent with these results, the number of root nodules formed in *G. max* treated with ACC, AVG, or STS was unchanged (Nukui et al. 2000).

ACC deaminase catalyzes the degradation of ACC into ammonium and α-ketobutyrate, thus reducing ethylene levels (Honma and Shimomura 1978). The ACC deaminase gene (*acdS*) has been found in many rhizosphere bacteria (Honma and Shimomura 1978; Grichko and Glick 2000), including *M. loti* MAFF303099 (Kaneko et al. 2000). In *M. loti*, *acdS* was found in the symbiosis island, and the enhancing effect of this gene on nodulation of *L. japonicus* roots was demonstrated by using an *M. loti acdS* disruption mutant (Uchiumi et al. 2004). Furthermore, DNA macroarray analysis showed that a clone containing *acdS* (mlr5932) was up-regulated in bacteroid cells. These studies also showed the negative effect of ethylene on root nodule formation (Fig. 9.1).

Legumes control nodule numbers through the systemic AON process (Gresshoff 2003). Mutants in this circuit have a supernodulating phenotype (Krusell et al. 2002; Nishimura et al. 2002). A causative gene is the CLAVATA1-related LRR receptor kinase *HAR1*. Because *L. japonicus* AON mutants are still ethylene-sensitive, and a double mutant of *Ljhar1/LjETR1-1* does not show an additive phenotype, the AON, and ethylene pathways are independent (Gresshoff et al. 2009).

9.5 Gibberellin (GA)

Gibberellin (GA) regulates seed germination, stem elongation, flower bud formation, and fruit enlargement. The effects of gibberellin on root nodulation in *L. japonicus* have been reported in detail (Maekawa et al. 2009). Exogenous application of a biologically active GA, GA_3, inhibited the formation of infection threads and

nodules, and this inhibition was counteracted by the application of Uniconazole-P, an inhibitor of GA_3 biosynthesis. Moreover, Nod factor–-induced root hair deformation was severely blocked in the presence of GA_3. GA treatment also suppressed nodule formation on the roots of *snf2* mutant plants, spontaneous nodulation triggered by a gain-of-function mutation of calcium/calmodulin-dependent kinase (CCaMK) (*snf1*), and cytokinin-dependent induction of NIN. These results indicate that GA inhibits the nodulation signaling pathway downstream of cytokinin (Fig. 9.1).

The F box–containing protein SLEEPY1 (SLY1) functions as a positive regulator in GA signaling. In the presence of GA, SLY1 interacts with negative regulators of GA signaling (e.g., DELLA domain–containing GRAS proteins), leading to their degradation (McGinnis et al. 2003; Dill et al. 2004). In *L. japonicus* overexpressing SLEEPY, root nodule formation was inhibited in spite of normal root development (Maekawa et al. 2009). The *la cry-s* constitutive GA signaling mutants in *P. sativum* also form significantly fewer nodules than wild-type plants. However, GA deficiency resulting from the *na* mutation in *P. sativum* also causes a reduction in nodulation, indicating that some level of GA signaling is required for normal root nodule development (Ferguson et al. 2011).

9.6 Abscisic Acid (ABA)

Abscisic acid (ABA) is involved in various stress responses, seed maturation, germination, and stomatal closure. ABA shows negative effects on root nodule formation. Exogenous application of ABA inhibited root nodule formation in *P. sativum* (Phillips 1971), *G. max* (Cho and Harper 1993; Bano and Harper 2002), *L. japonicus* (Suzuki et al. 2004), *T. repens* (Suzuki et al. 2004), and *M. truncatula* (Ding et al. 2008) (Fig. 9.1). ABA was also able to inhibit root nodule formation in the hypernodulation mutant *NOD1-3* of *G. max* (Cho and Harper 1993). In both, *L. japonicus* and *M. truncatula*, ABA treatment inhibited infection thread formation in the root hair (Suzuki et al. 2004; Nakatsukasa-Akune et al. 2005; Ding et al. 2008) (Fig. 9.1). Moreover, Ding et al. (2008) found that in *M. truncatula*, calcium spiking after Nod factor perception was inhibited by the application of ABA. Phillips (1971) observed that cell division induced by cytokinin was arrested by ABA treatment and postulated that ABA had an inhibitory effect on cortical cell division induced by cytokinin. Ding et al. (2008) proved that this idea was correct. ABA treatment inhibited spontaneous root nodule formation in the *snf2* mutant of *L. japonicus* and cytokinin-induced *ENOD40* gene expression in the cortex of wild-type *M. truncatula*. These results suggest that both infection thread formation in root hairs and induction of cortical cell division induced by cytokinin are regulated by ABA.

Due to the importance of root nodule function, methods to increase the number of root nodules have been of interest for many years. Studies of hypernodulation mutants are good examples of this interest. However, because production of many nodules is costly to the plant, growth is drastically affected in such mutants (Nishimura et al. 2002). The *enf1* (enhanced *nitrogen fixation 1*) mutant was isolated by screening *L. japonicus* seedlings for survival on an agar medium containing 70 μM ABA, indicating reduced sensitivity to ABA. The number of nodules formed on *enf1* roots was approximately 1.7 times that of wild-type MG20. The low ABA sensitivity of the *enf1* mutant was thought to result from its lower endogenous ABA concentration (Tominaga et al. 2009). Enhanced root nodule formation was also observed in transgenic *M. truncatula* carrying a dominant-negative allele of *abscisic acid insensitive1* from *Arabidopsis* (Ding et al. 2008). Taken together, the available data indicate that root nodule formation relies on an exquisite balance between cytokinin and ABA. Though the regulation of nodulation by ABA resembles the effects of the negative regulator ethylene, experiments with the ethylene-insensitive *sickle* mutant show that ABA and ethylene function independently in the regulation of lateral root initiation, nodulation, and Nod factor signal transduction (Ding et al. 2008).

ABA regulates not only root nodule formation but also nitrogen fixation activity. When mature nodules of *P. sativum* were treated with ABA, nitrogen fixation activity decreased (González et al. 2001). Conversely, the nitrogen fixation activity on a per-plant basis in *enf1* (reduced ABA) plants was 1.8 times that in wild-type plants (Tominaga et al. 2009). Nitric oxide (NO) is a strong inhibitor of nitrogen fixation (Trinchant and Rigaud 1982; Shimoda et al. 2009). The reduced endogenous ABA concentration in *enf1* plants correlated with elevated nitrogen fixation activity and reduced NO production in the root nodule. These results suggest that endogenous ABA concentration regulates nitrogen fixation activity by regulating NO production in nodules.

9.7 Jasmonic Acid (JA)

Jasmonic acid (JA) is involved in defense responses to pathogens, wound responses, leaf senescence, and tuber formation. JA has often been reported to be a negative regulator of root nodule formation. For example, in *L. japonicus*, foliar application of 1 μM or more methyl jasmonate (MeJA) inhibited nodule development (Nakagawa and Kawaguchi 2006). In *M. truncatula*, Nod factor–induced *ENOD11* and *RIP1* gene expression, calcium spiking, and root nodule development were suppressed by the addition of JA to the growth medium (Sun et al. 2006). In contrast, the JA concentration in leaves of the *G. max nts* hypernodulation mutant was higher than in the wild-type (Seo et al. 2007). The expression of genes related to JA biosynthesis and JA responses in the leaves of wild-type *G. max* is normally suppressed by inoculation with rhizobia, but no suppression was seen in the hypernodulation mutant *nts1007* (Kinkema and Gresshoff 2008), consistent with the findings of Seo et al. (2007). Furthermore, Kinkema and Gresshoff (2008) showed that spraying the shoots of hypernodulation mutant plants with *n*-propyl gallate, an inhibitor of JA biosynthesis, significantly reduced the number of root nodules. Recently, we reported that root nodule formation was enhanced by treatment with a low concentration of JA (0.1 μM) (Suzuki et al. 2011). Therefore, JA functions as a positive regulator of root nodulation over a certain range of concentrations in these plant species. Because plants secrete JA from the roots into the rhizosphere (Creelman and Mullet 1995), and *nod* gene expression and Nod factor production are upregulated by JA application (Rosas et al. 1998; Mabood et al. 2006), it is unclear whether plant response, rhizobial response, or both cause the positive effects of JA on root nodulation. We also found that nodule formation was suppressed on the roots of a *phytochrome B* (*phyB*) mutant of *L. japonicus* that had not only decreased levels of photoassimilates but also a reduced concentration of JA-Ile (the active JA derivative) (Suzuki et al. 2011). In fact, the number of root nodules in the *phyB* mutant was restored by JA treatment, providing further evidence that JA can act as a positive regulator of nodulation.

9.8 Salicylic Acid (SA)

Salicylic acid (SA) is an inducer of systemic acquired resistance (SAR) in defense responses to pathogens. Van Spronsen et al. (2003) reported that root nodule formation in *Vicia sativa* subsp. *nigra* (vetch, an indeterminate-type nodulating plant), but not in *L. japonicus*, was suppressed by treatment with SA. In contrast, when endogenous levels of SA were modulated through the transgenic expression of salicylate hydroxylase (*NahG*) in *L. japonicus*, the number of rhizobial infections and nodules increased (Stacey et al. 2006). Thus, endogenous SA concentration affects the number of root nodules regardless of nodule type (i.e., determinate or indeterminate). In *M. sativa*, endogenous SA concentration did not increase after inoculation with wild-type *S. meliloti*, but did increase after inoculation with an incompatible *nodC* mutant of *S. meliloti* (Martínez-Abarca et al. 1998). These results suggest that in the presence of a biotic stress, the plant defense response begins with an

increase in the endogenous concentration of SA, whereas this defense response is avoided in the case of a compatible symbiont.

9.9 Strigolactones (SLs)

Strigolactones (SLs) are newly identified hormones that regulate multiple aspects of plant development, infection by parasitic weeds, and mutualistic symbiosis in the root. The correlation between SLs and nodulation was previously investigated in *M. sativa* and *P. sativum*, both of which form indeterminate nodules. In these species, root nodulation is positively affected by the treatment of the SLs (Soto et al. 2010; Foo and Davies 2011; Foo et al. 2013). The role of SLs was studied in *L. japonicus* using transgenic lines silenced for *carotenoid cleavage dioxygenase 7* (*LjCCD7*), an ortholog of Arabidopsis *More Axillary Growth 3*. Silencing of *LjCCD7* is expected to reduce strigolactone levels. In *L. japonicus*, *LjCCD7*-silenced plants had 20 % fewer nodules than controls, suggesting that SLs have a slight positive effect on the formation of determinate nodules (Liu et al. 2013).

References

Bano A, Harper JE (2002) Plant growth regulators and phloem exudates modulate root nodulation of soybean. Funct Plant Biol 29:1299–1307

Cho M-J, Harper JE (1993) Effect of abscisic acid application on root isoflavonoid concentration and nodulation of wild-type and nodulation-mutant soybean plants. Plant Soil 152:145–149

Cooper JB, Long SR (1994) Morphogenetic rescue of Rhizobium meliloti nodulation mutants by trans-Zeatin secretion. Plant Cell 6:215–225

Creelman RA, Mullet JE (1995) Jasmonic acid distribution and action in plants: regulation during development and response to biotic and abiotic stress. Proc Natl Acad Sci USA 92:4114–4119

de Billy F, Grosjean C, May S et al (2001) Expression studies on AUX1-like genes in *Medicago truncatula* suggest that auxin is required at two steps in early nodule development. Mol Plant Microbe Interact 14:267–277

Dill A, Thomas SG, Hu J et al (2004) The Arabidopsis F-box protein SLEEPY1 targets gibberellin signaling repressors for gibberellin-induced degradation. Plant Cell 16:1392–1405

Ding Y, Kalo P, Yendrek C et al (2008) Abscisic acid coordinates Nod factor and cytokinin signaling during the regulation of nodulation in *Medicago truncatula*. Plant Cell 20:2681–2695

Ferguson BJ, Foo E, Ross JJ et al (2011) Relationship between gibberellin, ethylene and nodulation in *Pisum sativum*. New Phytol 189:829–842

Foo E, Davies NW (2011) Strigolactones promote nodulation in pea. Planta 234:1073–1081

Foo E, Yoneyama K, Hugill C et al (2013) Strigolactones and the regulation of pea symbioses in response to nitrate and phosphate deficiency. Mol Plant 6:76–87

González EM, Galvaz L, Arrese-Igor C (2001) Abscisic acid induces a decline in nitrogen fixation that involves leghaemoglobin, but is independent of sucrose synthase activity. J Exp Bot 52:285–293

Gresshoff PM (2003) Post-genomic insights into nodulation. Genome Biol 4:201

Gresshoff PM, Lohar D, Chan PK et al (2009) Genetic analysis of ethylene regulation of legume nodulation. Plant Sign Behav 4:818–823

Grichko VP, Glick BR (2000) Identification of DNA sequences that regulate the expression of the *Enterobacter cloacae* UW4 1-aminocyclopropane-1-carboxylic acid deaminase gene. Can J Microbiol 46:1159–1165

Heckmann AB, Sandal N, Bek AS et al (2011) Cytokinin induction of root nodule primordia in *Lotus japonicus* is regulated by a mechanism operating in the root cortex. Mol Plant Microbe Interact 24:1385–1395

Honma M, Shimomura T (1978) Metabolism of 1-aminocyclopropane-1-carboxylic acid. Agric Biol Chem 42:1825–1831

Kaneko T, Nakamura Y, Sato S et al (2000) Complete genome structure of the nitrogen-fixing symbiotic bacterium *Mesorhizobium loti*. DNA Res 7:331–338

Kinkema M, Gresshoff PM (2008) Investigation of downstream signals of the soybean autoregulation of nodulation receptor kinase GmNARK. Mol Plant Microbe Interact 21:1337–1348

Krusell L, Madsen LH, Sato S et al (2002) Shoot control of root development is mediated by a receptor-like kinase. Nature 420:422–426

Lee KH, LaRue TA (1992) Exogenous ethylene inhibits nodulation of *Pisum sativum L.* cv Sparkle. Plant Physiol 100:1759–1763

Liu J, Novero M, Chamikhova T et al (2013) CAROTENOID CLEAVAGE DIOXYGENASE 7 modulates plant growth, reproduction, senescence, and determinate nodulation in the model legume *Lotus japonicus*. J Exp Bot 64:1967–1981

Lohar DP, Schaff JE, Laskey JG et al (2004) Cytokinins play opposite roles in lateral root formation, and nematode and Rhizobial symbioses. Plant J 38:203–214

Lohar D, Stiller J, Kam J et al (2009) Ethylene insensitivity conferred by a mutated Arabidopsis ethylene receptor gene alters nodulation in transgenic *Lotus japonicus*. Ann Bot 104:277–285

Mabood F, Souleimanov A, Smith DL (2006) Jasmonates induce Nod factor production by *Bradyrhizobium japonicum*. Plant Physiol Biochem 44:759–765

Maekawa T, Maekawa-Yoshikawa M, Takeda N et al (2009) Gibberellin controls the nodulation signaling pathway in *Lotus japonicus*. Plant J 58:183–194

Martínez-Abarca F, Herrera-Cervera JA, Bueno P et al (1998) Involvement of salicylic acid in the establishment of the *Rhizobium meliloti-alfalfa* symbiosis. Mol Plant Microbe Interact 11:153–155

McGinnis KM, Thomas SG, Soule JD et al (2003) The Arabidopsis *SLEEPY1* gene encodes a putative F-box subunit of an SCF E3 ubiquitin ligase. Plant Cell 15:1120–1130

Murray JD, Karas BJ, Sato S et al (2007) A cytokinin perception mutant colonized by Rhizobium in the absence of nodule organogenesis. Science 315:101–104

Nakagawa T, Kawaguchi M (2006) Shoot-applied MeJA suppresses root nodulation in *Lotus japonicus*. Plant Cell Physiol 47:176–180

Nakatsukasa-Akune M, Yamashita K, Shimoda Y et al (2005) Suppression of root nodule formation by artificial expression of the *TrEnodDR1* (coat protein of *white clover cryptic virus 1*) gene in *Lotus japonicus*. Mol Plant Microbe Interact 18:1069–1080

Nishimura R, Hayashi M, Wu GJ et al (2002) HAR1 mediates systemic regulation of symbiotic organ development. Nature 420:426–429

Nukui N, Ezura H, Yuhashi K et al (2000) Effects of ethylene precursor and inhibitors for ethylene biosynthesis and perception on nodulation in *Lotus japonicus* and *Macroptilium atropurpureum*. Plant Cell Physiol 41:893–897

Nukui N, Ezura H, Minamisawa K (2004) Transgenic *Lotus japonicus* with an ethylene receptor gene Cm-ERS1/H70A enhances formation of infection threads and nodule primordia. Plant Cell Physiol 45:427–435

Oldroyd GE, Engstrom EM, Long SR (2001) Ethylene inhibits the Nod factor signal transduction pathway of *Medicago truncatula*. Plant Cell 13:1835–1849

Op den Camp RH, De Mita S, Lillo A et al (2011) A phylogenetic strategy based on a legume-specific whole genome duplication yields symbiotic cytokinin type-A response regulators. Plant Physiol 157:2013–2022

Pacios-Bras C, Schlaman HR, Boot K et al (2003) Auxin distribution in *Lotus japonicus* during root nodule development. Plant Mol Biol 52:1169–1180

Penmetsa RV, Cook DR (1997) A legume ethylene-insensitive mutant hyperinfected by its rhizobial symbiont. Science 275:527–530

Penmetsa RV, Frugoli JA, Smith LS et al (2003) Dual genetic pathways controlling nodule number in *Medicago truncatula*. Plant Physiol 131:998–1008

Penmetsa RV, Uribe P, Anderson J et al (2008) The *Medicago truncatula* ortholog of Arabidopsis EIN2, sickle, is a negative regulator of symbiotic and pathogenic microbial associations. Plant J 55:580–595

Phillips DA (1971) Abscisic acid inhibition of root nodule initiation in *Pisum sativum*. Planta 100:181–190

Rosas S, Soria R, Correa N et al (1998) Jasmonic acid stimulates the expression of nod genes in Rhizobium. Plant Mol Biol 38:1161–1168

Schmidt JS, Harper JE, Hoffman TK et al (1999) Regulation of soybean nodulation independent of ethylene signaling. Plant Physiol 119:951–960

Seo HS, Li J, Lee S-Y et al (2007) The hypernodulating nts mutation induces jasmonate synthetic pathway in soybean leaves. Mol Cells 24:185–193

Shimoda Y, Shimoda-Sasakura F, Kucho K et al (2009) Overexpression of class 1 plant hemoglobin genes enhances symbiotic nitrogen fixation activity between *Mesorhizobium loti* and *Lotus japonicus*. Plant J 57:254–263

Soto MJ, Fernández-Aparicio M, Castellanos-Morales V et al (2010) First indications for the involvement of strigolactones on nodule formation in alfalfa (*Medicago sativa*). Soil Biol Biochem 42:383–385

Stacey G, McAlvin CB, Kim SY et al (2006) Effects of endogenous salicylic acid on nodulation in the model legumes *Lotus japonicus* and *Medicago truncatula*. Plant Physiol 141:1473–1481

Subramanian S, Stacey G, Yu O (2006) Endogenous isoflavones are essential for the establishment of symbiosis between soybean and *Bradyrhizobium japonicum*. Plant J 48:261–273

Sun J, Cardoza V, Mitchell DM et al (2006) Crosstalk between jasmonic acid, ethylene and Nod factor signaling allows integration of diverse inputs for regulation of nodulation. Plant J 46:961–970

Suzaki T, Yano K, Ito M et al (2012) Positive and negative regulation of cortical cell division during root nodule development in *Lotus japonicus* is accompanied by auxin response. Development 139:3997–4006

Suzuki A, Akune M, Kogiso M et al (2004) Control of nodule number by the phytohormone abscisic acid in the roots of two leguminous species. Plant Cell Physiol 45:914–922

Suzuki A, Suriyagoda L, Shigeyama T et al (2011) *Lotus japonicus* nodulation is photomorphogenetically controlled by sensing the red/far red (R/FR) ratio through jasmonic acid (JA) signaling. Proc Natl Acad Sci USA 108:16837–16842

Takanashi K, Sugiyama A, Yazaki K (2011) Involvement of auxin distribution in root nodule development of *Lotus japonicus*. Planta 234:73–81

Thimann KV (1936) On the physiology of the formation of nodules on legumes roots. Proc Natl Acad Sci USA 22:511–514

Tirichine L, Sandal N, Madsen LH et al (2007) A gain-of-function mutation in a cytokinin receptor triggers spontaneous root nodule organogenesis. Science 315:104–107

Tominaga A, Nagata M, Futsuki K et al (2009) Enhanced nodulation and nitrogen fixation in the abscisic acid low-sensitive mutant *enhanced nitrogen fixation1* of *Lotus japonicus*. Plant Physiol 151:1965–1976

Trinchant JC, Rigaud J (1982) Nitrite and nitric oxide as inhibitors of nitrogenase from soybean bacteroids. Appl Environ Microbiol 44:1385–1388

Uchiumi T, Ohwada T, Itakura M et al (2004) Expression islands clustered on the symbiosis island of the *Mesorhizobium loti* genome. J Bacteriol 186:2439–2448

van Spronsen PC, Tak T, Rood AMM et al (2003) Salicylic acid inhibits indeterminate-type nodulation but not determinate-type nodulation. Mol Plant Microbe Interact 16:83–91

Part III
Metabolic Pathways, Secondary Metabolites and Defense Responses

Sucrose and Starch Metabolism

10

Cécile Vriet, Anne Edwards, Alison M. Smith, and Trevor L. Wang

Abstract

The metabolism of starch and sucrose fuels all aspects of plant growth and development. Over the last decade, significant advances have been made in our understanding of the metabolism of these compounds through the use of model systems, mainly *Arabidopsis*. Legume species are characterised by their capacity to form symbioses with *Rhizobium*, a nitrogen-fixing bacterium, leading to up to half the carbon assimilated in photosynthesis being sequestered to their roots. Study of a legume model may therefore increase our knowledge about carbohydrate turnover. We review here the resources available and the contribution that research on *Lotus japonicus* has made to our knowledge of sucrose breakdown and starch metabolism in relation to plant growth and development processes, especially processes that are legume specific.

10.1 Introduction

Plants are photoautotrophic organisms that capture energy from sunlight for the reduction of carbon dioxide to generate sucrose. The sucrose produced is exported from the source leaves to non-photosynthetic parts of the plant, where it is by far the major source of carbon for cellular metabolism. Sucrose and its metabolites are also important signalling molecules, influencing a host of growth and developmental processes. As well as producing sucrose for immediate use, most plants also store some of their assimilated carbon in the form of starch. Starch accumulated in leaves during the day is broken down to provide a supply of carbon for metabolism during the night. Starch accumulated in other organs of the plant, including roots, stems and seeds, fuels cellular metabolism at times when photosynthesis is not possible, for example during regrowth following defoliation and during germination.

Over the last decade, studies performed on the model plant *Arabidopsis thaliana* have greatly improved our understanding of starch and sucrose metabolism (see recent reviews by Smith et al. 2005; Zeeman et al. 2010; Stitt and Zeeman 2012). However, a broad understanding of the importance of these processes for plant growth

C. Vriet
Laboratoire de Génétique et Biophysique des Plantes, Aix-Marseille University, 13009 Marseille, France

A. Edwards · A.M. Smith · T.L. Wang (✉)
Metabolic Biology, John Innes Centre, Norwich Research Park, Colney, Norwich NR4 7UH, UK
e-mail: trevor.wang@jic.ac.uk

S. Tabata and J. Stougaard (eds.), *The Lotus japonicus Genome*, Compendium of Plant Genomes,
DOI 10.1007/978-3-662-44270-8_10,

and development and for yield-related traits of crop species will require research on a broader range of species (Stitt and Zeeman 2012). Such studies will be increasingly straightforward now that genomic resources are being developed for many plants, including cultivated species.

Although the starch metabolism pathway appears well conserved, the extent to which growth and development is dependent on this pathway differs between species (Stitt and Zeeman 2012). By virtue of the genomic resources and the numerous mutants already available, *Lotus japonicus* is a useful model plant in which to uncover some of these differences. Here, we outline the available resources and review progress in identifying and functionally characterising genes of starch and sucrose metabolism in this model legume.

10.2 The Pathways of Sucrose Breakdown and Starch Metabolism

The entry of sucrose into cellular metabolism is catalysed by two enzymes—sucrose synthase and invertase. Sucrose synthases form a small family of cytosolic enzymes, whereas invertases constitute a larger family that is divided into two classes based on properties and cellular location: acid (vacuolar or cell wall localised) invertases and neutral/alkaline invertases. The neutral/alkaline forms can be either cytosolic or organellar, located either in the plastid or the mitochondrion (Vargas and Salerno 2010). Sucrose synthases and invertases catalyse different reactions. The former convert sucrose and UDP to UDP–glucose and fructose in a physiologically reversible reaction, whereas the latter convert sucrose directly to its monomers, glucose and fructose, in a reaction that is physiologically irreversible. Entry of sucrose into metabolism via sucrose synthase requires less consumption of ATP than entry via invertases and offers more opportunities for coordination of this process with demand for carbon by cellular metabolism (Barratt et al. 2009). In this section, we will focus on sucrose synthases and neutral/alkaline invertases since these enzymes are present in the cytosol where sucrose enters general cellular metabolism.

The pathway of starch synthesis was defined by genetic and biochemical research on specialised starch-storing organs, including the cotyledons of legume seeds (Wang et al. 1998). The enzymes that convert glucose-6-phosphate into starch are essentially conserved among plant species (Streb et al. 2009; Stitt and Zeeman 2012) and simplified synthesis and breakdown pathways are shown in Fig. 10.1. Glucose 6-phosphate is converted to glucose 1-phosphate via phosphoglucomutase, and then, glucose 1-phosphate is converted to ADP glucose via ADP glucose pyrophosphorylase with the consumption of ATP. Four different classes of starch synthase isoforms use the glucosyl moiety of ADP glucose to elongate α-1,4-linked glucose chains, into which branch points are introduced by two classes of isoforms of starch-branching enzyme. Debranching enzymes called isoamylases subsequently cleave some of the α-1,6 linkages to create a branched polymer—amylopectin—that becomes organised to form the matrix of the starch granule. Within the matrix, a fifth class of isoform of starch synthase called granule-bound starch synthase generates amylose, an essentially linear α-1,4-linked polymer (Fig. 10.1a). Amylose makes up ca. 25 % of the starch granule in non-photosynthetic organs, and typically less than 15 % of the starch granule in leaves (Wang et al. 1998; Zeeman et al. 2010).

The pathway of starch synthesis occurs inside plastids in almost all plant organs. In chloroplasts, glucose 6-phosphate is synthesised from the Calvin–Benson cycle intermediate fructose 6-phosphate (Fig. 10.1a), whereas in non-photosynthetic plastids, it is imported via a phosphate-exchange translocator from the cytosol. The exception to this picture is the developing endosperm of cereals, where the conversion of glucose 6-phosphate to the starch synthase substrate ADP glucose occurs largely in the cytosol, and ADP glucose is subsequently imported into the plastid for starch synthesis (James et al. 2003).

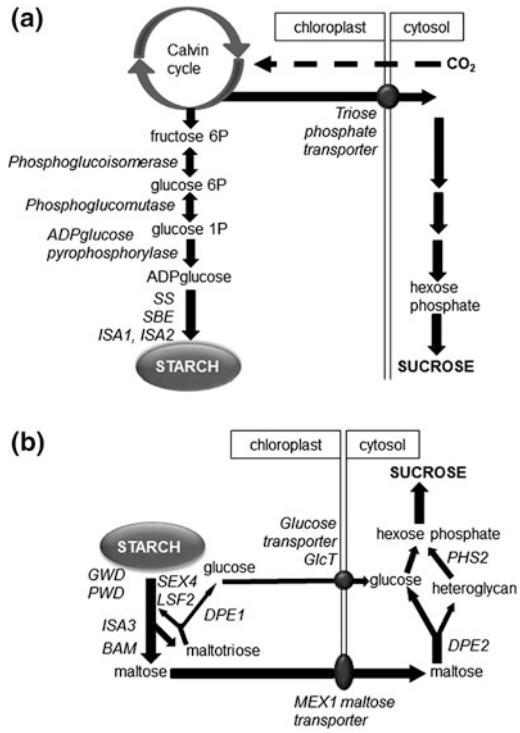


Fig. 10.1 Pathways of starch synthesis (**a**) and degradation (**b**) in the leaf. Much of this information has been obtained through studies in *Arabidopsis*. **a** Pathway of starch synthesis during the day. Intermediates from the Calvin–Benson cycle are used for the synthesis of sucrose and starch. Starch is synthesised inside the chloroplast; sucrose is synthesised in the cytosol from triose phosphates, exported from the chloroplast via the triose phosphate transporter. *SS* starch synthase; *SBE* starch-branching enzyme; *ISA* isoamylase. **b** Pathway of starch degradation at night. Phosphorylation by glucan water dikinase (GWD, also called GWD1) and phosphoglucan water dikinase (PWD, also called GWD3) makes the granule surface accessible to β-amylases (BAM) and isoamylases (ISA) following removal of the phosphate by the enzyme SEX4 (a glucan phosphate phosphatase). The hydrolysis products maltose, and some longer maltooligosaccharides are further metabolised by the glucanotransferase, DPE1, to glucose. Maltose is exported from the chloroplast via the maltose transporter, MEX1, and glucose via the GlcT transporter. In the cytosol, maltose is converted to hexose phosphates via the glucanotransferase, DPE2. The next reactions involve a putative heteroglycan as a glucosyl acceptor and the glucan phosphorylase, PHS2. The size of the arrows indicates approximately the importance of each pathway; for example, greater export is via MEX1 than GlcT

Major advances in understanding starch degradation have come from research in *Arabidopsis* (Fig. 10.1b) (Kötting et al. 2010; Smith et al. 2005; Smith 2012; Stitt and Zeeman 2012). The pathway in *Arabidopsis* leaves is very different from the well-established pathway in the endosperm of germinating cereal grains. Briefly, this pathway is composed of three major steps. The first is a reduction in the organisation of the starch granule surface, making it more accessible to hydrolytic attack. This is brought about by phosphorylation of small numbers of glucosyl residues within amylopectin molecules by the enzymes glucan, water dikinase (GWD) and phosphoglucan, water dikinase (PWD, also called GWD3). The second step is the hydrolysis of the starch polymers by ß-amylases (BAM) and isoamylase 3 (ISA3). Hydrolysis is facilitated by dephosphorylation of amylopectin chains by the phosphoglucan phosphatases SEX4 and LSF2 (Zeeman et al. 2010; Santelia et al. 2012; Smith 2012). The products of hydrolysis are maltose and short linear glucans. These glucans are further metabolised via a glucanotransferase (disproportionating enzyme) DPE1, yielding glucose. Finally, maltose and glucose are exported via specific transporters (MEX1 and GlcT, respectively) to the cytosol, where they are metabolised to sucrose via a complex pathway involving a second glucanotransferase DPE2 and a glucan phosphorylase PHS2 (Ruzanski et al. 2013).

10.3 Genes Encoding Enzymes of Starch and Sucrose Metabolism in *Lotus japonicus*

Candidate genes encoding enzymes involved in sucrose catabolism and in the synthesis and degradation of starch were identified by Vriet et al. (2010) by searching the Miyakogusa genome database (http://www.kazusa.jp/lotus/index.html) using the gene names as keywords and/or by performing Blastp searches using the protein

sequences of their *Arabidopsis* homologues (Kötting et al. 2010; Zeeman et al. 2010). Blastn or tBlastn searches were also performed against the LjEST database. Sequences of all the genes identified were submitted to multiple alignments to determine whether they indeed corresponded to the *L. japonicus* homologue of the *Arabidopsis* gene in question. Where definitive, all *L. japonicus* genes encoding enzymes of starch metabolism were given the same name as their *Arabidopsis* homologues to aid later comparisons. Whenever possible, these genes were mapped onto the genetic linkage map of *L. japonicus* using information provided by the Miyakogusa genome database (version 1.0) (Vriet et al. 2010). Progress in genome sequencing of *L. japonicus* since 2010 has allowed us to update the list of orthologues of the *Arabidopsis* genes in this species for this volume (using version 2.5 and Lj3.0, a prerelease version), and data from both these versions are presented in Table 10.1. We have also included genes encoding enzymes of the committed steps in sucrose synthesis and catabolism.

It is clear that almost all of the major enzymes and isoform classes of starch and sucrose metabolism present in *Arabidopsis* are also present in *L. japonicus*, although in several cases duplications have led to different numbers of isoforms of a particular class in the two species. For example, as in many species, *L. japonicus* has two genes encoding starch synthase 2 and granule-bound starch synthase, whereas *Arabidopsis* has only one gene encoding each isoform class. A notable exception to the generally high level of conservation is in the β-amylase gene family. The BAM4 isoform from *Arabidopsis* does not appear to be represented in *L. japonicus*. This is of particular interest for two reasons. First, in *Arabidopsis,* this isoform is in the chloroplast, together with BAM1 and BAM3, and thus is likely to be involved in some way with starch degradation. Isoforms BAM 5 to BAM 9 are not chloroplastic and are not involved in starch degradation. Second, BAM4 in *Arabidopsis* is catalytically inactive and yet is essential for normal rates of starch degradation. It has been proposed to play a regulatory role, perhaps as part of a protein complex (Fulton et al. 2008). Thus, *L. japonicus* apparently lacks a BAM that is regulatory and colocated with starch.

10.4 Spatial and Temporal Patterns of Expression of Genes Encoding Enzymes of Starch and Sucrose Metabolism

We have investigated the spatial and developmental expression patterns of *L. japonicus* genes encoding enzymes of starch metabolism using the transcriptomic data within the Lotus Gene Expression Atlas (http://mtgea.noble.org/v3/) (Verdier et al. 2013). Expression patterns of the genes encoding enzymes of sucrose catabolism in *L. japonicus* have been published previously (Horst et al. 2007; Welham et al. 2009) and were used to validate the robustness of this analysis. The pattern and level of expression obtained for four sucrose synthase genes from the Lotus Gene Expression Atlas were very similar to those published (Horst et al. 2007); *LjSUS1* was the main gene expressed in leaves, stem, uninoculated roots and flowers, whereas *LjSUS3* was the most highly expressed gene in nodules. Expression of *LjSUS3* was higher in nodules than in any other organ.

Figure 10.2 shows the absolute transcript levels of starch genes in a number of organs sampled according to the information in the Atlas. Genes encoding the plastidial isoforms of phosphoglucoisomerase and phosphoglucomutase and the major small subunit of ADP glucose pyrophosphorylase (*LjPGI1*, *LjPGM1* and *LjAPS1*) were significantly expressed in all organs analysed (Fig. 10.2). *LjAPL1* was the main gene encoding a large subunit of ADP glucose pyrophosphorylase that was expressed in leaves, with lower expression in stems and flowers, and very little to no expression in roots, nodules and seeds. In contrast, *LjAPL3* was the most highly expressed large subunit gene in all organs except leaves, where its expression was lower than that of *LjAPL1*. *LjAPL2* was only very weakly expressed in leaves, but it was the

Table 10.1 Sucrose and starch metabolism genes in *Arabidopsis thaliana* and their homologues in *Lotus japonicus*

Enzyme name (at gene abbreviation)	At gene number	Lj release 3.0 gene number	At/Lj % ID	Lj 2.5 genome clone	Gene name	Reference for Lj mutants
Sucrose metabolism						
Sucrose phosphate synthase 1F (ATSPS1F)	AT5G20280	Lj4g1893200.1	72	chr4.CM0003.1230.r2.m		
		Lj0g0058449.1	54			
		Lj2g3101760.1	53			
		Lj3g1955590.1	51			
Sucrose phosphate synthase 2F (ATSPS2F)	AT5G11110	Lj4g1893200.1	66	chr4.CM0003.1230.r2.m		
		Lj0g0058449.1	55			
		Lj2g3101760.1	53			
		Lj3g1955590.1	51	chr3.CM0047.470.r2.d		
Sucrose phosphate synthase 3F (ATSPS3F)	AT1G04920	Lj0g0058449.1	69			
		Lj2g3101760.1	67			
		Lj3g1955590.1	51	chr3.CM0047.470.r2.d		
		Lj4g1893200.1	52	chr4.CM0003.1230.r2.m		
Sucrose phosphate synthase 4F (ATSPS4F)	AT4G10120	Lj3g1955590.1	66	chr3.CM0047.470.r2.d		
		Lj0g0058449.1	53			
		Lj2g3101760.1	53			
		Lj4g1893200.1	50	chr4.CM0003.1230.r2.m		
Sucrose phosphate phosphatase 1 (ATSPP1)	AT1G51420	Lj0g0007959.2*	58	LjT04A22.110.r2.m		
Sucrose phosphate phosphatase 2 (ATSPP2)	AT3G52340	Lj0g0007959.2*	56	LjT04A22.110.r2.m		
Sucrose phosphate phosphatase	AT2G35840	Lj0g0007959.2*	68	LjT04A22.110.r2.m		
Sucrose phosphate phosphatase	AT3G54270	Lj0g0007959.2*	53	LjT04A22.110.r2.m		

(continued)

Table 10.1 (continued)

Enzyme name (at gene abbreviation)	At gene number	Lj release 3.0 gene number	At/Lj % ID	Lj 2.5 genome clone	Gene name	Reference forLj mutants
Sucrose synthase 1 (SUS1)	AT5G20830	Lj4g2215210.1	83	chr4.CM0006.540.r2.m	*LjSUS3*	Horst et al (2007)
		Lj6g1162830.2	80	chr6.CM0013.460.r2.m	*LjSUS1*	Horst et al (2007)
		Lj1g4875640.2	67	chr1.CM0122.2540.r2.m		
		Lj0g0242009.1	59	chr5.CM0239.860.r2.d		
		Lj2g2507780.1	54	chr2.LjT08I01.10.r2.a		
		Lj3g1238910.1	58	chr2.CM0435.700.r2.m		
		Lj6g1468270.2	64			
Sucrose synthase 2 (SUS2)	AT5G49190	Lj1g4875640.2	76	chr1.CM0122.2540.r2.m	*LjSUS2*	
		Lj6g1162830.2	68	chr6.CM0013.460.r2.m		
		Lj4g2215210.1	68	chr4.CM0006.540.r2.m		
		Lj0g0242009.1	66	chr5.CM0239.860.r2.d		
		Lj2g2507780.1	55	chr2.LjT08I01.10.r2.a		
		Lj6g1468270.2	76			
		Lj3g1238910.1	58	chr2.CM0435.700.r2.m		
Sucrose synthase 3 (SUS3)	AT4G02280	Lj1g4875640.2	81	chr1.CM0122.2540.r2.m	*LjSUS2*	
		Lj0g0242009.1	71	chr5.CM0239.860.r2.d	*LjSUS4*	
		Lj4g2215210.1	68	chr4.CM0006.540.r2.m		
		Lj6g1162830.2	67	chr6.CM0013.460.r2.m		
		Lj2g2507780.1	56	chr2.LjT08I01.10.r2.a		
		Lj6g1468270.2	73			
		Lj3g1238910.1	61	chr2.CM0435.700.r2.m		
Sucrose synthase 4 (SUS4)	AT3G43190	Lj4g2215210.1	83	chr4.CM0006.540.r2.m	*LjSUS3*	Horst et al (2007)

(continued)

Table 10.1 (continued)

Enzyme name (at gene abbreviation)	At gene number	Lj release 3.0 gene number	At/Lj % ID	Lj 2.5 genome clone	Gene name	Reference forLj mutants
		Lj6g1162830.2	81	chr6.CM0013.460.r2.m	*LjSUS1*	Horst et al (2007)
		Lj1g4875640.2	68	chr1.CM0122.2540.r2.m		
		Lj0g0242009.1	59	chr5.CM0239.860.r2.d		
		Lj2g2507780.1	54	chr2.LjT08I01.10.r2.a		
		Lj3g1238910.1	57	chr2.CM0435.700.r2.m		
		Lj6g1468270.2	65			
Sucrose synthase 5 (SUS5)	AT5G37180	Lj2g2507780.1	71	chr2.LjT08I01.10.r2.a	*LjSUS5*	
		Lj3g1238910.1	71	chr2.CM0435.700.r2.m	*LjSUS6*	
		Lj1g4875640.2	53	chr1.CM0122.2540.r2.m		
		Lj6g1162830.2	51	chr6.CM0013.460.r2.m		
		Lj4g2215210.1	52	chr4.CM0006.540.r2.m		
		Lj0g0242009.1	45*	chr5.CM0239.860.r2.d		
Sucrose synthase 6 (SUS6)	AT1G73370	Lj2g2507780.1	70	chr2.LjT08I01.10.r2.a	*LjSUS5*	
		Lj3g1238910.1	72	chr2.CM0435.700.r2.m	*LjSUS6*	
		Lj6g1162830.2	53	chr6.CM0013.460.r2.m		
		Lj4g2215210.1	53	chr4.CM0006.540.r2.m		
		Lj1g4875640.2	53	chr1.CM0122.2540.r2.m		
		Lj0g0242009.1	47*	chr5.CM0239.860.r2.d		
Cytosolic invertase 1 (CINV1/A/N-INVG)	AT1G35580	Lj5g1853130.1	82	chr5.CM1667.130.r2.a	*LjINV1*	Welham et al (2009)
		Lj1g3343910.1	73	chr1.CM0113.70.r2.a		
Cytosolic invertase 2 (CINV2A/N-INVI)	AT4G09510	Lj5g1853130.1	84	chr5.CM1667.130.r2.a	*LjINV1*	Welham et al (2009)

(continued)

Table 10.1 (continued)

Enzyme name (at gene abbreviation)	At gene number	Lj release 3.0 gene number	At/Lj % ID	Lj 2.5 genome clone	Gene name	Reference forLj mutants
		Lj1g3343910.1	75	chr1.CM0113.70.r2.a		
Alkaline/neutral invertase A (A/N-INVA)	AT1G56560	Lj5g1530080.1	76	chr5.CM1598.160.r2.m	*LjINV2*	Welham et al (2009)
		Lj4g2253510.1	70	chr4.CM0333.40.r2.m		
		Lj4g2785740.2	74	chr4.CM1622.190.r2.a		
		Lj1g4790720.1, Lj1g4790710.1*		chr1.CM0122.260.r2.m	*LjINV3*	
Alkaline/neutral invertase B (A/N-INVB)	AT4G34860	Lj1g3343910.1	77	chr1.CM0113.70.r2.a	*LjINV7*	
		Lj5g1853130.1	79	chr5.CM1667.130.r2.a		
Alkaline/neutral invertase C (A/N-INVC)	AT3G06500	Lj4g2253510.1	63	chr4.CM0333.40.r2.m	*LjINV4*	
		Lj5g1530080.1	69	chr5.CM1598.160.r2.m		
		Lj4g2785740.2	71	chr4.CM1622.190.r2.a		
		Lj1g4790720.1, Lj1g4790710.1*		chr1.CM0122.260.r2.m	*LjINV3*	
Alkaline/neutral invertase D (A/N-INVD)	AT1G22650	Lj5g1853130.1	83	chr5.CM1667.130.r2.a	*LjINV1*	Welham et al (2009)
		Lj1g3343910.1	74	chr1.CM0113.70.r2.a		
Alkaline/neutral invertase E (A/N-INVE)	AT5G22510	Lj4g2785740.2	87	chr4.CM1622.190.r2.a	*LjINV5*	
		Lj5g1530080.1	68	chr5.CM1598.160.r2.m		
		Lj4g2253510.1	71	chr4.CM0333.40.r2.m		
		Lj1g4790720.1, Lj1g4790710.1*		chr1.CM0122.260.r2.m		
Alkaline/neutral invertase F (A/N-INVF)	AT1G72000	Lj5g1853130.1	81	chr5.CM1667.130.r2.a	*LjINV1*	Welham et al (2009)
		Lj1g3343910.1	76	chr1.CM0113.70.r2.a		
Alkaline/neutral invertase F (A/N-INVH)	AT3G05820	Lj5g1530080.1	68	chr5.CM1598.160.r2.m	*LjINV2*	Welham et al (2009)
		Lj4g2253510.1	63	chr4.CM0333.40.r2.m		
		Lj4g2785740.2	70	chr4.CM1622.190.r2.a		

(continued)

Table 10.1 (continued)

Enzyme name (at gene abbreviation)	At gene number	Lj release 3.0 gene number	At/Lj % ID	Lj 2.5 genome clone	Gene name	Reference forLj mutants
		Lj1g4790720.1, Lj1g4790710.1*		chr1.CM0122.260.r2.m	*LjINV3*	
Starch metabolism						
Phosphoglucoisomerase, plastidial (PGI1)	At4g24620	Lj0g0299609.1	85	chr1.LjT29N14.30.r2.d	*LjPGI1*	Vriet et al (2010)
		Lj1g1582530.1	81			
Phosphoglucomutase, plastidial (PGM1)	At5g51820	Lj5g2029660.2	87	chr5.CM0200.200.r2.m	*LjPGM1*	Vriet et al (2010)
		Lj4g3099620.1	60			
ADP glucose pyrophosphorylase SS (APS1/ADG1)	At5g48300	Lj2g3337310.1	83	chr2.CM0191.60.r2.m	*LjAPS1*	Vriet et al (2010)
ADP glucose pyrophosphorylase LS1 (APL1/ADG2)	At5g19220	Lj4g2400900.1	76	chr4.CM0387.950.r2.m	*LjAPL1*	Vriet et al (2010)
		Lj1g2126130.1	70	chr1.CM0952.140.r2.m		
ADP glucose pyrophosphorylase LS2 (APL2)	At1g27680	Lj1g2126130.1	73	chr1.CM0952.140.r2.m	*LjAPL2a*	Vriet et al (2010)
				chr5.CM0077.220.r2.m		
		Lj1g3384550.1	67	chr1.CM0113.1080.r2.d		
		Lj3g3362020.2	72	chr3.CM0091.1230.r2.d		
		Lj1g4819990.2	66	chr1.CM0122.1120.r2.m		
		Lj4g2400900.1	65	chr4.CM0387.950.r2.m		
ADP glucose pyrophosphorylase LS3 (APL3)	At4g39210*	Lj3g3362020.2	75	chr3.CM0091.1230.r2.d		
		Lj1g3384550.1	72	chr1.CM0113.1080.r2.d	*LjAPL3*	
		Lj1g2126130.1	67	chr1.CM0952.140.r2.m		
				chr5.CM0077.220.r2.m		
ADP glucose pyrophosphorylase LS4 (APL4)	At2g21590	Lj1g3384550.1	71	chr3.CM0091.1230.r2.d	*LjAPL4*	

(continued)

Table 10.1 (continued)

Enzyme name (at gene abbreviation)	At gene number	Lj release 3.0 gene number	At/Lj % ID	Lj 2.5 genome clone	Gene name	Reference forLj mutants
		Lj1g2126130.1	70	chr1.CM0113.1080.r2.d		
				chr1.CM0952.140.r2.m		
				chr5.CM0077.220.r2.m		
Starch synthase I (SS1)	At5g24300	Lj1g0958830.1 Lj2g0909570.1	82	chr1.LjT38J19.60.r2.d		
Starch synthase II (SS2)	At3g01180	Lj2g1058180.1	64	chr2.CM1835.140.r2.m	*LjSSIIa*	
		Lj1g2975950.2	63	chr1.CM0141.60.r2.d	*LjSSIIb*	
Starch synthase III (SS3)	At1g11720	Lj6g1874670.1	61	LjT22N06.110.r2.d		
		Lj4g0151570.1	76			
		Lj0g0048129.1	47*			
Starch synthase IV (SS4/SSIV)	At4g18240	Lj4g0281070.1	59	chr4.CM0007.1020.r2.a	*LjSS4*	
Starch synthase V/glycogen synthase-like	AT5G65685	Lj2g1468340.1*	54	chr2.CM0177.640.r2.m	*LjSS5*	
Granule-bound starch synthase (putative GBSS)	AT1G32900	Lj5g1262950.2*	78	chr5.CM0431.180.r2.d	*LjGBSSIa*	
		Lj3g2515230.2*	70	chr3.CM0208.40.r2.d	*LjGBSSIb*	
Starch-branching enzyme, class II (SBE2.2)	At5g03650	Lj1g4699750.1	86	chr1.CM0178.250.r2.m		
		Lj1g2294860.1	56			
		Lj0g0276889.1	58			
Starch-branching enzyme, class II (BE3/SBE2.1)	At2g36390	Lj1g4699750.1	84	chr1.CM0178.250.r2.m		
		Lj1g2294860.1	56			
		Lj0g0276889.1	58			
Starch-branching enzyme, class III (BE1)	At3g20440	Lj0g0332909.1	73	LjSGA 048386.1		
Isoamylase 1 (ISA1)	At2g39930	Lj4g3015110.3	77	chr4.CM0004.2080.r2.a	*LjISA1*	
Isoamylase 2 (ISA2/DBE1)	At1g03310	Lj0g0304069.1	55	LjSGA_028198.1	*LjISA2*	
Isoamylase 3 (ISA3)	At4g09020	Lj1g1525840.1	77	LjSGA 033928.1		
Limit dextrinase (LDA/PU1)	At5g04360	Lj5g1698870.1	72	chr5.CM0909.690.r2.m	*LjLDA*	

(continued)

Table 10.1 (continued)

Enzyme name (at gene abbreviation)	At gene number	Lj release 3.0 gene number	At/Lj % ID	Lj 2.5 genome clone	Gene name	Reference forLj mutants
Glucan, water dikinase 1 (GWD1/SEX1)	At1g10760	Lj4g1083940.2	62	chr4.CM0161.100.r2.m	*LjGWD1*	Vriet et al (2010)
		Lj0g0189879.1	71			
		Lj0g0159529.1	71			
Glucan, water dikinase 2 (ATGWD2)	At4g24450	Lj4g1083940.2	46	chr4.CM0161.100.r2.m	*LjGWD2*	
		Lj0g0189879.1	49			
		Lj0g0159529.1	48			
Glucan, water dikinase 3 (ATGWD3, PWD)	At5g26570	Lj5g2302970.2	62	chr5.LjT42F22.160.r2.m	*LjGWD3*	Vriet et al (2010)
Phosphoglucan phosphatase (SEX4, DSP4)	At3g52180	Lj5g1699470.1	60*	n/a		
Phosphoglucan phosphatase (LSF2)	At3g10940	Lj1g1182240.1	76*	chr1.CM0125.220.r2.d	*LjLSF2*	
Alpha-amylase 1 (AMY1)	At4g25000	Lj5g0290420.1	65*	chr5.CM0852.140.r2.m	*LjAMY1*	
Alpha-amylase 2 (AMY2)	At1g76130	Lj5g0255820.2	74*	n/a		
Alpha-amylase 3 (AMY3)	At1g69830	Lj2g1338900.1	60	chr2.CM0608.1120.r2.m	*LjAMY3*	
Beta-amylase 1 (BAM1/BMY7)	At3g23920	Lj2g0632120.1, Lj2g0632130.1	87 87	chr2.CM0435.1010.r2.a	*LjBAM1*	
Beta-amylase 2 (BAM2/BMY9)	At4g00490	Lj1g2659300.1	71	chr1.LjT34L14.60.r2.m	*LjBAM2*	
Beta-amylase 3 (BAM3/BMY8)	At4g17090	Lj2g1988790.1	67	chr2.CM0021.1150.r2.a	*LjBAM3*	
Beta-amylase 4 (BAM4/BMY6)	At5g55700	n/a	n/a	n/a		
Beta-amylase 5 (BAM5/BMY1)	At4g15210	Lj3g2888320.1	65	chr3.CM0152.240.r2.m	*LjBAM5*	
Beta-amylase 6 (BAM6/BMY5)	At2g32290	Lj0g0359449.1	65			
		Lj3g2888320.1	66	chr3.CM0152.240.r2.m		
Beta-amylase 7 (BAM7/BMY4)	At2g45880	Lj1g2659300.1	67	chr1.LjT34L14.60.r2.m	*LjBAM7*	
Beta-amylase 8 (BAM8/BMY2)	At5g45300	Lj2g1550200.2	69	chr2.CM0803.520.r2.m	*LjBAM8*	
Beta-amylase 9 (BAM9/BMY3)	At5g18670	Lj6g1918250.1	56*	chr6.LjT15B19.110.r2.m	*LjBAM9*	
Glucanotransferase/disproportionating enzyme (DPE1)	At5g64860	Lj1g1218940.2	74	chr1.CM0032.10.r2.m	*LjDPE1*	

(continued)

Table 10.1 (continued)

Enzyme name (at gene abbreviation)	At gene number	Lj release 3.0 gene number	At/Lj % ID	Lj 2.5 genome clone	Gene name	Reference forLj mutants
Glucanotransferase/disproportionating enzyme (DPE2)	At2g40840	Lj1g4081830.2	71	chr1.CM1911.70.r2.m	*LjDPE2*	
Alpha-glucan phosphorylase, plastidial (PHS1)	At3g29320	Lj2g1079510.1	72	chr2.CM1882.170.r2.a	*LjPHS1*	
		Lj0g0360239.1	69	chr6.LjB08M07.90.r2.m		
		Lj6g2006830.2	45*	chr6.CM0114.510.r2.m		
Alpha-glucan phosphorylase, cytosolic (PHS2)	At3g46970	Lj6g2006830.2	47*	chr6.CM0114.510.r2.m		
				chr6.LjB08M07.90.r2.m		
Maltose transporter (MEX1/RCP1)	At5g17520	Lj3g3639950.2	64*	chr3.CM0127.650.r2.m	*LjMEX1*	
Plastidic glucose translocator (pGlcT)	At5g16150	Lj0g0342169.1 Lj1g3300010 Lj0g0342169.2	77 76 61	LjSGA 043891.1 LjSGA 062891.1 LjSGA 099348.1 LjSGA 137187.1.1	*LjPGlcT*	

List of genes coding for proteins involved in the core pathways of sucrose and starch metabolism in *Arabidopsis* and their homologues in *Lotus japonicus*. Genome clones in *L. japonicus* were identified by performing Blastp with the *Arabidopsis* protein sequences against the Miyakogusa.jp v2.5 and v3.0 (early access) genome databases developed by the Kasuza DNA Research Institute (http://www.kazusa.or.jp/lotus/; Sato et al. 2008). Unless otherwise indicated (with an asterisk), only hits with a null E-value, and a percentage identity of sequence around 50 % were retrieved. Where the chromosome number is 0, the position of the sequence on the genome is not yet determined. For *L. japonicus* genes for which mutants and/or transgenic plants have been generated and characterised, reference is given to the corresponding publication. *Abbreviations At Arabidopsis thaliana*; *Lj Lotus japonicus*; for ADP glucose pyrophosphorylase, *LS* large subunit, *SS* small subunit. Commentary on enzymes not discussed in the text for which there is no further information in *Lotus*: sucrose phosphate synthase and sucrose phosphate phosphatase are the two committed enzymes of sucrose synthesis in plants (Lunn and MacRae 2003); isoamylases 1 and 2 are involved in the synthesis of amylopectin; isoamylase 3 is involved in starch degradation (Delatte et al. 2006; Streb et al. 2008); limit dextrinase catalyses a similar reaction to isoamylase, and in *Arabidopsis,* the enzyme is plastidial, but largely redundant with respect to isoamylases (Delatte et al. 2006); glucan water dikinase 2 is extra-plastidal and not required for starch degradation in *Arabidopsis* (Glaring et al. 2007); α-amylases are not required for starch degradation in *Arabidopsis* leaves (AMY3 is plastidial and may play a redundant role in starch degradation, whereas AMY1 and AMY2 are not plastidial; Yu et al. 2005); β-amylases 5–9 are not plastidial in *Arabidopsis*; BAM7 and BAM8 are nuclear DNA-binding proteins (Fulton et al. 2008; Reinhold et al. 2011); α-glucan phosphorylase 1 (PHS1) is not required for starch degradation in *Arabidopsis* leaves in normal conditions, but may be important in stress conditions (Zeeman et al. 2004)

second most highly expressed gene encoding a large subunit in every other organ. The transcript level of *LjAPL4* was low or at the threshold of detection in all organs analysed (Fig. 10.2).

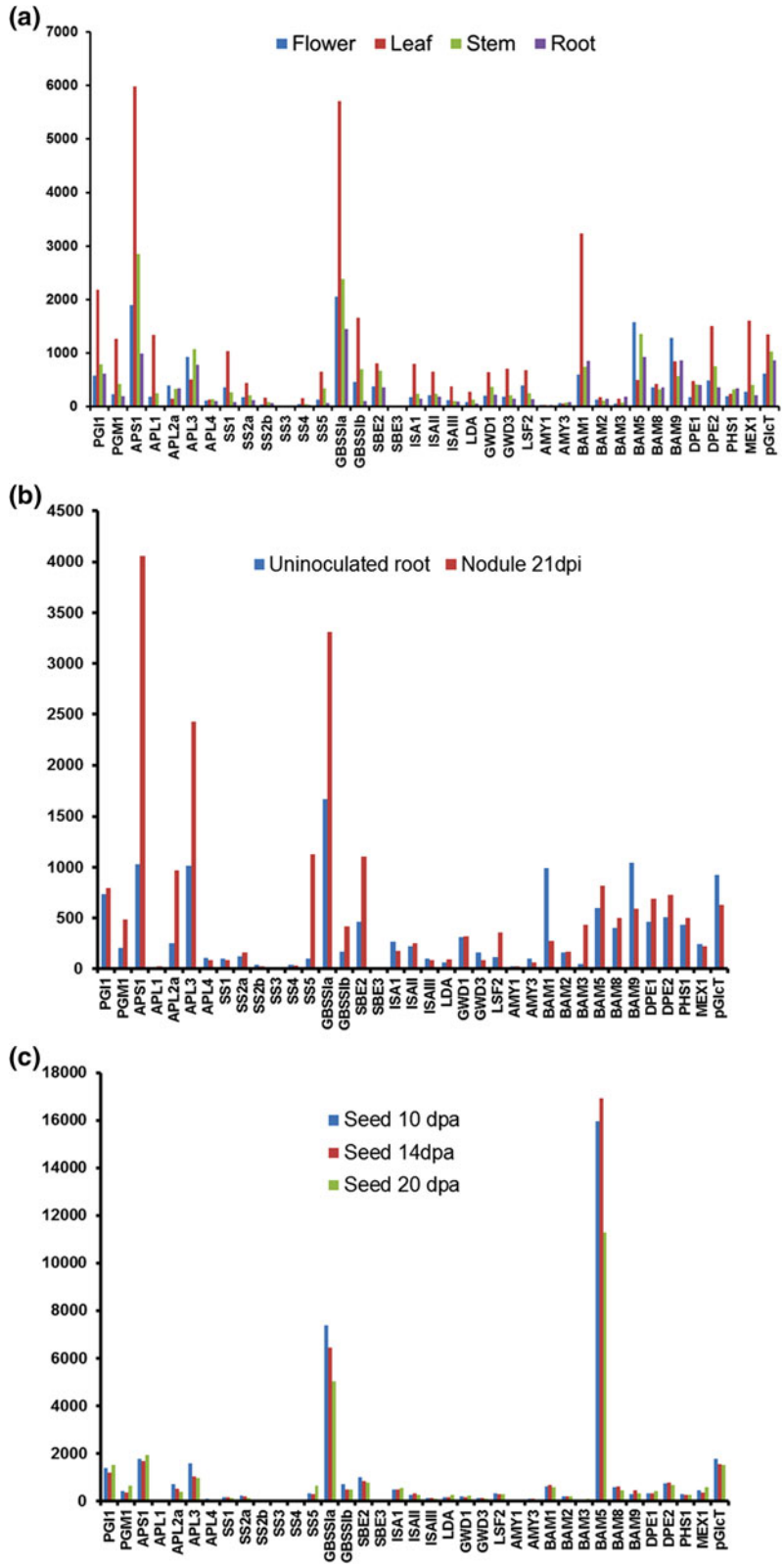
Tissue-specific expression patterns were also observed for genes encoding members of several families of starch-synthesising and starch-degrading enzymes. Thus, transcript levels of *GBSS1a* and *GBSS1b* were similar in shoots, but levels for *GBSS1a* were much greater than those of *GBSS1b* in most other organs including roots and seeds. *LjBAM1* was the main gene encoding a β-amylase expressed in leaves, whereas BAM3 was expressed only at low levels. This finding (together with the one mentioned in Sect. 10.3 regarding family members) points to major differences between *Arabidopsis* and *L. japonicus* in the types of BAM responsible for starch degradation. In *Arabidopsis*, BAM1 is a minor isoform that may be expressed largely in guard cells rather than mesophyll cells, whereas BAM3 is the major isoform responsible for starch degradation in the chloroplast (Sparla et al. 2006; Fulton et al. 2008). *LjDPE1*, *LjDPE2*, *LjpGlcT* and *LjMEX1* genes were ubiquitously expressed, suggesting that the encoded glucanotransferases and glucose and maltose transporters are involved in starch degradation in all organs. However, in contrast to leaves in which their expression was approximately equal, *LjpGlcT* expression was ca. threefold higher than that of *LjMEX1* in seeds, roots and nodules.

The expression of several genes was higher in nodules at 21 days after inoculation of roots with *Rhizobium* bacteria (Fig. 10.2b) than in roots at the time of inoculation, suggesting that the encoded proteins may be involved in starch metabolism in nodules. This was the case for the starch synthesis genes *LjPGM1, LjAPS1, LjAPL2a* and *LjAPL3* as well as *LjGBSS1a, LjGBSS1b, LjSS5* and *LjSBE2*. The expression of *LjSS5,* in particular, was higher in nodules than in any other organ (Fig. 10.2). In *Arabidopsis,* this starch-synthase-like protein has not been shown to have starch synthase activity or to be involved in starch synthesis. It would be interesting to investigate whether the protein makes a contribution to nodule starch metabolism in *L. japonicus*. Transcript levels of several genes encoding enzymes involved in starch degradation also significantly differed between nodules and roots. Thus *LjBAM3* transcript levels were higher in nodules than in uninoculated roots, whereas the opposite was true for *LjBAM1*.

There was little change in transcript levels of genes encoding enzymes of starch synthesis and degradation during seed development (Fig. 10.2c). Starch turnover has not been studied in detail in *L. japonicus* seeds, but starch is not a major component of the seed once it is mature (Dam et al. 2009). Of especial note, however, is the high level of *BAM5* expression. As previously mentioned, it is not believed to have a role in starch degradation in *Arabidopsis* leaves because it is extra-plastidial (Streb and Zeeman 2012), but the high transcript level in *L. japonicus* seeds indicates that the protein may be abundant. BAM activity is also high during seed development in soybean; this activity is largely attributable to a gene with 85 % identity to *LjBAM5* (Kim and Krishnan 2010). Both *L. japonicus* and soybean (Monma et al. 1991) produce starch during seed development, and then towards the end of development, the starch is degraded (Fig. 10.3). The dry seed of Lotus contains less than 1 % starch (Dam et al. 2009). However, a soybean mutant that lacks β-amylase activity in the seed because of a deletion in the BAM5-like gene is not defective in starch degradation or in oil or total protein accumulation (Hildebrand and Hymowitz 1981). The fact that β-amylase is not required for seed starch degradation suggests that it may function as a seed storage protein. A storage function has also been proposed for the large amounts of β-amylase that accumulate in the tap roots of alfalfa (Gana et al. 1998).

10.5 *Lotus japonicus* Mutants Defective in Starch or Sucrose Metabolism

Although transcript profiling can provide clues about the function of genes, direct evidence requires the isolation of plants with altered levels


(a)
Flower
Leaf
Stem
Root
7000
6000
5000
4000
3000
2000
1000
0
PGI1
PGM1
APS1
APL1
APL2a
APL3
APL4
SS1
SS2a
SS2b
SS3
SS4
SS5
GBSSIa
GBSSIb
SBE2
SBE3
ISA1
ISAII
ISAIII
LDA
GWD1
GWD3
LSF2
AMY1
AMY3
BAM1
BAM2
BAM3
BAM5
BAM8
BAM9
DPE1
DPE2
PHS1
MEX1
pGlcT
(b)
Uninoculated root
Nodule 21dpi
4500
4000
3500
3000
2500
2000
1500
1000
500
0
(c)
Seed 10 dpa
Seed 14dpa
Seed 20 dpa
18000
16000
14000
12000
10000
8000
6000
4000
2000
0

◀ **Fig. 10.2** Transcript levels (absolute) of *Lotus* starch genes in various organs. The expression of genes identified in Table 10.1 were analysed using the *Lotus japonicus* gene expression atlas versus 3 from the Noble Foundation (http://ljgea.noble.org/v3/). Where genes were represented by more than one probeset, median values were used. Organs and tissue used by the atlas: *leaf* trifolia without their petioles from 28-*d*-old plants; *stem* stems from 28-*d*-old plants; *flower* fully developed flowers; *root* roots from 28-*d*-old plants; *nod0d* N-starved roots at 0 dpi (control for nodules); *nod21d* mature N-fixing nodules at 21 dpi; *seed10d* developing seeds, 10 dpa (late embryogenesis); *seed14d* developing seeds, 14 dpa (accumulation of storage compounds); *seed20d* developing seeds, 20 dpa (physiologically mature seeds, onset of dessication). **a** A comparison of the transcript levels in different organs; **b** a comparison of transcript levels following inoculation with *Rhizobium*; **c** a comparison at three different stages of seed development. *Dpa* days post-anthesis; *dpi* days post-inoculation

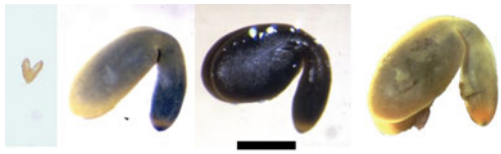


Fig. 10.3 Starch accumulation in embryos during development. Embryos of *L. japonicus* Gifu were removed from their seed coat and depigmented by heating in an ethanol/chloroform/water mix. The starch was then stained with Lugol's iodine solution (indicated by *blue/black* colouration). Stages of development from *left* to *right* early heart shaped; mid-cotyledonary; maximum fresh weight; drying, near-mature embryo (testa peelable). Note the root cap staining as well as the cotyledons and hypocotyl. Bar = ca. 1 mm

of gene expression. Table 10.1 shows mutants of *L. japonicus* that carry mutations in specific genes encoding enzymes of starch and sucrose metabolism. Both forward genetic screens and TILLING (see Chap. 22) have been used to isolate mutants with altered starch content (revealed by iodine staining). For the pathway of starch biosynthesis, the mutant phenotypes indicate that the same enzymes are required in this model legume as in pea and that *L. japonicus* conforms to the standard dicotyledonous pathway established in *Arabidopsis*. However, the effect of elimination of starch synthesis on plant growth and development differs significantly between the legumes and *Arabidopsis*. Whereas the almost starchless *pgm* mutant of *Arabidopsis* grows more slowly than wild-type plants in most day lengths (Caspar et al. 1985), the equivalent mutants of *L. japonicus* and pea (the *rug3* mutant), which also lack starch, are not impaired in growth (Harrison et al. 1998; Vriet et al. 2010) suggesting that the transitory storage of starch is dispensable in these legume species. They also retain functional root nodules and—at least for *L. japonicus*—mycorrhizal associations, showing that starch storage and turnover is not essential for the establishment of successful *Rhizobium* and mycorrhizal symbioses (Vriet et al. 2010; Gutjahr et al. 2011). It seems likely that, as in *Arabidopsis*, the starchless mutants of legumes have altered patterns of sugar export and utilisation over the day (Stitt and Zeeman 2012). Why these patterns are able to compensate completely for the loss of starch turnover in legumes, but not in *Arabidopsis* remains to be investigated.

The phenotypes of mutants lacking specific enzymes of starch degradation show that at least the first step in the *Arabidopsis* starch degradation pathway—the phosphorylation of the starch granule surface—is also important for the process in *L. japonicus*. Both *gwd* and *pwd* mutants of *L. japonicus* have increased levels of starch in their leaves (Vriet et al. 2010). However, impaired starch degradation caused by loss of GWD activity has a far greater effect on the plant in *L. japonicus* than it does in *Arabidopsis*. The *L. japonicus gwd* mutants grow more slowly in comparison with wild-type plants than *Arabidopsis gwd* mutants, and unlike the *Arabidopsis* mutants, they are largely infertile.

As well as identifying genes already known to encode enzymes of starch metabolism, forward screens for abnormal starch levels in *L. japonicus* also identified several loci not previously known to encode proteins important for starch metabolism (Vriet et al. 2010). Recent data from map-based cloning indicate one of these encodes a pentatricopeptide repeat-containing protein (Vriet, Welham, Brachmann, Edwards, Parniske,

Smith and Wang, unpublished), loss of which results in a slow-growing plant with a low starch content. No protein of this class has previously been reported to be necessary for starch metabolism. A knockout of the equivalent gene in *Arabidopsis* results in slow growth, but not reduced starch content, indicating that the influence on starch content in *L. japonicus* may reflect an indirect rather than a direct requirement for the gene product.

Mutant analysis in pea showed that one isoform of sucrose synthase that is highly expressed in the nodule is the main enzyme responsible for delivering carbon for nitrogen fixation. Mutants lacking this isoform derived their nitrogen largely from the soil rather than from the *Rhizobium* symbiosis (Craig et al. 1999). These mutants retained a small amount of sucrose synthase activity in the nodule, indicating a potential contribution from a second isoform. Using *L. japonicus*, Horst et al. (2007) showed that two isoforms are present in the nodule, *LjSUS1* and *LjSUS3*, and that both contribute to sucrose catabolism. There was no nitrogen fixation in nodules of a double mutant lacking both isoforms, leading to the conclusion that sucrose synthase is vital for nodule function and that its actions cannot be replaced by those of invertases also present in the nodule.

Although neutral/alkaline invertase does not seem to contribute to nodule function, mutant analysis shows that a cytosolic isoform of this class is essential for the cellular development and growth of the *L. japonicus* plant as a whole. Loss of expression of *Ljinv1* resulted in dramatically reduced growth of root and shoot, and changes to tissue organisation, with enlarged root meristems and thickened roots with more cells. The change in cell organisation was also observed in leaves, stems and nodules, and there was a lack of pollen (Welham et al. 2009). A knockout of the equivalent gene in *Arabidopsis* had only a very mild phenotype, but a double mutant lacking two cytosolic invertases closely related to *Ljinv1* resulted in a very strong effect on *Arabidopsis* growth and development (Barratt et al. 2009). The reasons for the importance of this class of invertases are not yet understood, but they may be related to the control of sugar signalling in meristems (Pignocchi, Edwards, Welham, Wang and Smith, unpublished).

10.6 The Relationship Between Starch Content and Growth in *L. japonicus*

There is a general expectation that fast-growing plant species maximise resource acquisition, whereas slow-growing species or accessions rather maximise the conservation of resources. With regard to the metabolism of carbohydrates, this means that faster-growing plants are expected to hold less carbohydrate in reserve as starch that would allow them to cope with unexpected changes in their environmental conditions. This general expectation holds true for *Arabidopsis*: a study of 94 *Arabidopsis* accessions showed a negative correlation between starch level at the end of the night and biomass production (Sulpice et al. 2009). However, the picture for *L. japonicus* is more complex. For the first 6 weeks of growth, two ecotypes with different growth rates accumulated similar amounts of starch during the day and degraded about 95 % of it at night. Starch levels then increased dramatically in both ecotypes. At 8 weeks, only about 70 % of starch accumulated during the day was degraded at night, and this value dropped to 30 % or less by 10 weeks when the plants started to flower. Importantly, the faster-growing ecotype contained 80 % more starch at the end of the day than the slower-growing ecotype at this point (Vriet et al. 2010). We suggest that although there may be a general relationship between starch storage and growth rate during vegetative growth, this relationship no longer holds in perennial plants approaching flowering. Presumably, storage of carbon at this point is controlled by signals that reflect a future requirement for carbon to fuel growth of new vegetative meristems once flowering (and perhaps subsequent winter dormancy) has finished.

The genetic resources available in *L. japonicus* have provided further insights into the role and importance of starch storage for regrowth

and the perennial habit. It had long been speculated that starch stored in roots and stolons of perennial legumes provides an important source of carbon for their regrowth after dormancy or cutting back. However, attempts to test this idea experimentally had led to contradictory conclusions (Vriet et al. 2013). By comparing wild-type and starchless mutant plants of *L. japonicus*, Vriet et al. (2013) showed unambiguously that under at least some growth conditions, starch storage in roots is essential for regrowth following removal of all shoots. A survey of root starch content and the response to cutting back in 24 species of *Lotus* similarly revealed a general correlation between vigorous regrowth and root starch content. Furthermore, species with high root starch and vigorous regrowth were mostly perennials, while species with little root starch and poor regrowth were mostly annuals. Thus, although the capacity to store starch in roots is unlikely to be a primary determinant of the perennial habit, it may well confer advantages on perennial species subject to grazing and/or winter dormancy (Vriet et al. 2013).

10.7 Conclusions: *Lotus japonicus* Resources Provide New Insights into the Relationship Between Growth, and Starch and Sucrose Metabolism

The collection of *L. japonicus* mutants defective in starch and sucrose metabolism is one of the most extensive available, arguably second only to *Arabidopsis*. While comparative genomic analysis reveals that the components of pathways of starch and sucrose metabolism are largely conserved between *L. japonicus* and *Arabidopsis* (Table 10.1), mutant analyses show that their importance for plant growth and development differs substantially between the two species. At least, part of these differences may be attributed to the fact that *Arabidopsis* is an annual, whereas *L. japonicus* is a perennial. For example, in contrast to *Arabidopsis*, loss of starch has little consequence for normal growth in *L. japonicus*. However, starch stored in the roots is important for regrowth following cutting back in this perennial plant. The substantial accumulation of starch in plants approaching flowering also suggests that stored starch may be important for continued vegetative growth post-flowering under at least some conditions. The genetic resources available in *L. japonicus* have also given new insights into the way in which carbon is supplied for nitrogen fixation and subsequent ammonia assimilation in root nodules. The resources may well prove to be of great importance in understanding the engineering of carbon storage and partitioning that will be required in the future to enable non-leguminous crop plants to form nitrogen-fixing associations (Oldroyd 2013).

References

Barratt DH, Derbyshire P, Findlay K, Pike M, Wellner N, Lunn J, Feil R, Simpson C, Maule AJ, Smith AM (2009) Normal growth of *Arabidopsis* requires cytosolic invertase but not sucrose synthase. Proc Natl Acad Sci USA 106:13124–13129

Caspar T, Huber SC, Somerville C (1985) Alterations in growth, photosynthesis, and respiration in a starchless mutant of *Arabidopsis thaliana* (L.) deficient in chloroplast phosphoglucomutase activity. Plant Physiol 79:11–17

Craig J, Barratt P, Tatge H, Dejardin A, Handley L, Gardner CD, Barber L, Wang T, Hedley C, Martin C, Smith AM (1999) Mutations at the rug4 locus alter the carbon and nitrogen metabolism of pea plants through an effect on sucrose synthase. Plant J 17:353–362

Dam S, Laursen BS, Ornfelt JH, Jochimsen B, Staerfeldt HH, Friis C, Nielsen K, Goffard N, Besenbacher S, Krusell L, Sato S, Tabata S, Thøgersen IB, Enghild JJ, Stougaard J (2009) The proteome of seed development in the model legume *Lotus japonicus*. Plant Physiol 149:1325–1340

Delatte D, Umhang M, Trevisan M, Eicke S, Thorneycroft D, Smith SM, Zeeman SC (2006) Evidence for distinct mechanisms of starch granule breakdown in plants. J Biol Chem 281:12050–12059

Fulton DC, Stettler M, Mettler T, Vaughan CK, Li J, Francisco P, Gil M, Reinhold H, Eicke S, Messerli G, Dorken G, Halliday K, Smith AM, Smith SM, Zeeman SC (2008) β-AMYLASE4, a noncatalytic protein required for starch breakdown, acts upstream of three active β-amylases in *Arabidopsis chloroplasts*. Plant Cell 20:1040–1058

Gana JA, Kalengamaliro NE, Cunningham SM, Volenec JJ (1998) Expression of β-amylase from alfalfa taproots. Plant Physiol 118:1495–1505

Glaring MA, Zygadlo A, Thorneycroft D, Schulz A, Smith SM, Blennow A, Baunsgaard L (2007) An extra-plastidial α-glucan, water dikinase from *Arabidopsis* phosphorylates amylopectin in vitro and is not necessary for transient starch degradation. J Exp Bot 58:3949–3960

Gutjahr C, Novero M, Welham T, Wang T, Bonfante P (2011) Root starch accumulation in response to arbuscular mycorrhizal colonization differs among *Lotus japonicus* starch mutants. Planta 234:639–646

Harrison CJ, Hedley CL, Wang TL (1998) Evidence that the rug3 locus of pea (*Pisum sativum* L.) encodes plastidial phosphoglucomutase confirms that the imported substrate for starch synthesis in pea amyloplasts is glucose-6-phosphate. Plant J 13:753–762

Hildebrand D, Hymowitz T (1981) Role of β-amylase in starch metabolism during soybean seed development and germination. Physiol Plant 53:429–434

Horst I, Welham T, Kelly S, Kaneko T, Sato S, Tabata S, Parniske M, Wang TL (2007) TILLING mutants of *Lotus japonicus* reveal that nitrogen assimilation and fixation can occur in the absence of nodule-enhanced sucrose synthase. Plant Physiol 144:806–820

James MG, Denyer K, Myers AM (2003) Starch synthesis in the cereal endosperm. Curr Opin Plant Biol 6:215–222

Kim W-S, Krishnan HB (2010) The lack of beta-amylase activity in soybean cultivar *Altona* sp 1 is associated with a 1.2 kb deletion in the 5′ region of beta-amylase I gene. Crop Sci 50:1942–1949

Kötting O, Kossmann J, Zeeman SC, Lloyd JR (2010) Regulation of starch metabolism: the age of enlightenment? Curr Opin Plant Biol 13:321–329

Lunn JE, MacRae E (2003) New complexities in the synthesis of sucrose. Curr Opin Plant Biol 6:208–214

Oldroyd GED (2013) Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. Nat Rev Microbiol 11:252–263

Monma M, Sugimoto T, Monma M, Kawamura Y, Saio K (1991) Starch breakdown in developing soybean seeds (*Glycine max* cv. Enrei). Agric Biol Chem 55:67–71

Reinhold H, Soyk S, Simkova K, Hostettler C, Marafino J, Mainiero S, Vaughan CK, Monroe JD, Zeeman SC (2011) β-Amylase–like proteins function as transcription factors in *Arabidopsis*, controlling shoot growth and development. Plant Cell 23:1391–1403

Ruzanski C, Smirnova J, Rejzek M, Cockburn D, Pedersen HL, Pike M, Willats WGT, Svensson B, Steup M, Ebenhöh O, Smith AM, Field RA (2013) A bacterial glucanotransferase can replace the complex maltose metabolism required for starch-to-sucrose conversion in leaves at night. J Biol Chem 288:28581–28598

Santelia D, Kötting O, Seung D, Schubert M, Thalmann M, Bischof S, Meekins DA, Lutz A, Patron N, Gentry MS, Allain FH-T, Zeeman SC (2012) The phosphoglucan phosphatase LSF2 (like sex four 2) dephosphorylates starch at the C3-position in *Arabidopsis*. Plant Cell 23:4096–4111

Sato S, Nakamura Y, Kaneko T, Asamizu E, Kato T, Nakao M, Sasamoto S, Watanabe A, Ono A, Kawashima K, Fujishiro T, Katoh M, Kohara M, Kishida Y, Minami C, Nakayama S, Nakazaki N, Shimizu Y, Shinpo S, Takahashi C, Wada T, Yamada M, Ohmido N, Hayashi M, Fukui K, Baba T, Nakamichi T, Mori H, Tabata S (2008) Genome structure of the legume, *Lotus japonicus*. DNA Res 15:227–239

Smith AM (2012) Starch in the *Arabidopsis* plant. Starch/Staerke 64:421–434

Smith AM, Zeeman SC, Smith SM (2005) Starch degradation. Annu Rev Plant Biol 56:73–97

Sparla F, Costa A, Schiavo FL, Pupillo P, Trost P (2006) Redox regulation of a novel plastid-targeted β-amylase of *Arabidopsis thaliana*. Plant Physiol 141:840–850

Stitt M, Zeeman SC (2012) Starch turnover: pathways, regulation and role in growth. Curr Opin Plant Biol 15:282–292

Streb S, Delatte T, Umhang M, Eicke S, Schorderet M, Reinhardt D, Zeeman SC (2008) Starch granule biosynthesis in *Arabidopsis* is abolished by removal of all debranching enzymes but restored by the subsequent removal of an endoamylase. Plant Cell 20:3448–3466

Streb S, Zeeman S (2012) Starch metabolism in *Arabidopsis*. In: The *Arabidopsis* Book 10:e0160, doi: 10.1199/tab.0160

Streb S, Egli B, Eicke S, Zeeman SC (2009) The debate on the pathway of starch synthesis: a closer look at low-starch mutants lacking plastidial phosphoglucomutase supports the chloroplast-localized pathway. Plant Physiol 151:1769–1772

Sulpice R, Pyl ET, Ishihara H, Trenkamp S, Steinfath M, Witucka-Wall H, Gibon Y, Usadel B, Poree F, Piques MC, Von Korff M, Steinhauser MC, Keurentjes JJB, Guenther M, Hoehne M, Selbig J, Fernie AR, Altmann T, Stitt M (2009) Starch as a major integrator in the regulation of plant growth. Proc Natl Acad Sci U S A 106:10348–10353

Vargas WA, Salerno GL (2010) The Cinderella story of sucrose hydrolysis: alkaline/neutral invertases, from cyanobacteria to unforeseen roles in plant cytosol and organelles. Plant Sci 178:1–8

Verdier J, Torres-Jerez I, Wang M, Andriankaja A, Allen SN, He J, Tang Y, Murray JD, Udvardi MK (2013) Establishment of the *Lotus japonicus* Gene Expression Atlas (LjGEA) and its use to explore legume seed maturation. Plant J 74:351–362

Vriet C, Welham T, Brachmann A, Pike M, Pike J, Perry J, Parniske M, Sato S, Tabata S, Smith AM, Wang TL (2010) A suite of *Lotus japonicus* starch mutants reveals both conserved and novel features of starch metabolism. Plant Physiol 154:643–655

Vriet C, Smith AM, Wang TL (2013) Root starch reserves are necessary for vigorous re-growth following cutting back in *Lotus japonicus*. PLOS ONE (in press)

Wang TL, Bogracheva TY, Hedley CL (1998) Starch: As simple as A, B, C? J Exp Bot 49:481–502

Welham T, Pike J, Horst I, Flemetakis E, Katinakis P, Kaneko T, Sato S, Tabata S, Perry J, Parniske M, Wang TL (2009) A cytosolic invertase is required for normal growth and cell development in the model legume, *Lotus japonicus*. J Exp Bot 60:3353–3365

Yu TS, Zeeman SC, Thorneycroft D, Fulton DC, Dunstan H, Lue WL, Hegemann B, Tung SY, Umemoto T, Chapple A, Tsai DL, Wang SM, Smith AM, Chen J, Smith SM (2005) α-amylase is not required for breakdown of transitory starch in *Arabidopsis* leaves. J Biol Chem 280:9773–9779

Zeeman SC, Thorneycroft D, Schupp N, Chapple A, Weck M, Dunstan H, Halidmann P, Bechtold N, Smith AM, Smith SM (2004) Plastidial α-glucan phosphorylase is not required for starch degradation in *Arabidopsis* leaves but has a role in the tolerance of abiotic stress. Plant Physiol 135:849–858

Zeeman SC, Kossmann J, Smith AM (2010) Starch: its metabolism, evolution, and biotechnological modification in plants. Annu Rev Plant Biol 61:209–234

Genes Involved in Ammonium Assimilation

11

Carmen M. Pérez-Delgado, Margarita García-Calderón, Alfredo Credali, José M. Vega, Marco Betti, and Antonio J. Márquez

Abstract

Ammonium resulting from primary nitrate reduction, dinitrogen fixation, or nitrogen remobilization has to be efficiently assimilated. In this chapter, we describe the main enzymes and genes responsible for ammonium assimilation in *Lotus japonicus* plants. We summarize the nomenclature and codes available in Kazusa 2.5 for the main genes involved in the ammonium assimilatory process, as well as the levels of expression found by qRT-PCR for these genes in different tissues of the plant.

11.1 Introduction

In the model legume *Lotus japonicus,* different forms of inorganic nitrogen (NO_3^- or NH_4^+) can be taken up by the plants, depending on nitrogen availability (Márquez et al. 2005; Orea et al. 2005). The utilization of NO_3^- requires its reduction to NH_4^+ produced by the consecutive action of nitrate reductase and nitrite reductase enzymes, prior to ammonium assimilation. Previous works have characterized the nitrate and nitrite reduction systems in *L. japonicus* (Harrison et al. 2004; Márquez et al. 2005; Orea et al. 2001; Pajuelo et al. 2002; Prosser et al. 2006). On the other hand, *L. japonicus*, as a legume plant, also has the chance to establish symbiosis with *Mesorhizobium loti* bacteria in order to use atmospheric N_2, which is reduced to NH_4^+ in the nodules by the action of bacterial nitrogenase (see other chapters of this book). Consequently, the process of primary ammonium assimilation, either derived from nitrate reduction or dinitrogen fixation, is of crucial importance in *L. japonicus* plants. This is also the case for other processes that produce an internal release of ammonium in *L. japonicus* plants, such as photorespiration, phenylpropanoid biosynthesis, or amino acid catabolism, called in general terms secondary ammonium assimilation (Betti et al. 2012; Márquez et al. 2005).

In the following sections, we will describe first the main enzymes in charge of ammonium assimilation, basically associated with glutamine/glutamate and asparagine metabolisms (Sect. 11.2). This is followed by the description (Sect. 11.3) of the different genes from *L.*

C.M. Pérez-Delgado · M. García-Calderón · A. Credali · J.M. Vega · M. Betti · A.J. Márquez (✉)
Departamento de Bioquímica Vegetal y Biología Molecular, Facultad de Química, Universidad de Sevilla, Calle Profesor García González, 1, 41012 Sevilla, Spain
e-mail: cabeza@us.es

S. Tabata and J. Stougaard (eds.), *The Lotus japonicus Genome*, Compendium of Plant Genomes,
DOI 10.1007/978-3-662-44270-8_11,

japonicus, which encode for the aforementioned enzymes, their nomenclature, and the codes available for them in Kazusa 2.5 database as well as the corresponding ortholog genes in *A. thaliana*. We also comparatively analyze the levels of expression determined by qRT-PCR that we have obtained for all of these genes in leaves, roots, and nodules of *L. japonicus* plants growing under N-sufficient conditions. All these results are also summarized in Table 11.1.

11.2 Enzymes Involved in Ammonium Assimilation

11.2.1 The GS-GOGAT Cycle

Glutamine is the first organonitrogen compound that is synthesized in the plants as a result of both primary and secondary ammonium assimilation, by means of the glutamine synthetase (GS)/glutamate synthase (GOGAT) cycle. Glutamine

Table 11.1 Genes for ammonium assimilation in *L. japonicus*

Gene name	Code in Kazusa 2.5	Levels of expression			Ortholog genes in *A. thaliana*
		Leaves	Roots	Nodules	
LjGln1.1	*chr2.CM0312.1480.r2.m*	H	H	H	*At5g37600* (*AtGln1.1*) *At1g66200* (*AtGln1.2*) *At3g17820* (*AtGln1.3*) *At5g16570* (*AtGln1.4*) *At1g48470* (*AtGln1.5*)
LjGln1.2	*chr6.CM0014.300.r2.m*	VH	VH	VH	
LjGln1.3	*LjSGA_030247.1*	L	VL	VL	
LjGln1.4	*LjSGA_058827.1*	L	ND	ND	
LjGln1.5	*LjSGA_019428.1*	VL	ND	ND	
LjGln2	*chr6.CM0139.890.r2.m*	VH	VH	VH	*At5g35630* (*AtGln2*)
LjGlu1	*chr1.CM0009.170.r2.d*	VH	H	H	*At5g04140* (*AtGlu1*) *At2g41220* (*At Glu2*)
LjGlt1	*LjSGA_035611.1*	VL	L	L	*At5g53460* (*AtGlt1*)
LjGlt2	*LjSGA_037992.1*	H	H	H	
LjGdh1	*chr1.CM0104.2530.r2.m*	H	VH	L	*At5g18170* (*AtGdh1*) *At5g07440* (*AtGdh2*) *At3g03910* (*AtGdh3*)
LjGdh2	*chr4.CM2142.210.r2.a*	VL	L	VL	
LjGdh3	*chr2.CM0021.1320.r2.m*	ND	VL	ND	
LjGdh4	*chr3.CM1488.210.r2.d* *chr3.CM1488.260.r2.d* *chr3.CM1488.250.r2.d* *chr3.CM1488.230.r2.d*	H	H	H	*At1g51720* (*AtGdh4*)
LjAsn1	*chr5.CM0071.330.r2.d*	VH	VH	H	*At3g47340* (*AtAsn1*) *At5g65010* (*AtAsn2*) *At5g10240* (*AtAsn3*)
LjAsn2	*LjT47C13.80.r2.d*	VL	VH	VH	
LjAsn3	*LjT09J04.190.r2.d*	L	L	L	
LjNse1	*chr5.CM0096.20.r2.m*	VH	H	L	*At3g16150* (*AtAspGB1*)
LjNse2	*chr4.CM0087.1740.r2.m*	L	H	H	*At5g08100* (*AtAspGA1*)
LjNse3	*LjSGA_021574.1*	H	L	VL	*At3g16150* (*AtAspGB1*)

Level of expression (relative units)	Level of expression	Symbol
0–0.001	Undetectable	ND
0.001–0.01	Very low	VL
0.01–0.1	Low	L
0.1–1	High	H
1–10	Very high	VH

synthetase (EC 6.3.1.2) catalyzes the biosynthesis of L-glutamine from L-glutamate, ATP, and NH_4^+ according to the following reaction, which also requires Mg^{2+} cations as cofactors:

$$\text{L}-\text{glutamate} + \text{ATP} + \text{NH}_4{}^+ \overset{\text{GS}}{\rightarrow} \text{L}-\text{glutamine} + \text{ADP} + P_i$$

Subsequently, the glutamate synthase enzyme (EC 1.4.7.1 or EC 1.4.1.14) catalyzes the transfer of the amide group of glutamine into 2-oxoglutarate, yielding two molecules of glutamate, a reaction also requiring two electrons coming from either reduced ferredoxin (Fd-GOGAT) or pyridine nucleotides (NADH-GOGAT), as follows:

$$\text{L}-\text{glutamine} + 2-\text{oxoglutarate} + 2e^- + 2\text{H}^+ \overset{\text{GOGAT}}{\rightarrow} 2\text{L}-\text{glutamate}$$

The global balance of the consecutive action of these two enzymes forms the GS-GOGAT cycle by which one of the two molecules of glutamate formed by the GOGAT would be used by the reaction of GS. Consequently, the GS-GOGAT pathway results in the net formation of one molecule of L-glutamate at the expense of one molecule of 2-oxoglutarate, one molecule of NH_4^+ and one molecule of ATP as follows:

$$2-\text{oxoglutarate} + \text{NH}_4^+ + 2e^- + 2\text{H}^+ + \text{ATP} \overset{\text{GS}-\text{GOGAT}}{\rightarrow} \text{L}-\text{glutamate} + \text{ADP} + \text{P}_\text{i}$$

11.2.2 Glutamate Dehydrogenase

In addition to GS and GOGAT, which catalyze irreversible reactions, a third enzyme, glutamate dehydrogenase (GDH; EC 1.4.1.2/4), catalyzes a reversible amination/deamination reaction, which could lead to either the synthesis or the catabolism of glutamate, according to the following equation:

$$2-\text{oxoglutarate} + \text{NH}_4{}^+ + \text{NAD}(P)\text{H} + \text{H}^+ \overset{\text{GDH}}{\leftrightarrow} \text{L}-\text{glutamate} + \text{NAD}(P)^+$$

The role of GDH in glutamate catabolism is quite well established. However, the possible anabolic role of GDH for the assimilation of ammonium has been the subject of continuous controversy because most lines of evidence support the idea that glutamate biosynthesis takes place through the GS-GOGAT pathway, although a role for GDH under different plant stress situations has been also reported. It has been proposed that GDH has an important role in terms of metabolic signaling in relation to the partitioning of C and N assimilates, most likely that GDH contributes to the control of the homeostasis of leaf glutamate, a process of crucial importance (Fontaine et al. 2012).

11.2.3 Asparagine Metabolism: Asparagine Synthetase and Asparaginase

In most temperate legumes, it is proposed that asparagine, rather than glutamine, is the principal molecule used to transport reduced nitrogen within the plant, in contrast to many other plant species (Credali et al. 2013). This is the case for *L. japonicus* where it has been shown that asparagine can account for almost 90 % of the nitrogen transported from root to shoot (Waterhouse et al. 1996).

Asparagine synthetase (AS, EC 6.3.5.4) is the main enzyme in charge of asparagine biosynthesis in plants. This enzyme catalyzes the transfer of the amide group from glutamine to aspartate in an ATP-dependent reaction:

$$\text{L}-\text{glutamine} + \text{L}-\text{aspartate} + \text{ATP} \overset{\text{AS}}{\rightarrow} \text{L}-\text{glutamate} + \text{L}-\text{asparagine} + \text{AMP} + \text{PP}_i$$

It has also been proposed that the enzyme can use high concentrations of ammonia directly as substrate, but this is not clearly demonstrated (Lea et al. 2007).

Considering that asparagine is a nitrogen transport compound in *Lotus*, asparagine breakdown is also a process of crucial importance for

this plant (Credali et al. 2013). Asparaginase (NSE, EC 3.5.1.1) catalyzes the hydrolysis of asparagine to yield aspartate and ammonia as follows:

$$\text{Asparagine} + \text{H}_2\text{O} \xrightarrow{\text{NSE}} \text{Aspartate} + \text{NH}_4{}^+$$

The ammonia released by the asparaginase reaction has to be subsequently reassimilated by GS (Lea et al. 2007).

11.3 Genes for Ammonium Assimilation

11.3.1 Glutamine Synthetase (GS)

As in other plant species, a small multigene family is responsible for GS enzymatic activity in *L. japonicus*. Five gene sequences encoding for cytosolic GS (also called *GS1* or *Gln1*) and another one for plastidic GS (also called *GS2* or *Gln2*) were found in the available databases. Two of the cytosolic genes are expressed at high (*LjGln1.1*) or very high (*LjGln1.2*) levels in leaves, roots, and nodules from this plant (Table 11.1). The level of expression of the *LjGln1.2* gene is about threefold higher in nodules than in roots or leaves, while in the case of the *LjGln1.1* gene, the level of expression is lower in nodules and leaves than in roots (results not shown). The third gene (*LjGln1.3*) is expressed at low levels in leaves and at very low levels in roots and nodules. A fourth and a fifth cytosolic gene (*LjGln1.4* and *LjGln1.5*) are only expressed at low (*LjGln1.4*) or very low (*LjGln1.5*) levels in leaves. Cytosolic GS polypeptides are the most abundant in both roots and nodules of *L. japonicus* plants (García-Calderón et al. 2012; Orea et al. 2002). Previous studies in *L. japonicus* had indicated that a high level of GS activity in the root is negatively correlated with above-ground biomass (Limami et al. 1999). On the other hand, lowering GS activity in nodules results in an increase in fresh weight in nodules, roots, and shoots (Harrison et al. 2003). This change in biomass could be explained by more efficient ammonium assimilation in the nodules of transformed plants, as indicated by a large increase in amino acids (mostly asparagine) with a concomitant decrease in carbohydrate content (Harrison et al. 2000). Other results established that constitutive overexpression of GS1 in *Lotus* produce higher amino acid levels and soluble protein concentration, higher chlorophyll content, and a higher biomass accumulation in the transgenic plants (Ortega et al. 2004), while overexpression in shoots may accelerate plant development, leading to early senescence and premature flowering when plants are grown on an ammonium-rich medium (Vincent et al. 1997). In addition, it was also noted that overexpression of GS1 in reproductive organs critically affects their development and might be a reason for sterility of *L. japonicus* plants (Suárez et al. 2003). In other plant species, 2–5 functional *Gln1* genes have been generally reported. Different versions of these genes have been found to be associated with different roles in primary ammonium assimilation and other forms of N recycling in response to nitrogen availability in the external medium, plant nitrogen status, light/dark conditions, or abiotic/biotic stressors (Bernard and Habash 2009). Cytosolic GS assimilates ammonium from the three major types of nitrogen-fixing symbiotic associations involving plants and either *Rhizobium*, actinomycetes such as *Frankia*, or cyanobacteria. In some species, a specific cytosolic GS isoenzyme is induced in nitrogen-fixing root nodules, whereas in others, cytosolic GS that is already present is involved (Bernard and Habash 2009).

Regarding *Gln2*, a single gene has been detected in *L. japonicus*, which has very high levels of expression either in leaves, roots, or nodules (Table 11.1). Nevertheless, the levels of transcript detected in leaves were found to be about fivefold higher than those present in roots or nodules from this plant (results not shown). In most plant species, plastidic GS (GS2) is exclusively, or very predominantly, expressed in green tissues. However, the presence of GS2 was clearly demonstrated in non-photosynthetic tissues of temperate legumes (García-Calderón et al. 2012). In *Medicago truncatula*, a second *Gln2* gene was recently shown to be exclusively expressed in developing seeds (Seabra et al. 2010). The first

gln2 mutants available from legume plants were isolated from *L. japonicus* belonging to the class of photorespiratory mutants previously described in other plant species (Márquez et al. 2005; Orea et al. 2002). The *Ljgln2* mutants were further characterized at the molecular level and shown to be affected by single point mutations within the structural part of the *LjGln2* gene, which leads to amino acid replacements that abolish GS2 enzymatic activity completely (Betti et al. 2006). These mutants were used to analyze how photorespiratory metabolism affects nodule function in *L. japonicus* plants (García-Calderón et al. 2012). The results obtained indicated that in this plant, and, particularly, in photorespiration, GS2 deficiency results in profound limitations in carbon metabolism that affect the nodulation process and nitrogen fixation. An anticipated senescence phenotype linked to an important reduction in starch and sucrose levels was observed (García-Calderón et al. 2012). In a separate study, a single *LjGln2* locus encoding for GS2 was mapped together with other symbiotic loci (Sandal et al. 2006). More recently, GS2 was also involved in drought stress, nitrogen nutrition, and photorespiratory metabolism transcriptomic responses in *L. japonicus* plants (Betti et al. 2012; Díaz et al. 2010; Pérez-Delgado et al. 2013).

11.3.2 Glutamate Synthase (GOGAT)

One *LjGlu1* gene encoding for Fd-GOGAT and two different genes encoding for NADH-GOGAT (*LjGlt1* and *LjGlt2*) were identified in the Kazusa database. Expression levels of *LjGlu1* were very high in photosynthetic tissues but were also high in roots and nodules, and similarly, *LjGlt2* expression levels were also high in the three types of tissues. However, this was not the case for the *LjGlt1* gene that was poorly expressed in roots, nodules, and leaves (Table 11.1). Early studies indicated that NADH-GOGAT appeared to play a major role in legume root nodules, in which the activity increases dramatically following the onset of nitrogen fixation. Two different isoforms of NADH-GOGAT have been described in other plant species, one of them being clearly associated with effective nodules (Ireland and Lea 1999, and references therein). Measurements of mRNA levels and promoter-GUS fusions of the NADH-GOGAT genes in alfalfa and *Lotus* have shown the tight relationship of the regulated expression of NADH-GOGAT to the nodulation process in legumes (Vance et al. 1995).

11.3.3 Glutamate Dehydrogenase (GDH)

Three different genes encoding for the NAD^+-dependent GDH were identified in *L. japonicus* (Table 11.1). One of them (*LjGdh1*) showed high or very high levels of expression in leaves and roots, respectively, and low level in nodules. The other two genes (*LjGdh2* and *LjGdh3*) were poorly expressed in these tissues. The majority of recent studies performed on NAD^+-GDH in higher plants have been focused on deciphering the role of the α and β subunits in the formation of seven isoenzymes, which are encoded by two distinct nuclear genes, *Gdh2* and *Gdh1*, respectively. More recently, it has been found that in *Arabidopsis,* there is a third gene (*AtGdh3*) encoding a putative NAD^+-GDH that is actively transcribed and perhaps regulated by cytokinin. Similarly, in rice and soybean, three genes encoding NAD^+-GDH were reported, although it seems that the physiological functions of the GDH isoenzymes is a complex issue and may vary from one species to another (Fontaine et al. 2012, and references therein).

A $NADP^+$-dependent form of GDH also exists, which appears to be localized in the chloroplast (in contrast to NAD^+-GDH that appears to be localized in mitochondria). However, the role of $NADP^+$-GDH is not clear. Consequently, a fourth expressed gene (*AtGdh4*) encoding a putative $NADP^+$-GDH has been identified in *Arabidopsis* and rice, which is 50 % longer than the NAD^+-GDH (Fontaine et al. 2012, and references therein). This is also the case in *L. japonicus*, where a *LjGdh4* gene was also identified in the Kazusa database, and is associated with four

different code names (Table 11.1). There is a high level of expression of the *LjGdh4* gene in leaves, roots, and nodules from this plant.

11.3.4 Asparagine Synthetase (AS)

Three genes for asparagine synthetase (*LjAsn1, LjAsn2,* and *LjAsn3*) have been identified in *L. japonicus* (also called, respectively, *LjAS1, LjAS2, and LjAS3*). *LjAsn1* is highly or very highly expressed in mature leaves, roots, and nodules. *LjAsn2* is very highly expressed in roots and nodules and barely detectable in leaves. The third gene (*LjAsn3*) is poorly expressed in the three types of tissues (Table 11.1). The molecular cloning and characterization of AS from *L. japonicus* in relation to the dynamics of asparagine biosynthesis in N-sufficient conditions has been reported previously (Waterhouse et al. 1996). Three genes encoding AS have been identified in other plant species, such as *A. thaliana*, which appear to be regulated in different manners (Lea et al. 2007). Although there is considerable variation between plants in the exact mechanisms involved in the regulation of the expression of AS, there is an overall consensus: The expression of one gene (often that which is most highly expressed) is induced by a reduction in soluble carbohydrate supply and in some cases darkness, while a second gene is more widely expressed but may be stimulated by carbohydrate and light. An increased supply of reduced nitrogen, either as ammonium or amino acids, induces expression of AS genes (Lea et al. 2007). Analysis of the amino acid sequences of plant's ASs shows that the proteins contain glutamine, aspartate, and AMP-binding sites and are related to the *E. coli* asparagine synthetase ASB glutamine-dependent enzymes (Lea et al. 2007).

11.3.5 Asparaginase (NSE)

Three different *Nse* genes encoding for asparaginase have been identified in *L. japonicus*, which showed a different pattern of expression among leaves, roots, and nodules of the plants (Table 11.1). The *LjNse1* gene is by far the most highly expressed one, particularly in leaves. Two of the genes present in *L. japonicus*, named *LjNse1* and *LjNse3*, encode for different K^+-dependent versions of the asparaginase enzyme, while a third one, named *LjNse2,* corresponds to a K^+-independent isoform. All these enzymes have a $\alpha_2\beta_2$ tetrameric quaternary structure, where the α and β subunits correspond, respectively, to the N-terminal (20 kDa) or C-terminal (17 kDa) domains arising from a single proteolytic event of each precursor, which is encoded by the different *Nse* genes. Structural and kinetic studies revealed the crucial importance of K^+ for the higher enzymatic activity and stability as well as lower Km for asparagine and proper orientation of asparagine substrate within the LjNSE1 enzyme molecule (Credali et al. 2011). It was proposed that LjNSE1 must be the main enzyme responsible for the utilization of asparagine in *L. japonicus* plants, while the K^+-independent isoform LjNSE2 is probably a detoxifying enzyme in charge of the release of isoaspartyl peptides arising from proteolytic degradation of post-translationally altered proteins. TILLING mutants affected in *LjNSE1* asparaginase isoform were recently used to demonstrate by reverse genetics, the importance of this particular isoform in plant growth and seed production. In fact, the level of both legumin and convicilin seed storage proteins was affected in the mutants (Credali et al. 2013). Nevertheless, *nse1* mutants indicated that there was no apparent involvement of NSE1 protein in nodulation. Interestingly, these results illustrate a key difference between *L. japonicus* and *Arabidopsis* where asparaginase activity seems to be dispensable; in fact, insertional mutants from *Arabidopsis* lacking of one or both K^+-dependent (*AspGB1*) and K^+-independent (*AspGA1*) asparaginases develop normally (Ivanov et al. 2011). This differential behavior among *Lotus* and *Arabidopsis* regarding asparaginases was attributed to the fact that asparagine only accounts for approximately 5 % of the total amino acids in the phloem sap of Brassicaceae, whereas it accounts for almost

90 % of nitrogen translocated in *L. japonicus* (Waterhouse et al. 1996), thus highlighting the importance of asparagine for nitrogen remobilization in *L. japonicus* plants (Credali et al. 2013).

Acknowledgments Authors acknowledge the funding by Consejería de Economía, Innovación y Ciencia from Junta de Andalucía (project P10-CVI-6368 and BIO-163). CMP acknowledges the receipt of PIF and V Plan Propio fellowships from University of Seville.

References

Bernard SM, Habash DZ (2009) The importance of cytosolic glutamine synthetase in nitrogen assimilation and recycling. New Phytol 182:608–620

Betti M, Arcondeguy T, Marquez AJ (2006) Molecular analysis of two mutants from *Lotus japonicus* deficient in plastidic glutamine synthetase: functional properties of purified GLN2 enzymes. Planta 224:1068–1079

Betti M, García-Calderón M, Pérez-Delgado CM et al (2012) Glutamine synthetase in legumes: recent advances in enzyme structure and functional genomics. Int J Mol Sci 13:7994–8024

Credali A, Díaz-Quintana A, García-Calderón M et al (2011) Structural analysis of K^+ -dependence in L-asparaginases from *Lotus japonicus*. Planta 234:109–122

Credali A, García-Calderón M, Dam S et al (2013) The K^+ -dependent asparaginase, NSE1, is crucial for plant growth and seed production in *Lotus japonicus*. Plant Cell Physiol 54:107–118

Díaz P, Betti M, Sánchez D et al (2010) Deficiency in plastidic glutamine synthetase alters proline metabolism and transcriptomic response in *Lotus japonicus* under drought stress. New Phytol 188:1001–1013

Fontaine J-X, Tercé-Laforgue T, Armengaud P et al (2012) Characterization of a NADH-dependent glutamate dehydrogenase mutant of *Arabidopsis* demonstrates the key role of this enzyme in root carbon and nitrogen metabolism. Plant Cell 24:4044–4065

García-Calderón M, Chiurazzi M, Espuny MR et al (2012) Photorespiratory metabolism and nodule function: behavior of *Lotus japonicus* mutants deficient in plastid glutamine synthetase. Mol Plant-Microbe Interact 25:211–219

Harrison J, Brugière N, Phillipson B et al (2000) Manipulating the pathway of ammonia assimilation through genetic engineering and breeding: consequences to plant physiology and plant development. Plant Soil 221:81–93

Harrison J, Pou de Crescenzo M-A, Sené O et al (2003) Does lowering glutamine synthetase activity in nodules modify nitrogen metabolism and growth in *Lotus japonicus*? Plant Physiol 133:253–262

Harrison J, Hirel B, Limami A (2004) Variation in nitrate uptake and assimilation between two ecotypes of *Lotus japonicus* L and their recombinant inbred lines. Physiol Plant 120:124–131

Ivanov A, Kameka A, Pajak A et al (2011) Arabidopsis mutants lacking asparaginases develop normally but exhibit enhanced root inhibition by exogenous asparagine. Amino Acids 42:2307–2318

Lea PJ, Sodek L, Parry MAJ et al (2007) Asparagine in plants. Ann Appl Biol 150:1–26

Limami A, Phillipson B, Ameziane R et al (1999) Does root glutamine synthetase control plant biomass in *Lotus japonicus* L.? Planta 209:495–502

Márquez AJ, Betti M, García-Calderón M et al (2005) Nitrate assimilation in *Lotus japonicus*. J Exp Bot 56:1741–1749

Orea A, Pajuelo P, Pajuelo E et al (2001) Characterisation and expression studies of a root cDNA encoding for ferredoxin-nitrite reductase from *Lotus japonicus*. Physiol Plant 113:193–202

Orea A, Pajuelo P, Pajuelo E et al (2002) Isolation of photorespiratory mutants from *Lotus japonicus* deficient in glutamine synthetase. Physiol Plant 115: 352–361

Orea A, Pajuelo P, Romero JM et al (2005) Nitrate assimilation: influence of nitrogen supply. In: Márquez AJ (ed) Lotus japonicus handbook. Springer, Dordrecht, pp 295–313

Ortega JL, Temple SJ, Bagga S et al (2004) Biochemical and molecular characterization of transgenic *Lotus japonicus* plants constitutively over-expressing a cytosolic glutamine synthetase gene. Planta 219: 807–818

Pajuelo P, Pajuelo E, Orea A et al (2002) Influence of plant age and growth conditions on nitrate assimilation in roots of *Lotus japonicus* plants. Funct Plant Biol 29:485–494

Pérez-Delgado CM, García-Calderón M, Sánchez DH et al (2013) Transcriptomic and metabolic changes associated to photorespiratory ammonium accumulation in the model legume *Lotus japonicus*. Plant Physiol 162:1834–1848

Prosser IM, Massonneau A, Smyth AJ et al (2006) Nitrate assimilation in the forage legume *Lotus japonicus* L. Planta 223:821–834

Sandal N, Petersen TR, Murray J et al (2006) Genetics of symbiosis in *L. japonicus*: recombinant inbred lines, comparative genetic maps and map position on 35 symbiotic loci. Mol. Plant-Microbe Interact 19:80–91

Seabra AR, Vieira CP, Cullimore JV et al (2010) *Medicago truncatula* contains a second gene encoding a plastid located glutamine synthetase exclusively expressed in developing seeds. BMC Plant Biol 10:183

Suárez R, Márquez J, Shishkova S et al (2003) Overexpression of alfalfa cytosolic glutamine synthetase in nodules and flowers of transgenic *Lotus japonicus* plants. Physiol Plant 117:326–336

Vance CP, Miller SS, Gregerson RG et al (1995) Alfalfa NADH-dependent glutamate synthase: structure of the gene and importance in symbiotic N_2 fixation. Plant J 8:345–348

Vincent R, Fraisier V, Chaillou S et al (1997) Overexpression of a soybean gene encoding glutamine synthetase in shoots of transgenic *Lotus corniculatus* L. plants triggers changes in assimilation and plant development. Planta 201:424–433

Waterhouse RM, Smyth AJ, Massoneau A et al (1996) Molecular cloning and characterisation of asparagine synthetase from *Lotus japonicus*: dynamics of asparagine synthesis in N-sufficient conditions. Plant Mol Biol 30:883–897

Nitrate Transport and Signaling

12

Vladimir Totev Valkov and Maurizio Chiurazzi

Abstract

As sessile organisms, plants have developed a sophisticated network of mechanisms to adapt and optimize their growth to the constantly and rapidly changing nutritional environmental conditions. The transport of nitrate in higher plants is a paradigm of this regulatory control as either external or internal cues can govern the root uptake ability depending on the nutrient demand and nitrogen availability in the soil. Plant adaptations also include the potential to respond to changes of nitrate concentration in the soil by modulating the root system developmental plan. It is known that in leguminous plants, nitrate availability in the soil can also strongly affect nodule formation as low and high concentrations exert a positive and negative effect on initiation of the organogenesis process, respectively. Nitrate can act both as a nutrient and a signal for the induction of plant root responses, and members of the nitrate and peptide transporters family (NPF and NRT2) play crucial roles in the control of such signaling pathways. This chapter presents an overview of the genomic and transcriptomic data reported for the *Lotus japonicus NPF* and *NRT2* family members, and their possible roles in the control of the nodulation program are discussed.

12.1 NPF and NRT2 Families

Nitrate transporter proteins are involved in the control of nitrate flux from soil to root tissues and allocation throughout the whole plant body (Miller et al. 2007). Plant roots have two different uptake systems to cope with low or high NO_3^- concentrations in soil, the high affinity and low affinity (HATS and LATS). Two types of NO_3^- transporters, known as the NRT1 and NRT2, contribute to LATS and HATS, respectively. The AtNRT1.1 and MtNRT1.3 proteins represent the only exceptions as they display a dual HATS/LATS nitrate uptake activity (Liu and Tsay 2003; Morère-Le Paven et al. 2011), although very recently the high-affinity action of AtNRT1.1 *in planta* was not confirmed as it appears to be confined to the experimental system constituted by Xenopus oocytes (Glass and Kotur 2013).

V.T. Valkov · M. Chiurazzi (✉)
Institute of Biosciences and Bioresources, Via P. Castellino 111, 80131 Naples, Italy
e-mail: maurizio.chiurazzi@ibbr.cnr.it

S. Tabata and J. Stougaard (eds.), *The Lotus japonicus Genome*, Compendium of Plant Genomes,
DOI 10.1007/978-3-662-44270-8_12,

Low-affinity nitrate transporter members share sequence similarity with peptide transporters (PTR) forming a superfamily that includes 53 and 80 members in *A. thaliana* and rice, respectively. Nitrate and/or di/tripeptides transport activity has been demonstrated for eighteen Arabidopsis members, and in most of the cases, these proteins were reported to act as proton-coupled transporters (Table 12.1) (Liu et al. 1999; Chiang et al. 2004). However, a biochemical approach to determine the transported substrate specificity is crucial for members of this superfamily, as this cannot be argued by sequence data alone. Several reports indicate they encompass proteins capable of transporting different substrates such as nitrate, di/tripeptides, amino acids, glucosinolates, malate, auxin, and ABA (Frommer et al. 1994; Liu et al. 1999; Jeong et al. 2004; Waterworth and Bray 2006; Krouk et al. 2010; Kanno et al. 2012; Nour-Eldin et al. 2012). In a few cases, a dual transport capacity for different substrates was also found on the same member (Table 12.1). For this reason, a new unified nomenclature based exclusively on phylogenetic relationship has been very recently established (Leran et al. 2014). Proteins of the nitrate/peptide transporters family (NPF) from 33 fully sequenced plant genomes were analyzed and eight unambiguous clades (subfamilies) identified. Therefore, NPF members are now identified by a two number code, indicating subfamily and relative position within this, respectively (Leran et al. 2014).

NRT2 proteins form small families of plant transporters in plants, including seven and four members in Arabidopsis and rice, respectively. In contrast to NPF genes, NRT2 members can be identified by sequence data and all the NRT2 proteins identified in higher plants transport nitrate. The NRT2 proteins are not functional alone, as an additional component, called NAR2/NRT3, is required for their nitrate transport activities in plants (Tsay et al. 2007). It is believed that the NRT2s are also proton-coupled transporters and four out of the seven NRT2 genes found in Arabidopsis show a nitrate-related phenotype when mutated (Lin et al. 2007; Chopin et al. 2007; Kiba et al. 2012). Recently, the AtNRT2.1 and AtNRT2.6 proteins have also been involved in the plant response to bacterial pathogen infection, as *nrt2.1* and *nrt2.6* mutants show a reduced susceptibility to *Pseudomonas syringae* pv tomato and *Erwinia amylovora* bacteria (Camanes et al. 2012; Dechorgnat et al. 2012).

12.2 Nitrate Effects on Root Architecture and Nodulation Programs

Nitrate may trigger the signaling pathways controlling lateral root development either systemically or locally (Zhang and Forde 1998). In Arabidopsis, increasing evidence indicates a role for NPF and NRT2 proteins in signaling transduction pathways. AtNPF6.3 (old name AtNRT1.1) plays a dual nutrient transporter/sensor role (transceptor) in the perception of external nitrate concentrations (Ho et al. 2009; Gojon et al. 2011). A switch between HATS and LATS affinity in AtNPF6.3 is controlled by phosphorylation at the T101 residue (Liu and Tsay 2003). Moreover, an activity of AtNPF6.3 as an auxin transport facilitator, allowing the definition of a functional linkage between nitrate and auxin signaling, controlling secondary root elongation, was reported (Krouk et al. 2010). The high-affinity complex NRT2.1–NAR2.1 also participates in regulating lateral root development, independently of uptake functions (Little et al. 2005). In legumes, the *M. truncatula* MtNPF1.7/NIP-LATD, recently characterized as a high-affinity nitrate transporter (Bagchi et al. 2012), is involved in lateral root and nodule development and primary root meristem maintenance (Bright et al. 2005).

As for secondary root developmental control, the nitrate effect on nodulation is exerted through both local and systemic controls (Carroll and Gresshoff 1983; Day et al. 1989; Carroll and Mathews 1990; Fujikake et al. 2003; Omrane and Chiurazzi 2009; Jeudy et al. 2010). High nitrate concentration in the growth medium (>1 mM) inhibits nodule formation, and old observations indicate that this effect does not

Table 12.1 Subclassification of the thirty-nine *L. japonicus* NPF proteins in the eight subfamilies identified by Leran et al. (2014)

Clade	*L. japonicus* Locus id.	New name	*A. thaliana/old name*	*M. truncatula/* old name	*O. sativa*, *A. glutinosa, B. napus*
1	chr3. LjT0H20.20	LjNPF1.1	AtNPF1.1	MtNPF1.1	
	chr4. CM0617.810	LjNPF1.2	AtNPF1.2/AtNRT1.11 (nitrate)	MtNPF1.2	
			AtNPF1.3	MtNPF1.3	
				MtNPF1.4	
				MtNPF1.5	
				MtNPF1.6	
				MtNPF1.7/NIP-LATD (nitrate)	
				MtNPF1.8	
2	chr4. CM0170.180	**LjNPF2.1**	AtNPF2.1	MtNPF2.1	
	chr2. CM0608.1290	LjNPF2.2	AtNPF2.2	MtNPF2.2	
	chr4. CM0170.40	LjNPF2.3	AtNPF2.3	MtNPF2.3	
	chr4. CM0170.210	LjNPF2.4	AtNPF2.4	MtNPF2.4	
	chr1. CM0147.130	LjNPF2.5	AtNPF2.5	MtNPF2.5	
			AtNPF2.6	MtNPF2.6	
			AtNPF2.7/AtNAXT1 (nitrate)	MtNPF2.7	
			AtNPF2.8	MtNPF2.8	
			AtNPF2.9/AtNRT1.9 (nitrate; glucosinolate)	MtNPF2.9	
			AtNPF2.10/AtGTR1 (glucosinolate)	MtNPF2.10	
			AtNPF2.11/AtGTR2/NRT1.10 (nitrate; glucos)	MtNPF2.11	
			AtNPF2.12/AtNRT1.6 (nitrate)	MtNPF2.12	
			AtNPF2.1/AtNRT1.7 (nitrate; glucosinolate)		
			AtNPF2.14		
3	chr2. CM0903.350	**LjNPF3.1**	AtNPF3.1/AtNitr (nitrate)	MtNPF3.1	
	chr1. CM1911.210	LjNPF3.2		NPF3.2	
	chr1. CM1911.220	LjNPF3.3		NPF3.3	

(continued)

Table 12.1 (continued)

Clade	*L. japonicus* Locus id.	New name	*A. thaliana/old name*	*M. truncatula/* old name	*O. sativa*, *A. glutinosa*, *B. napus*
4	chr2. CM0608.1210	LjNPF4.1	AtNPF4.1/AtAIT3 (ABA)	MtNPF4.1	
	chr4. CM0170.290	LjNPF4.2	AtNPF4.2/AtAIT4	MtNPF4.2	
	chr4. CM0046.1690	LjNPF4.3	AtNPF4.3/AtNRT1.14	MtNPF4.3	
	chr6. CM0118.580	LjNPF4.4	AtNPF4.4/AtNRT1.13	MtNPF4.4	
	chr1. CM0017.480	LjNPF4.5	AtNPF4.5/AtAIT2	MtNPF4.5	
			AtNPF4.6/AtNRT1.2/AIT1 (nitrate; ABA)	MtNPF4.6	
			AtNPF4.7	MtNPF4.7	
				MtNPF4.8	
				MtNPF4.9	
				MtNPF4.10	
				MtNPF4.11	
				MtNPF4.12	
				MtNPF4.13	
				MtNPF4.14	
5	chr6. CM1625.50	LjNPF5.1	AtNPF5.1	MtNPF5.1	
	chr1. CM0295.1000	LjNPF.2	AtNPF5.2/AtPTR3 (di/tripeptide)	MtNPF5.2	
	chr1. CM0295.980	**LjNPF5.3**	AtNPF5.3	MtNPF5.3	
	chr1. CM0295.970	**LjNPF5.4**	AtNPF5.4	MtNPF5.4	
	chr2. CM0081.1270	LjNPF5.5	AtNPF5.5	MtNPF5.5	
	chr1. CM0125.390	**LjNPF5.6**	AtNPF5.6	MtNPF5.6	
			AtNPF5.7	MtNPF5.7	
			AtNPF5.8	MtNPF5.8	
			AtNPF5.9	MtNPF5.9	
			AtNPF5.10	MtNPF5.10	
			AtNPF5.11	MtNPF5.11	
			AtNPF5.12	MtNPF5.12	
			AtNPF5.13/AtNRT1.16 (nitrate)	MtNPF5.13	
			AtNPF5.14/AtNRT1.15	MtNPF5.14	
			AtNPF5.15	MtNPF5.15	
			AtNPF5.16	MtNPF5.16	
				MtNPF5.17	

(continued)

Table 12.1 (continued)

Clade	*L. japonicus* Locus id.	New name	*A. thaliana/old name*	*M. truncatula/* old name	*O. sativa*, *A. glutinosa, B. napus*
				MtNPF5.18	
				MtNPF5.19	
				MtNPF5.20	
				MtNPF5.21	
				MtNPF5.22	
				MtNPF5.23	
				MtNPF5.24	
				MtNPF5.25	
6	chr1. CM0017.480	LjNPF6.1	AtNPF6.1	MtNPF6.1	AgDCAT1 (malate)
	chr2. CM0826.350	**LjNPF6.2**	AtNPF6.2/AtNRT1.4 (nitrate)	MtNPF6.2	BnNRT1.2 (histid; nitrate)
	chr2. CM0021.3040	LjNPF6.3	AtNPF6.3/AtNRT1.1 (nitrate; auxine)	MtNPF6.3	
	chr4. LjB20H09.30	LjNPF6.4	AtNPF6.4/AtNRT1.3 (nitrate)	MtNPF6.4	
	chr2. CM0826.370	**LjNPF6.5**		MtNPF6.5	
	chr2. CM0545.330	LjNPF6.6		MtNPF6.6	
	chr2. CM0021.2180	LjNPF6.7		MtNPF6.7	
	chr2. CM0021.2200	LjNPF6.8		MtNPF6.8/MtNRT1.3 (nitrate)	
7	chr1. CM0017.770	LjNPF7.1	AtNPF7.1	MtNPF7.1	
	chr4. CM0247.130	LjNPF7.2	AtNPF7.2/AtNRT1.8 (nitrate)	MtNPF7.2	
	chr1. CM0141.10	LjNPF7.3	AtNPF7.3/AtNRT1.5 (nitrate)	MtNPF7.3	OsNPF7.3 (di/ tripeptide)
				MtNPF7.4	
				MtNPF7.5	
				MtNPF7.6	
				MtNPF7.7	
				MtNPF7.8	
				MtNPF7.9	
8	Lj24M05.60	LjNPF8.1	AtNPF8.1/AtPTR1 (di/tripep; histidine)	MtNPF8.1	
	chr4. CM0026.890	LjNPF8.2	AtNPF8.2/AtPTR5 (di/tripeptide)		
	chr4. CM0026.930	LjNPF8.3	AtNPF8.3/AtPTR2 (di/tripeptide; histidine)		

(continued)

Table 12.1 (continued)

Clade	*L. japonicus* Locus id.	New name	*A. thaliana/old name*	*M. truncatula/* old name	*O. sativa*, *A. glutinosa*, *B. napus*
	chr4. CM0026.860	LjNPF8.4	AtNPF8.4/AtPTR4		
	chr4. CM0026.870	LjNPF8.5	AtNPF8.5/AtPTR6		
	chr2. LjT15I01.230	**LjNPF8.6**			
	chr4. CM0026.880	LjNPF8.7			
					OsNPF8.9 (nitrate)

The *A. thaliana* and *M. truncatula* NPF members are also included for comparison, as well as the functionally characterized *O. sativa* NPF7.3 and NPF8.9 (Lin et al. 2000), *Brassica napus* NRT1.2 (Zhou et al. 1998), and *Alnus glutinosa* DCAT1 (Jeong et al. 2004) members. Clade numbers indicate the different subfamilies. When known, the transported substrates are indicated in brackets. In bold are the *L. japonicus* NPF members with a nodule-induced profile of expression. At; *Arabidopsis thaliana.* Mt *Medicago truncatula.* Os *Oriza sativa.* Bn *Brassica napus.* Ag *Alnus glutinosa*

depend on its assimilation by nitrate reductase, and hence, it is unlikely to be due to a nutritional effect (Carroll and Mathews 1990). Two clear-cut early phenotypes consisting of the inhibition of cortical cell division and down-regulation of the NIN gene expression were reported in *Lotus japonicus* wild-type plants inoculated with *M. loti* in the presence of high KNO_3 concentrations (Barbulova et al. 2007). However, the mechanisms and factors involved in the signaling pathways leading to the nitrate-dependent nodule organogenesis inhibition are still largely unknown. A possible role played by nitrate in the control of auxin, ethylene, and/or flavonoid signaling pathways has been postulated (Cho and Harper 1991; Caba et al. 2000). Recently, a role for nitrate-responsive CLE peptides as the main actors in the transduction of the root signal to the Leucine-rich Repeat Receptor Kinases (HAR1/NARK1/NST1/SUNN)-dependent mechanism governing autoregulation of nodule numbers has been reported in *L. japonicus*, Soybean, and *Medicago truncatula* and even in these cases both local and/or systemic mechanisms of control were proposed (Okamoto et al. 2009; Mortier et al. 2010; Reid et al. 2011).

12.3 *Lotus Japonicus* NPF and NRT2 Families: Genomic Organization and Expression Profiles

The retrieval of NPF sequences from the *L. japonicus* whole-genome sequence resource (Sato et al. 2008; http://www.kazusa.or.jp/lotus/) leading to the identification of 37 putative members has been recently published (Criscuolo et al. 2012). Sequence analyses predicted the conserved structural arrangement of 12 transmembrane domains connected by short peptide loops for almost the totality of the *L. japonicus* NPF members (Table 12.2). A further search led to the identification of two additional complete NPF sequences and 32 un-completed unique sequences of predicted *NPF* genes, indicating a size of around 70 members for the *L. japonicus* family. Fifty-one genes are physically mapped on the Lotus genome indicating a distribution on all the six chromosomes. Most are found on chromosomes 1, 2, and 4 with sixteen, fifteen, and fifteen genes, respectively, whereas one gene is located on chromosome 3 and two on chromosomes 5

and 6. At least seven gene clusters were identified that co-localize to the same contigs with paralogous genes and short intergenic regions (e.g., 2,032 bp between *LjNPF5.3* and *LjNPF5.4* genes) (Criscuolo et al. 2012).

The assignment of the 39 *L. japonicus* complete members to the eight clades of the NPF superfamily identified by Leran et al. (2014) and their relative position within these was based on a BLAST analysis where each of the *L. japonicus* proteins were queried against the already assigned members of the *A. thaliana* and *M. truncatula* families. A nomenclature for the provisional list of complete *Lj*NPF members is indicated in Tables 12.1 and 12.2. *L. japonicus* NPF members are distributed between all the eight subfamilies with a peak of eight and a minimum of two members in clades six and one, respectively.

Transcription of *NPF1* and *NRT2* plant genes is reported to be regulated by nitrate, nitrite, ammonium, glutamine, N starvation, light, sucrose, diurnal rhythm, and/or pH (Wang et al. 2012), and in some cases, the transcriptional regulation is linked to a modulation of the nitrate uptake activities. In the molecular characterization reported by Criscuolo et al. (2012), a transcriptional analysis of a subset of *L. japonicus NPF* and *NRT2* genes showed a repressible, inducible, and constitutive response to provision of nitrate, auxin, or cytokinin. This analysis also allowed the identification of *L. japonicus NPF* and *NRT2* genes induced during the symbiotic interaction with *M. loti* with two genes, LjNPF6.5/chr2.CM0826.370 and chr1.CM0001.20, that were specifically expressed in nodular tissue (Criscuolo et al. 2012). The pattern of expression of the *L. japonicus NPF* and *NRT2* genes can be integrated by the large amount of data reported in genome-wide analysis comparing expression profiles in inoculated and un-inoculated plants (Colebatch et al. 2004; Kouchi et al. 2004; Høgslund et al. 2009). In particular, we queried the large set of *L. japonicus* transcriptome data encompassing different organs, stages of the symbiotic interaction, and root nodules development in wild-type and mutant genotypes (Høgslund et al. 2009). A large number of Lotus *NPF* and *NRT2* genes are identified by the probe sets exploited in this GeneChip approach, and the profiles of expression in roots and young and mature nodules are reported in Table 12.3 (Høgslund et al. 2009). Interestingly, seven members of the family that do not include the *LjNPF6.5* cited above show a clear-cut induction profile in nodular tissue. Moreover, the level of expression of these nodule-induced genes is not dependent on nitrogen fixation as it is not affected in the fix-nodules obtained in the *sen1* and *sst1* genetic backgrounds (Høgslund et al. 2009). These data have been further updated by the analysis reported by Takahashi et al. (2012) that indicate the infection zone as the nodule region where a peak of transcription level is detected for five out of the eight NPF nodule-induced genes (Table 12.3).

The subclassification of the *L. japonicus* NPF members shown in Table 12.1 indicates that the eight Lotus proteins preferentially expressed in nodular tissue are distributed in five out of the eight phylogenetic subfamilies identified by Leran et al. (2014) with a peak of three nodule-induced NPF genes in the clade 5. This distribution suggests different biochemical and physiological roles. Interestingly, the subfamily six with the LjNPF6.2 and LjNPF6.5 proteins also includes the *A. glutinosa* AgDCAT1 protein involved in the supply of intracellular bacteria with dycarboxylates (malate) in the actinorhizal/frankia symbiotic interaction (Table 12.1) (Jeong et al. 2004).

The molecular characterization of the *L. japonicus NRT2* family allowed the identification of a single gene (chr1.CM0001.20) that was strongly induced in young and mature nodular tissue (Criscuolo et al. 2012). In Table 12.3, the expression profiles of the *LjNRT2.1* and *LjNRT2.2* genes exported from the data reported by Høgslund et al. (2009) are also indicated, showing a significant down-regulation in young and mature nodules. This pattern is consistent with the one published for *LjNRT2.1* in Criscuolo et al. (2012), while for *LjNRT2.2,* only a slight decrease of the transcript level was reported.

Table 12.2 List of complete *LjNPF* and *LjNRT2* genes with indication of structural features of the predicted proteins

LjNPF old name/new name	aa length	Number of TM domains
chr3.LjT07H20.20/ LjNPF1.1	585	12
chr4.CM0617.810/ LjNPF1.2	576	12
chr4.CM0170.180/ LjNPF2.1	579	12
chr2.CM0608.1290/ LjNPF2.2	635	12
chr4.CM0170.40/ LjNPF2.3	601	12
chr4.CM0170.210/ LjNPF2.4	593	12
chr1.CM0147.130/ LjNPF2.5	590	12
chr2.CM0903.350/ LjNPF3.1	580	12
chr1.CM1911.210	616	11
chr1.CM1911.220	599	11
chr2.CM0608.1210/ LjNPF4.1	581	12
chr4.CM0170.290/ LjNPF4.2	557	12
chr4.CM046.1690/ LjNPF4.3	591	12
chr6.CM0118.580/ LjNPF4.4	605	12
chr1.CM0017.480/ LjNPF4.5	634	12
chr6.CM1625.50/ LjNPF5.1	587	11
chr1.CM0295.1000/ LjNPF5.2	592	12
chr1.CM0295.980/ LjNPF5.3	606	12
chr1.CM0295.970/ LjNPF5.4	608	12
chr2.CM0081.1270/ LjNPF5.5	539	12
chr1.CM0125.390/ LjNPF5.6	573	12
chr1.CM0017.480/ LjNPF6.1	634	12
chr2.CM0826.350/ LjNPF6.2	583	12
chr2.CM0021.3040/ LjNPF6.3	613	12
chr4.LjB20H09.30/ LjNPF6.4	593	12
chr2.CM0826.370/ LjNPF6.5	603	12
chr2.CM0545.330/ LjNPF6.6	581	12
chr2.CM0021.2180/ LjNPF6.7	598	12
chr2.CM0021.2200/ LjNPF6.8	582	12
chr1.CM0017.770/ LjNPF7.1	603	13
chr4.CM0247.130/ LjNPF7.2	594	12
chr1.CM0141.10/ LjNPF7.3	582	12
LjT24M05.60/ LjNPF8.1	570	11
chr4.CM0026.890/ LjNPF8.2	586	11
chr4.CM0026.930/ LjNPF8.3	586	12
chr4.CM0026.860/ LjNPF8.4	566	12
chr4.CM0026.870/ LjNPF8.5	591	12
chr2.LjT15I01.230/ LjNPF8.6	570	11
chr4.CM0026.880/ LjNPF8.7	579	12
LjNRT2		
chr1CM0001.20	459	12
chr3.CM0649.30/ LjNRT2.1	531	12
chr3.CM0649.40/ LjNRT2.2	531	12
chr4.CM0161.180	508	12

12.4 Potential Cross Talk Between NPF, NRT2 Transporters and SNF

The functional characterization of the NPF and NRT2 families in a legume plant offers the opportunity to evaluate the potential roles played by these transporters in (i) a specific root organogenesis pathway occurring in response to biotic and abiotic signals (e.g., *Rhizobium* and N conditions), (ii) a specific tissue context represented by nodules, (iii) a specific functional constrain devoted to the exchanges occurring between the symbiotic partners.

A potential role of the NPF and NRT2 proteins could be played in several early steps of nodule formation. One possibility would be a direct involvement in the control of nitrate uptake rate and/or an indirect role through transporting the endogenous CLE peptides. A regulatory action may also reflect the involvement in the local pathway that senses and transduces the external nitrate signal to the root machinery involved in nodule organogenesis. Furthermore, as indicated in Table 12.1, a direct auxin and abscisic acid uptake capacity has been reported in the cases of *At*NPF6.3, AtNPF4.1, and AtNPF4.6 (Krouk et al. 2010; Kanno et al. 2012), while a negative feedback loop between NRT2.1 expression and ethylene biosynthesis was discovered (Zheng et al. 2013). These three hormones have important and antagonistic actions on epidermal and cortical cells responses in the early steps of the nodule organogenesis program (Ding and Oldroyd 2009). A dual transport capacity for nitrate and hormones as in the cases of ATNPF4.6 and ATNPF6.3 (Table 12.1) would also be functional for NPF proteins involved in the regulation of nodule development. The high-affinity nitrate transporter MtNIP/LATD plays a role in nodule development, but its action appears to be independent by its nitrate transport function suggesting the involvement of a different biochemical activity (Salehin et al. 2013).

Furthermore, the development and functioning of nodules is based on a coordinated differentiation of plant and bacterial cells, which produce a mature organ with both infected and uninfected plant cells. Gene expression studies reveal that nodules have a distinct metabolic phenotype. A complex network of transport and exchanges takes place in nodules, which provides reduced carbon and other nutrients from the plant to bacteroids and fixed nitrogen from bacteroids to the plants. Many transporters must be involved in the dynamic of metabolite exchanges, but at the moment, the knowledge of their molecular basis is still limited. The increased amount of data indicating high transport plasticity for members of the NPF family, including the reported capacity of amino acids (Waterworth and Bray 2006) and dicarboxylic acids (Jeong et al. 2004), makes these proteins candidates for playing important roles in the control of this nutrient traffic.

The nitrate itself plays also an important role as regulator of legume nodules activity as it is known that a few days after nitrate exposure, nodule activity is almost completely lost and the nodules become senescent (Matamoros et al. 1999), but the mechanisms through which this action takes place is still controversial. The effect of nitrate is mediated by significant changes at the gene expression level occurring in nodules (Cabeza et al. 2014). The analysis of the global response of nodule transcriptome apparently suggests that nitrate targets the very heart of the N_2 reduction process, i.e., the formation of the nitrogenase complex itself and ATP generation (Cabeza et al. 2014). According to this, a nitrate reductase-dependent nitric oxide (NO) synthesis process, involved in the maintenance of the energy status required for N fixation under oxygen-limiting conditions, has been reported (Horchani et al. 2011). Therefore, the involvement of nitrate transporters into this complicate network of activities supporting the nodule functioning must be taken in consideration.

A crucial drive for the understanding of the NPF and NRT2 roles during the symbiotic interaction will certainly come from the exploitation of the recently released LORE1-tagged collection that includes 40.000 LORE1 *L. japonicus* lines comprising more than 120.000

Table 12.3 Log2 expression values of the *NPF* and *NRT2* genes are exported by the Web-accessible resource http://cgi-www.cs.au.dk/cgi-compbio/Niels/index.cgi (Høgslund et al. 2009). Nodule-induced genes are in bold

	Root	Nodule 14 dpi	Nodule 21 dpi
NPF genes			
chr3.LjT07H20.20/ LjNPF1.1	9.49	6.74	6.69
chr4.CM0617.810/ LjNPF1.2	2.12	2.48	2.11
chr4.CM0170.180/ LjNPF2.1	7.73	9.86	10.05
chr4.CM0170.210/ LjNPF2.4	2.29	2.29	2.28
chr1.CM0147.130/ LjNPF2.5	4	4	4
chr2.CM0903.350/ LjNPF3.1[a]	9	12.76	13.03
chr2. CM0608.1210/ LjNPF4.1	9.1	9.1	8.8
chr4.CM0170.290/ LjNPF4.2	3.27	2.98	3.19
chr4.CM046.1690/ LjNPF4.3	7.08	2.51	2.6
chr6.CM1625.50/ LjNPF5.1	2.29	2.3	2.28
chr1. CM0295.1000/ LjNPF5.2	2.55	2.46	2.45
chr1.CM0295.980/ LjNPF5.3[a]	2.6	13.2	13.53
chr1.CM0295.970/ LjNPF5.4[a]	2.21	13.7	13.86
chr2. CM0081.1270/ LjNPF5.5	9.67	9.67	9.5
chr1.CM0125.390/ LjNPF5.6[a]	2.4	14.3	14.3
chr2.CM0826.350/ LjNPF6.2	3.06	5.24	5.44
chr2. CM0021.3040/ LjNPF6.3	2.13	2.12	2.11
chr4.LjB20H09.30/ LjNPF6.4	7.5	7.54	7.14
chr2.CM0545.330/ LjNPF6.6	1.99	11.46	11.44
chr4.CM0247.130/ LjNPF7.2	6.22	4.8	6.19
chr1.CM0141.10/ LjNPF7.3	2.48	2.58	2.52
chr2. LjT15I01.230/ LjNPF8.6[a]	2.82	10.45	10.5
NRT2 genes			
chr3.CM0649.40/ LjNRT2.1	8.87	6.93	6.18
chr3.CM0649.30/ LjNRT2.2	12.48	10.79	9.86

[a] Indicates a consistent nodule-induced profile in Takahashi et al. (2012)

annotated insertion events (Urbanski et al. 2012; Fukai et al. 2012). So far, LORE1 element was identified in 15 out of the 38 annotated members of the NPF family and three out of the four NRT2 members and functional characterizations of correlated knock out mutants are in progress.

References

Bagchi R, Slaehin M, Adeyemo OS et al (2012) Functional assessment of the *Medicago truncatula* NIP/LATD protein demonstrates that is a high-affinity nitrate transporter. Plant Physiol 160:906–916

Barbulova A, Rogato A, D'apuzzo E et al (2007) Differential effects of combined N sources on early steps of the Nod factor-dependent transduction pathway in Lotus japonicus. Mol Plant Microbe Interact 20:994–1003

Bright LJ, Liang Y, Mitchell DM et al (2005) The LATD gene of *Medicago truncatula* is required for both nodule and root development. Mol Plant Microbe Inter 18:521–532

Caba JM, Centeno ML, Fernandez B et al (2000) Inoculation and nitrate alter phytohormone levels in soybean roots: differences between a supernodulating mutant and the wild type. Planta 211:98–104

Cabeza R, Koester B, Liese R et al (2014) An RNA sequencing transcriptome analysis reveals novel insights into molecular aspects of the nitrate impact on the nodule activity of *Medicago truncatula*. Plant Physiol 164:400–411

Camanes G, Pastor V, Cerezo M et al (2012) A deletion in NRT2.1 attenuates Pseudomonas syringae-induced hormonal perturbation, resulting in primed plant defenses. Plant Physiol 158:1054–1066

Carroll B, Gresshoff PM (1983) Nitrate inhibition of nodulation and nitrogen fixation in white clover. Z Pflanzenphysiol 110:69–76

Carroll B, Mathews A (1990) Nitrate inhibition of nodulation in legumes. In: Gresshoff PM (ed) Molecular biology of symbiotic nitrogen fixation, pp 159–180

Chiang CS, Stacey G, Tsay YF (2004) Mechanisms and functional properties of two peptide transporters, AtPTR2 and fPTR2. J Biochem Chem 279:30150–30157

Cho M, Harper JE (1991) Effect of localized nitrate application on isoflavonoid concentration and nodulation in split-root systems of wild-type and nodulation-mutant soybeans plants. Plant Physiol 95:1106–1112

Chopin F, Orsel M, Dorbe MF et al (2007) The *Arabidopsis* ATNRT2.7 nitrate transporter controls nitrate content in seeds. Plant Cell 19:1590–1602

Colebatch G, Desbrosses G, Ott T (2004) Global changes in transcription orchestrate metabolic differentiation during symbiotic nitrogen fixation in *Lotus japonicus*. Plant J 39:487–512

Criscuolo G, Valkov VT, Parlati A et al (2012) Molecular characterization of the *Lotus japonicus* NRT1(PTR) and NRT2 families. Plant Cell Environ 35:1567–1581

Day DA, Carroll BJ, Delves AC et al (1989) Relationship between autoregulation and nitrate inhibition of nodulation in soybeans. Physiol Plant 75:37–42

Dechorgnat J, Patrit O, Krapp A et al (2012) Characterization of the Nrt2.6 gene in Arabidopsis thaliana: a link with plant response to biotic and abiotic stress. PLoS ONE 7:e42491. doi:10.1371/journal.pone.0042491

Ding Y, Oldroyd GE (2009) Positioning the nodule, the hormone dictum. Plant Signal Behav 4:89–93

Frommer WB, Hummel S, Rentsch D (1994) Cloning of an *Arabidopsis* histidine transporting protein related to nitrate and peptide transporters. FEBS Lett 347:185–189

Fujikake H, Yamazaki A, Ohtake N et al (2003) Quick and reversible inhibition of soybean root nodule growth by nitrate involves a decrease in sucrose supply to nodules. J Exp Bot 54:1379–1388

Fukai E, Soyano T, Umehara Y et al (2012) Establishment of a *Lotus japonicus* gene tagging population using the exon-targeting endogenous retrotransposon LORE1. Plant J 69:720–730

Glass AD, Kotur Z (2013) A reevaluation of the role of Arabidopsis NRT1.1 in high-affinity nitrate transport. Plant Physiol 163:1103–1106

Gojon A, Krouk G, Perrine-Walker F et al (2011) Nitrate transceptor(s) in plants. J Exp Bot 62:2299–2308

Ho CH, Lin SH, Hu HC et al (2009) CHL1 functions as a nitrate sensor in plants. Cell 138:1184–1194

Høgslund N, Radutoiu S, Krusell L et al (2009) Organ development by integrated transcriptome analysis of *Lotus japonicus* mutant and wild-type plants. PloS One 7; 4(8):e6556. doi:10.1371/journal.pone.0006556

Horchani F, Prèvot M, Boscari A et al (2011) Both plant and bacterial nitrate reductase contribute to nitric oxide production in *Medicago truncatula* nitrogen-fixing nodules. Plant Physiol 155:1023–1036

Jeong J, Suh S, Guan C et al (2004) A nodule-specific dicarboxylate transporter from alder is a member of the peptide transporter family. Plant Physiol 134:969–978

Jeudy C, Ruffell S, Freixes S et al (2010) Adaptation of *Medicago truncatula* to nitrogen limitation is modulated via local and systemic nodule developmental responses. New Phytol 185:817–828

Kanno Y, Hanada A, Chiba Y et al (2012) Identification of an abscisic acid transporter by functional screening using the receptor complex as a sensor. Proc Natl Acad Sci 109:9653–9658

Kiba T, Feria-Bourrellier AB, Lafouge F et al (2012) The *Arabidopsis* nitrate transporter NRT2.4 plays a double role in roots and shoots of nitrogen-starved plants. Plant Cell 24:245–258

Kouchi H, Shimomura K, Hata S et al (2004) Large-scale analysis of gene expression profiles during early stages of root nodule formation in a model legume, *Lotus japonicus*. DNA Res 11:263–274

Krouk G, Lacombe B, Bielach A et al (2010) Nitrate-regulated auxin transport by NRT1.1 defines a mechanism for nutrient sensing in plants. Dev Cell 18:927–937

Leran S, Varala K, Boyer J-C et al (2014) A unified nomenclature of Nitrate Transporter 1/Peptide transporter family members in plants. Trends Plant Sci 19:5–9

Lin W, Wang Y, Okamoto M, Crawford NM et al (2007) Dissection of the AtNRT2.1:AtNRT2.2 inducible high-affinity nitrate transporter gene cluster. Plant Physiol 143:425–433

Lin CM, Koh S, Stacey G et al (2000) Cloning and functional characterization of a constitutively expressed nitrate transporter gene OsNRT1 from rice. Plant Physiol 122:379–388

Little D, Rao H, Oliva S, Daniel-Vedele F et al (2005) The putative high-affinity nitrate transporter NTR2.1 represses lateral root initiation in response to nutritional cues. Proc Natl Acad Sci 102:13693–13698

Liu KH, Huang CY, Tsay YF (1999) CHL1 is a dual-affinity nitrate transporter of *Arabidopsis* involved in multiple phases of nitrate uptake. Plant Cell 11:865–874

Liu KH, Tsay YF (2003) Switching between the two action modes of the dual-affinity nitrate transporter CHL1 by phosphorylation. EMBO J 22:1005–1013

Matamoros MA, Baird LM, Escuredo PR et al (1999) Stress-induced legume root nodule senescence. Physiological, biochemical, and structural alterations. Plant Physiol 121:97–112

Miller AJ, Fan X, Orsel M et al (2007) Nitrate transport and signalling. J Exp Bot 58:2297–2306

Morère-Le Paven MC, Viau L, Hamon A et al (2011) Characterization of a dual-affinity nitrate transporter MtNRT1.3 in the model legume *Medicago truncatula*. J Exp Bot 62:5595–5605

Mortier V, Den Herder G, Whitford R et al (2010) CLE peptides control *Medicago truncatula* nodulation locally and systemically. Plant Physiol 153:222–237

Nour-Eldin HH, Andersen TG, Burow M et al (2012) NRT/PTR transporters are essential for translocation of glucosinolate defence compounds to seeds. Nature 488:531–534

Okamoto S, Ohnishi E, Sato S et al (2009) Nod factor/nitrate-induced CLE genes that drive HAR1-mediated systemic regulation of nodulation. Plant Cell Physiol 50:67–77

Omrane S, Chiurazzi M (2009) A variety of regulatory mechanisms are involved in the nitrogen-dependent modulation of the nodule organogenesis program in legume roots. Plant Sig Behav 4:1066–1068

Reid DE, Ferguson BJ, Gresshoff PM (2011) Inoculation- and nitrate-induced CLE peptides of soybean control NARK-dependent nodule formation. Mol Plant Microbe Interact 24:608–618

Salehin M, Huang Y-S, Bagchi R et al (2013) Allelic differences in *Medicago truncatula* NIP/LATD mutants correlate with their encoded proteins transport activities in planta. Plant Signal Behav doi:10.4161/psb.22813

Sato S, Nakamura Y, Kanedo T et al (2008) Genome structure of the legume, *Lotus japonicus*. DNA Res 15:227–239

Takanashi K, Takahashi H, Sakurai N et al (2012) Tissue-specific transcriptome analysis in nodules of *Lotus japonicus*. Mol Plant Microbe Interact 25:869–876

Tsay YF, Chiu CC, Tsai CB et al (2007) Nitrate transporters and peptide transporters. FEBS Lett 581:2290–2300

Urbański DF, Małolepszy A, Stougaard J et al (2012) Genome-wide LORE1 retrotransposon mutagenesis and high-throughput insertion detection in *Lotus japonicus*. Plant J 69:731–741

Wang X, Bian Y, Cheng K et al (2012) A comprehensive differential proteomic study of nitrate deprivation in *Arabidopsis* reveals complex regulatory networks of plant nitrogen responses. J Proteome Res 11:2301–2315

Waterworth WM, Bray CM (2006) Enigma variations for peptides and their transporters in higher plants. Ann Bot 98:1–8

Zhang HM, Forde BG (1998) An *Arabidopsis* MADS box gene that controls nutrient-induced changes in root architecture. Science 279:407–409

Zheng D, Han X, An Y et al (2013) The nitrate transporter NRT2.1 functions in the ethylene response to nitrate deficiency in *Arabidopsis*. Plant Cell Environ 36:1328–1337

Zhou JJ, Theodoulou FL, Muldin I et al (1998) Cloning and functional characterization of a Brassica napus transporter that is able to transport nitrate and histidine. J Biol Chem 273:12017–12023

Reactive Oxygen/Nitrogen Species and Antioxidant Defenses in *Lotus japonicus*

13

Manuel Becana, Manuel A. Matamoros, Javier Ramos, Maria C. Rubio, and Martha Sainz

Abstract

Reactive oxygen/nitrogen species (ROS/RNS) are potentially cytotoxic molecules because they can generate oxidative/nitrosative stress. However, ROS and RNS, at concentrations tightly regulated by antioxidants, also serve useful purposes in processes such as organ development, abiotic and biotic stress response, and redox signaling. Antioxidant enzymes and metabolites are abundant in plants and particularly in legume nodules. Most of the enzymes involved in antioxidant defense are encoded by multigene families and occur as multiple isoforms in various cellular compartments, forming a dynamic network that is spatiotemporally regulated. Genomic, transcriptomic, and proteomic analyses of model legumes, such as *Lotus japonicus* and *Medicago truncatula*, are unveiling a complex regulation of antioxidant pathways in different tissues and especially during the symbiotic interaction with rhizobia. This regulation includes alternatively spliced forms of the genes and post-translational modifications of the proteins, which with no doubt will be the subject of intense research over the next years.

13.1 Reactive Oxygen and Nitrogen Species in Legumes

Plants, like other aerobic organisms, produce reactive oxygen species (ROS), mainly superoxide radicals ($O_2^{\cdot-}$) and hydrogen peroxide (H_2O_2), during mitochondrial respiration (Halliwell and Gutteridge 2007). Other major sources of ROS in plant cells are the chloroplasts and peroxisomes. In legume nodules, ROS are generated in mitochondria, plastids, peroxisomes, and bacteroids (Becana et al. 2010). Although $O_2^{\cdot-}$ and H_2O_2 have moderate reactivity, they can interact giving rise to ROS with highly oxidizing potential. This is the case of metal-catalyzed Fenton reactions, in which H_2O_2 is reduced by trace amounts of ferrous iron to hydroxyl radicals, which are then able to oxidize virtually all molecules at nearly diffusion rates (Halliwell and Gutteridge 2007). Also, alkoxyl and peroxyl radicals can be formed in processes such as the peroxidation of membrane lipids. The

M. Becana (✉) · M.A. Matamoros · J. Ramos · M.C. Rubio · M. Sainz
Departamento de Nutrición Vegetal,
Estación Experimental de Aula Dei,
Consejo Superior de Investigaciones Científicas,
Apartado 13034, 50080 Zaragoza, Spain
e-mail: becana@eead.csic.es

S. Tabata and J. Stougaard (eds.), *The Lotus japonicus Genome*, Compendium of Plant Genomes,
DOI 10.1007/978-3-662-44270-8_13,

peroxisomal, symbiosomal, and plasma membranes contain short electron transfer chains that generate ROS. In the plasma membrane, NADPH oxidases generate O_2^- and H_2O_2 and perform important functions both in plant immunity and in the symbiotic interaction (Marino et al. 2012). Production of O_2^- has been detected in root hairs during infection by compatible rhizobia (Santos et al. 2001; Cárdenas et al. 2008), and accumulation of H_2O_2 and formation of hydroxyl radicals have been observed in senescent nodules (Becana and Klucas 1992; Rubio et al. 2004).

Similarly to ROS, reactive nitrogen species (RNS) are generated in many subcellular compartments, including the mitochondria, peroxisomes, plastids, and bacteroids. Two RNS of major relevance in vivo are nitric oxide (NO), which acts as a signal in multiple developmental and stress responses, and *S*-nitrosoglutathione (GSNO), which is implicated in *trans*-nitrosylation reactions with cysteine thiol groups. NO has been detected in intact nodules and found to be produced by bacteroid and plant nitrate reductases (Meakin et al. 2007; Horchani et al. 2011) and by a plant NO synthase-like activity (Cueto et al. 1996). Other RNS with highly oxidizing and nitrating potential can be formed by the interaction of ROS with NO or heme groups. Peroxynitrite ($ONOO^-$) is produced by a reaction between O_2^- and NO, or between H_2O_2 and nitrite. Also, nitrogen dioxide (NO_2) can be produced by reaction of NO with the ferryl form of hemoglobin (Hb), a non-functional state of Hb generated by oxidation of the heme with H_2O_2. Formation of both $ONOO^-$ and NO_2 have been detected in plant tissues and a RNS species that is

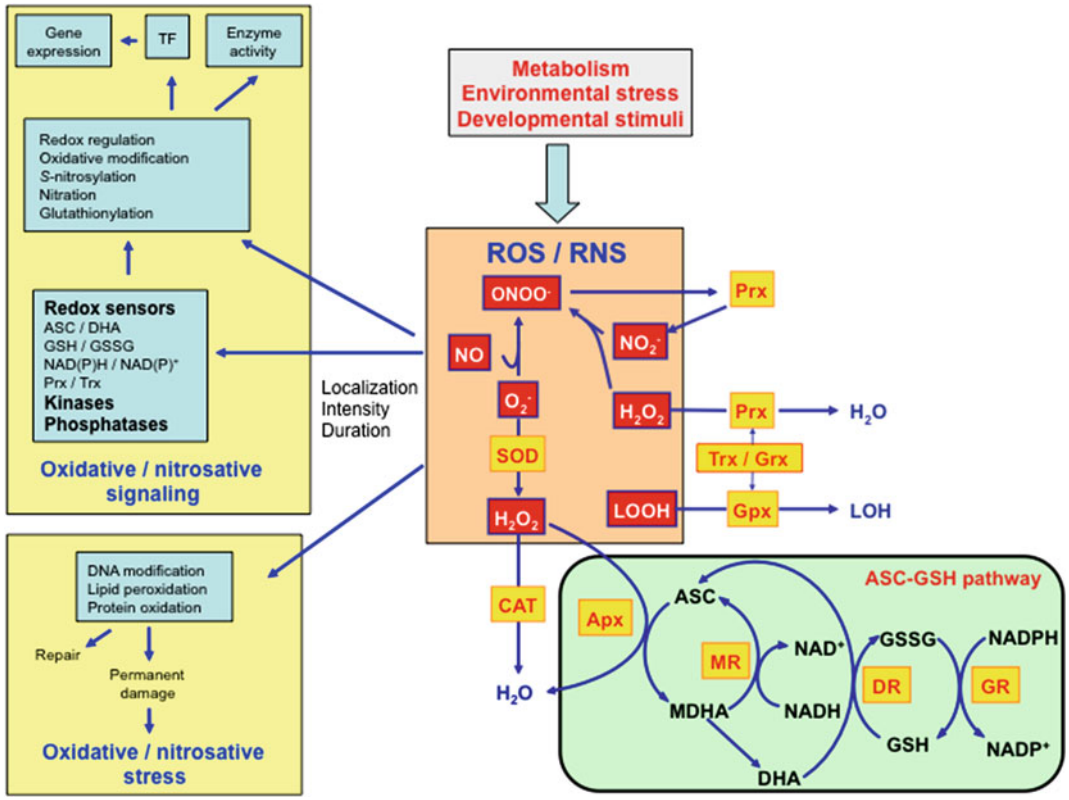


Fig. 13.1 Generalized scheme showing processes for generation and removal of ROS and RNS in legume root nodules. Additional abbreviations: *ASC* ascorbate, *CAT* catalase, *DHA* dehydroascorbate, *Grx* glutaredoxin, *LOOH* lipid peroxide, *MDHA* monodehydroascorbate, and *TF* transcription factor(s). Reproduced with permission from Becana et al. (2010)

capable of nitrating leghemoglobin (Lb) heme is produced in soybean nodules, especially during senescence (Navascués et al. 2012).

Many ROS and RNS perform useful functions in vivo, and their concentrations need to be kept under strict control to avoid cytotoxicity. This task is carried out by a highly complex and dynamic network of antioxidant enzymes and metabolites. In fact, ROS and RNS may, on their own or by interacting between them, act as molecular signals that trigger activation of genes involved in antioxidative protection and other defense processes. The subtle frontier between the useful roles of ROS and RNS, such as in signaling, organogenesis and stress responses, and the oxidative and nitrosative stress that they trigger when antioxidants fail to cope with them, is illustrated in Fig. 13.1.

13.2 Antioxidant Enzymes in Legumes

Plants are endowed with an impressive variety of antioxidant metabolites and enzymes. In particular, nodules are very rich in antioxidants, probably as a result of the diverse reactions that generate ROS and RNS in nodule host cells and bacteroids (Puppo et al. 2005; Becana et al. 2010). Here, we will briefly describe the antioxidant genes and proteins that are expressed in legumes and especially in nodules. A list of these genes has been compiled for the model legume *Lotus japonicus*, which is the subject of this collective book. Readers are referred to Table 13.1 and to the bibliography given in the text for a more complete description of antioxidants, which have been grouped, for clarity, according to their biochemical activities.

13.2.1 Superoxide Dismutases

The superoxide dismutase (SOD) family of enzymes catalyzes the dismutation of O_2^- to H_2O_2 and O_2, and is considered a primary line of defense against ROS. However, the resulting H_2O_2 also needs to be kept under control by other antioxidant enzymes (see below). There are three types of SODs depending on their metal cofactor: Cu and Zn, Fe or Mn. All of them have been found in nodules, roots, and leaves of *L. japonicus*. The genome of this legume encodes five SODs (Rubio et al. 2007; Table 13.1). The two CuZnSOD isoforms and the two FeSOD isoforms are each localized to the cytosol and plastids, whereas the MnSOD isoform is localized to the mitochondria. In addition, nodule bacteroids contain a MnSOD with significant homology to the plant isoform. Immunolocalization studies showed that CuZnSOD and FeSOD are also present in the nuclei, where they may perform useful roles by preventing oxidative damage of DNA and/or by modulating ROS levels and signaling. In *L. japonicus*, cytosolic CuZnSOD is localized in infection threads of incipient nodules and in the infected cells of young nodules, but FeSODs are localized in the cortex, vascular bundles, and infected zone at all stages of nodule development. Cytosolic CuZnSOD and mitochondrial MnSOD are transcriptionally downregulated during nodule development, whereas cytosolic FeSOD is upregulated. Based on these results, we have proposed that cytosolic CuZnSOD and FeSOD may functionally compensate each other at the late stages of nodule development (Rubio et al. 2007).

13.2.2 Catalases

These tetrameric heme proteins catalyze the decomposition of H_2O_2 to water and O_2 (Scandalios et al. 1997). However, the affinity of catalases for H_2O_2 is low compared with that of ascorbate peroxidase (Apx) and they may be efficient only at high H_2O_2 levels such as those produced in the peroxisomes, where the enzymes are primarily located. A single catalase gene has so far been identified in the *L. japonicus* genome (Table 13.1), although other plants such as *Arabidopsis thaliana* and maize express three catalase genes that are differentially regulated during development and in response to light and other

Table 13.1 Antioxidant proteins of *Lotus japonicus*

	CDS[a]	TC[b]	EST[b]	Localization[c]	UniRef100[d]
Superoxide dismutases					
CuZnSODc	chr1.CM0544.890.r2.m	76,148	27	cyt	Q56VR6
CuZnSODp	chr3.CM0396.620.r2.m	65,636	17	pl	O65198
FeSODc	chr5.CM0909.300.r2.m	57,161	17	cyt	Q53D71
FeSODp	chr6.CM0472.130.r2.m	57,158	12	pl	Q53D73
MnSOD	chr1.CM0125.150.r2.a	69,295	70	mit	Q56VR0
Catalase					
CAT	chr1.CM0178.260.r2.m	64,128	53	px	A0PG71
Ascorbate–glutathione pathway					
Apx1	chr3.CM0616.30.r2.d	72,396	118	cyt	Q43758
Apx3	chr3.LjT15F17.80.r2.d	69,072	32	px, mit?	Q5QIA9
Apx4	LjSGA_054875.1	57,581	35	pl (stroma)	Q5QHW6
Apx5	-	73,169	4	pl (thylakoid)	Q5QHW7
MR1	chr4.CM0004.930.r2.d	68,255, 63,143	11	pl, mit	Q94IB7, P92947
MR2	LjSGA_039892.1	65,440	16	px	Q66PF9
MR3	LjSGA_065200.1 LjSGA_039604.1 LjSGA_011141.3	57,632	14	sec, cyt	Q9LK94
DR1	LjSGA_026118.1	58,049	19	pl	Q4U3Z3
DR2	chr5.CM0180.360.r2.d	57,836	24	cyt	Q84UH4
GR1	chr5.CM1574.730.r2.m	61,990, 59,115, 67,131	20	pl, mit	P48640, P27456
GR2	LjSGA_027688.1, LjSGA_018577.1, LjSGA_073463.2	61,392, 66,284, 61,118	13	cyt	Q43621, A7XTY1, Q072J9
Ascorbate and thiol biosynthesis					
L-GalLDH	chr5.CM1813.90.r2.m chr5.CM1813.120.r2.m	57,542	22	mit	Q0ZHB8
γECS	chr4.CM0004.360.r2.m	60,483	13	pl	Q6XXZ2
GSHS	chr1.CM0544.610.r2.m	57,139	13	pl, cyt	Q93XE5
hGSHS	chr1.CM0544.600.r2.m	57,141	13	pl, cyt	Q6XPU3
Peroxiredoxins					
1CPrx	chr4.LjT20M01.140.r2.a	57,452	4	nu, cyt	Q6E2Z6
2CPrxA	chr1.CM0029.120.r2.m	75,376	80	pl	F1C3E5
2CPrxB	chr5.CM1005.20.r2.m	76,501	54	pl	F1C3E5
PrxQ1	chr4.CM0006.430.r2.m chr4.CM0006.420.r2.m	62,358	26	pl	G7JS60
PrxIIB	chr2.CM0660.240.r2.m	64,422	22	cyt	B3GV28
PrxIIE	chr1.CM0378.390.r2.d	76,090	20	pl	G1JT87
PrxIIF	chr6.CM0139.310.r2.m	60,826	16	mit	Q6KBB1
Glutathione peroxidases					
Gpx1	chr4.CM0004.300.r2.m	58,368	44	pl, mit	Q56VU1
Gpx2	chr4.CM0004.310.r2.m	59,133	8	cyt	A3FNZ8

(continued)

Table 13.1 (continued)

	CDS[a]	TC[b]	EST[b]	Localization[c]	UniRef100[d]
Gpx3	chr4.CM0042.1400.r2.m	60,457, 63,299	26	sec	Q56VS3
Gpx4	chr4.CM0042.1400.r2.m	76,738	2	cyt?	A7PU76
Gpx5	chr1.LjT23J20.90.r2.m	62,813	5	cyt?	A7PU76
Gpx6	chr5.CM0345.30.r2.m	57,520, 80,230	18	pl	O24296
Thioredoxins					
Trx*h1*	chr5.CM0077.790.r2.d	65,928	34	cyt	Q45NL7
Trx*h3*	chr2.CM0249.770.r2.m	68,183	12	cyt	I3S0R6
Trx*h4*	chr1.CM0051.240.r2.m	65,406	6	cyt	I3S917
Trx*h6*	chr2.CM0608.1340.r2.m	65,208	2	cyt	I3S341
Trx*h8*	LjSGA_031277.0.1	58,009	16	cyt	I3S3D9
Trx*h9*	LjSGA_132520.1	63,066	33	cyt	I3S146
Trx*f*	LjSGA_082631.1	59,402	23	pl	G7KRK1
Trx*m1*	chr5.CM1439.450.r2.d	60,229	8	pl	Q2PXN7
Trx*m2*	LjSGA_126827.0.1	71,331	15	pl	Q95AH9
Trx*m4*	LjSGA_126077.1 LjSGA_061545.0.1	67,299	20	pl	Q95AH9
Trx*x*	chr5.CM1439.450.r2.d	61,897	9	pl	G7IE85
Trx*y*	chr5.CM0052.830.r2.m	61,826	6	pl	A9PGA7
Trx*z*	LjSGA_025025.0.1	62,611	9	pl	Q9M7X9
Trx*o*	chr4.CM0007.270.r2.d chr4.CM0007.280.r2.d	61,209	19	mit	Q257C6
NTRA	-	63,269, 73,407	49	mit, cyt	K7LNQ6
NTRB	LjT16K13.10.r2.a	73,044, 80,146	6	mit, cyt	K7LNQ6
NTRC	LjSGA_039049.1 LjSGA_051261.0.1 LjSGA_046269.1	57,567, 68,679	9	pl	I1L9Y6
Nitrosoglutathione reductase					
GSNOR	chr1.CM0295.1290.r2.m	70,357	10	?	Q96533
Ferritins					
Fer1	chr3.CM0116.300.r2.m	68,510	237	pl, mit?	A5HKJ9
Fer2	LjSGA_064891.1 LjSGA_063587.1	64,989	39	pl, mit?	Q41709
Fer3	LjSGA_077559.1	57,335	25	pl, mit?	Q948P6
Phytochelatin synthases					
PCS1	chr1.CM0295.820.r2.m	57,148	2	?	Q2TSC7
PCS2	chr1.CM0295.830.r2.m	57,154	2	?	Q2TE74
PCS3	chr1.CM0295.840.r2.m	57,153	4	?	Q2QKL5
Hemoglobins					
Lb1	chr5.CM0089.1180.r2.m	74,540	153	cyt	Q3C1F7
Lb2	chr5.CM0089.1200.r2.m	67,341	321	cyt	Q3C1F6
Lb3	chr5.CM0034.610.r2.m	68,542	195	cyt	Q9FEP8
Glb1-1	chr3.CM0091.620.r2.m	61,058	36	nu, pl	Q3C1F4
Glb1-2	chr3.CM0091.630.r2.m	60,275	8	nu, pl	Q3C1F3

(continued)

Table 13.1 (continued)

	CDS[a]	TC[b]	EST[b]	Localization[c]	UniRef100[d]
Glb2	chr5.CM0909.850.r2.a	64,839	11	nu, pl	P14848
Glb3-1	-	65,804	3	nu, pl	A2TDC3
Glb3-2	chr1.CM2121.130.r2.a	59,615	4	nu, pl	A2TDC3

[a] Gene coding sequence (CDS) in release 2.5 of the *L. japonicus* genome. Detailed information on these CDS can be accessed through the Web database (http://www.kazusa.or.jp/lotus/)
[b] Tentative consensus (TC) sequences and number of expressed sequence tags (ESTs) according to the DFCI Lj Gene Index v6.0 (http://compbio.dfci.harvard.edu/tgi/plant.html)
[c] Predicted or observed subcellular localization: *cyt* cytosol, *mit* mitochondrion, *pl* plastid, *px* peroxisome, *sec* secretory pathway, and *nu* nucleus
[d] Best hit for the protein in the UniRef100 database

environmental factors (Scandalios et al. 1997). An alternative splice form of *L. japonicus* catalase could be detected by careful analysis of expressed sequence tags (ESTs). The predicted protein contains a modified C terminus, but the physiological role of this isoform is unknown.

13.2.3 Ascorbate-(homo)glutathione Biosynthetic Enzymes

The most abundant water-soluble antioxidants in plants and nodules are ascorbate (vitamin C), glutathione (GSH; γGlu-Cys-Gly), and homoglutathione (hGSH; γGlu-Cys-βAla). They can act as antioxidants on their own, by intercepting and destroying ROS and RNS, but also as substrates of enzymes with ROS scavenging activities. Ascorbate is synthesized mainly *via* the D-mannose/L-galactose (Smirnoff-Wheeler) pathway, involving multiple and complex sequential enzymatic reactions, the last of which is catalyzed by mitochondrial L-galactono-1,4-lactone dehydrogenase (Wheeler et al. 1998). Apparently, a single gene encoding this enzyme is present in the genome of *L. japonicus* (Table 13.1), as may be the case of other plants. Thiol tripeptides are synthesized *via* two sequential steps catalyzed by γ-glutamylcysteine synthetase (γECS), localized in the plastids, and by either glutathione synthetase (GSHS) or homoglutathione synthetase (hGSHS), both localized in the cytosol and plastids. However, hGSH and hGSHS are only present in some legumes such a *L. japonicus*, soybean, and common bean, where they may functionally replace GSH and GSHS. Some other legumes, such as *M. truncatula*, only express hGSHS in certain tissues (Frendo et al. 1999). In *L. japonicus*, two γECS genes and one gene each for GSHS and hGSHS appear to be present. The major γECS gene and the GSHS and hGSHS genes have been characterized (Matamoros et al. 2003).

Ascorbate and (homo)glutathione are also substrates for some enzymes of the ascorbate-(homo)glutathione (Foyer-Halliwell-Asada) pathway. The key enzyme of this pathway is Apx, which uses ascorbate to reduce H_2O_2 to water, has a high affinity for its substrate (~70 μM to achieve half maximum velocity), and accounts for 0.9 % of the total protein in soybean nodules (Dalton et al. 1987). Multiple isoforms of Apx exist in all plant tissues, which are localized to the chloroplasts (with both thylakoidal and stromal isoforms), mitochondria, peroxisomes, and cytosol. In *L. japonicus*, the five genes encoding the expected Apx isoforms have been identified (Table 13.1), but in nodules, most Apx activity corresponds to the cytosolic isoform. As a result of Apx activity, ascorbate is oxidized to monodehydroascorbate free radical and dehydroascorbate. These compounds are reduced back to ascorbate by monodehydroascorbate reductase (MR) and dehydroascorbate reductase (DR) using, respectively, NADH and GSH as reductants (Table 13.1). Plants contain MR isoforms in the plastids, mitochondria, and peroxisomes. Immunolocalization studies with an antibody raised against soybean MR showed that the enzyme in *L. japonicus* nodules is mainly associated with the cell walls, where MR may recycle

the monodehydroascorbate produced as a result of ascorbate oxidation during lignification. By contrast, the two DR isoforms are located to the plastids/mitochondria (dual targeting) and cytosol. The disulfide forms of glutathione (GSSG) and homoglutathione (hGSSG), generated as a result of DR activity, are reduced back to GSH and hGSH by (homo)glutathione reductases (GR) at the expense of NADPH. As occurs for DR, the two GR isoforms of *L. japonicus* and other plants are located in the plastids/mitochondria (dual targeting) and cytosol (Table 13.1). Therefore, the ascorbate-(homo) glutathione pathway comprises four enzymes with multiple isoforms and subcellular locations in order to keep H_2O_2 under control using NADH or NADPH as electron donors.

13.2.4 Nitrosoglutathione Reductase

The enzyme *S*-nitrosoglutathione reductase (GSNOR), previously known as glutathione-dependent formaldehyde dehydrogenase or class III alcohol dehydrogenase, is amply distributed from bacteria to humans. GSNOR activity does not release NO but produces GSSG and ammonia and regulates the levels of GSNO and *S*-nitrosylated proteins (Espunya et al. 2012). The enzyme is encoded by a single gene (*ADH2*) in *A. thaliana*. We have obtained the full sequence of the orthologous gene in *L. japonicus* (Table 13.1) and produced the recombinant enzyme, but it has not been characterized yet.

13.2.5 Peroxiredoxins, Glutathione Peroxidases, and Thioredoxins

Peroxiredoxins (Prxs) and glutathione peroxidases (Gpxs) are non-heme thiol peroxidases widespread in all organisms (Rouhier and Jacquot 2005). They show closely related biochemical properties and are present as multiple isoforms located in different cellular compartments. Prxs catalyze the reduction of H_2O_2 and organic peroxides and, in some cases, peroxynitrite. Gpxs also catalyze the reduction of H_2O_2 but are much more active with lipid peroxides as substrates and thus protect membranes from lipid peroxidation. Gpxs are considered a fifth class of Prxs, and all of them use preferentially thioredoxins (Trxs) as reductants.

In *L. japonicus*, we have identified most if not all Prx genes (Tovar-Méndez et al. 2011; Table 13.1). They belong to four classes and encode the following isoforms: one 1C-Prx (nuclei), two 2C-Prxs (A and B, in plastids), PrxQ (plastids), and three PrxII (B in cytosol, E in plastids, F in mitochondria). The 1C-Prx and PrxIIB genes are highly expressed in the embryo and pollen, respectively, whereas 2C-PrxA and PrxQ are more expressed in leaves than in roots and nodules, and PrxIIB and PrxIIF are similarly expressed in leaves, roots, and nodules. In *L. japonicus,* there are at least six Gpx genes, which have been mapped and functionally characterized (Ramos et al. 2009). The mRNA levels of Gpx1 (plastids/mitochondria) and Gpx6 (plastids) are especially abundant in leaves, and those of Gpx3 (cytosol/secretory pathway) and Gpx6 in nodules. The expression of Gpx6 was increased 30-fold after exposure to NO donors, which suggests a role of at least this isoform in stress and metabolic signaling.

Plants such as *A. thaliana* and rice have more than twenty thioredoxin (Trx) genes that are classified in seven groups. The Trx*f*, Trx*m*, Trx*x*, Trx*y,* and Trx*z* are localized in the chloroplasts, the Trx*h* isoforms in the cytosol, and Trx*o* in the mitochondria (Meyer et al. 2009). Oxidized Trxs produced as a result of reactions with Prxs and other substrates are reduced back to the functional state by NADPH–thioredoxin reductases (NTRA and NTRB) in the cytosol and mitochondria or by ferredoxin–thioredoxin reductase (FTR) in the chloroplasts (Jacquot et al. 2009). Another NADPH–thioredoxin reductase (NTRC) has been recently found in green tissues; this peculiar enzyme contains both NTR and Trx domains in the same polypeptide and may act as a complete NTR–Trx system (Spínola et al. 2008). A search of *L. japonicus* EST and genomic databases allowed us to identify fourteen Trx

genes (six Trx*h*, three Trx*m,* and one each of Trx*f*, Trx*x*, Trx*y*, Trx*z,* and Trx*o*), three NTR genes, and one FTRB gene (Tovar-Méndez et al. 2011) (Table 13.1). The NTRA, NTRB, and NTRC genes are expressed in leaves, roots, and nodules, but mRNA levels of NTRA are higher than those of NTRB and NTRC in all three organs. Based on gene expression and proteomic analyses, we have proposed that three NTR–Trx–Prx systems, localized to the cytosol, mitochondria, and plastids, may be operative in nodules (Tovar-Méndez et al. 2011).

13.2.6 Ferritins and Phytochelatins

Plants have evolved multiple strategies to maintain physiological concentrations of essential metals and to cope with heavy metal toxicity. One of them involves chelation of metal ions by ferritins and phytochelatins. Ferritins are large proteins of 24 subunits capable of concentrating up to 4,500 atoms of iron in a safe form. They are transcriptionally regulated, play an essential role in iron homeostasis, and protect plant cells against oxidative stress by preventing the participation of ferrous iron in damaging Fenton reactions (Briat et al. 2010). There are four active ferritin genes in cowpea and *A. thaliana* that display tissue-specific expression and differential regulation during development and in response to environmental cues. The proteins of legumes and other plants have been mainly localized to the plastids (Lucas et al. 1998), although a mitochondrial isoform has been recently detected (Briat et al. 2010). The ferritin genes of *L. japonicus* have been identified but not characterized yet (Table 13.1).

Phytochelatins are polypeptides of general structure $(\gamma\text{Glu-Cys})_{2-11}$-Gly that are synthesized by dipeptidyl transferases called phytochelatin synthases (PCS). The reaction entails the net transfer of a γGlu-Cys unit from GSH to another GSH molecule or to an elongating PC polypeptide (Clemens 2006). In *L. japonicus* and some other hGSH-producing legumes, homophytochelatins of general structure $(\gamma\text{Glu-Cys})_{2-11}$-βAla can be synthesized also by PCS using GSH plus hGSH or hGSH alone as substrates (Loscos et al. 2006). The PCS reactions are strictly dependent on the presence of metal ions. We have identified three functional PCS genes in *L. japonicus* and found that they are differentially expressed in response to Cd (Ramos et al. 2007, 2008) (Table 13.1). The PCS1 gene encodes a protein with high homology to soybean PCS1 (84 % amino acid identity). The PCS2 and PCS3 genes encode proteins that are closely related to each other (90 % identity) but are distant in evolutionary terms (53–56 % identity) from PCS1. The PCS2 and PCS3 genes each show two alternatively spliced forms. Interestingly, the nodule form of PCS2 (PCS2N) conferred tolerance to cadmium when expressed in yeast, whereas the root form (PCS2R) did not, indicating a complex regulation of PCS expression in organisms in response to heavy metals (Ramos et al. 2007, 2008).

13.2.7 Hemoglobins

Plants can express up to three classes of Hbs: nonsymbiotic, symbiotic, and truncated (see reviews by Garrocho-Villegas et al. 2007; Hoy and Hargrove 2008; Gupta et al. 2011). Nonsymbiotic Hbs occur at concentrations of ~100 nM in many tissues and are further categorized into class 1 and class 2 based on phylogenic relationships, gene expression profiles, and O_2-binding properties. Class 1 Hbs display high O_2 affinities and modulate NO concentration in stressful conditions (Igamberdiev and Hill 2004; Gupta et al. 2011). Class 2 Hbs have O_2 affinities resembling those of symbiotic Hbs (Hunt et al. 2001), and their functions are largely unknown. Symbiotic Hbs include the Lbs found at concentrations of 1–5 mM in legume nodules, where they facilitate a steady low O_2 supply to the bacteroids, thus avoiding nitrogenase inactivation. Class 3 or 'truncated' Hbs have a 2-on-2 α-helical sandwich secondary structure instead of the canonical 3-on-3 structure of other Hbs. Although virtually nothing is known about their function in plants, some of their bacterial counterparts have been

implicated in tolerance to nitrosative stress (Garrocho-Villegas et al. 2007; Hoy and Hargrove 2008).

In addition to three genes of symbiotic Lbs, the genome of *L. japonicus* contains genes encoding non-symbiotic and truncated Hbs, which are expressed in nodules and other plant organs (Nagata et al. 2008; Bustos-Sanmamed et al. 2011) (Table 13.1). These genes encode two class 1 Hbs (Glb1-1 and Glb1-2), one class 2 Hb (Glb2), and two class 3 Hbs (Glb3-1 and Glb3-2). This gene profile may be extended to other legumes because two class 3 Hbs are expressed in *M. truncatula* (Vieweg et al. 2005) but is in contrast with *A. thaliana*, which only contains one globin of each class (Hunt et al. 2001). In particular, the Glb1-1, Glb2, and Glb3-2 mRNAs are abundant in nodules and mainly localized to the vascular bundles, cortex, and infected tissue (Bustos-Sanmamed et al. 2011). Expression of Hb genes is greatly affected by hormonal treatment of plants, and these effects are organ dependent. Cytokinins suppress expression of Glb2 and Glb3-1 in nodules but induce Glb1-1 in roots, whereas polyamines and jasmonic acid induce Glb1-1 only in roots (Bustos-Sanmamed et al. 2011). These observations suggest that Hbs act downstream of hormones in signaling or regulatory pathways and that their functions are rather specific for the corresponding hormones and target tissues.

The very high O_2 affinities of class 1 Hbs make them not able to act as O_2 carriers, and a number of alternative functions have been proposed, including modulation of NO levels, maintenance of cellular energetics under hypoxic conditions, and O_2 scavenging (Igamberdiev and Hill 2004; Gupta et al. 2011). We have produced *L. japonicus* Hbs in recombinant form and have characterized them by measuring O_2 affinities and other biochemical properties. Glb1-1 displays the highest O_2 affinity (K^{O_2} ~50 pM) known for a plant or animal Hb. Glb1-2 (K^{O_2} ~0.9 nM) still has too high affinity for O_2 transport, whereas Glb2 (K^{O_2} ~11 nM) has an O_2 affinity similar to soybean Lb*a* (K^{O_2} ~43 nM) and thus is suitable for O_2 transport and delivery in plant cells. However, the low concentration of Glb2 in cells may hamper this function. Most Hbs of *L. japonicus* were expressed in yeast and found to confer tolerance to oxidative stress, probably as a result of ROS scavenging by the hemes. By contrast, only Glb1-2 and Glb2 afford protection against nitrosative stress induced by GSNO, suggesting that cysteine residues are implicated in NO detoxification. These results, along with those of others (Igamberdiev and Hill 2004; Gupta et al. 2011), indicate that Hbs can act as antioxidants by regulating ROS and RNS concentrations in vivo.

Acknowledgments We thank our friend and colleague Shusei Sato for invaluable help in identifying CDS for Table 13.1. Research of our laboratory was funded by the Ministry of Economy and Competitivity (grants AGL2008-01298 and AGL2011-24524, cofunded by Fondo Europeo de Desarrollo Regional), and by Gobierno de Aragón (group A53).

References

Becana M, Klucas RV (1992) Transition metals in legume root nodules. Iron dependent free radical production increases during nodule senescence. Proc Natl Acad Sci USA 89:8958–8962

Becana M, Matamoros MA, Udvardi M, Dalton DA (2010) Recent insights into antioxidant defenses of legume root nodules. New Phytol 188:960–976

Briat JF, Ravet K, Arnaud N, Duc C, Boucherez J, Touraine B, Cellier F, Gaymard F (2010) New insights into ferritin synthesis and function highlight a link between iron homeostasis and oxidative stress in plants. Ann Bot 105:811–822

Bustos-Sanmamed P, Tovar-Méndez A, Crespi M, Sato S, Tabata S, Becana M (2011) Regulation of nonsymbiotic and truncated hemoglobin genes of *Lotus japonicus* in plant organs and in response to nitric oxide and hormones. New Phytol 189:765–776

Cárdenas L, Martínez A, Sánchez F, Quinto C (2008) Fast, transient and specific intracellular ROS changes in living root hair cells responding to Nod factors (NFs). Plant J 56:802–813

Clemens S (2006) Evolution and function of phytochelatin synthases. J Plant Physiol 163:319–332

Cueto M, Hernández-Perera O, Martín R, Bentura ML, Rodrigo J, Lamas S, Golvano MP (1996) Presence of nitric oxide synthase activity in roots and nodules of *Lupinus albus*. FEBS Lett 398:159–164

Dalton DA, Hanus FJ, Russell SA, Evans HJ (1987) Purification, properties, and distribution of ascorbate

peroxidase in legume roots nodules. Plant Physiol 83:789–794

Espunya MC, De Michele R, Gómez-Cadenas A, Martínez MC (2012) *S*-nitrosoglutathione is a component of wound- and salicylic acid-induced systemic responses in *Arabidopsis thaliana*. J Exp Bot 63:3219–3227

Frendo P, Gallesi D, Turnbull R, Van de Sype G, Hérouart D, Puppo A (1999) Localisation of glutathione and homoglutathione in *Medicago truncatula* is correlated to a differential expression of genes involved in their synthesis. Plant J 17:215–219

Garrocho-Villegas V, Gopalasubramaniam SK, Arredondo-Peter R (2007) Plant hemoglobins: what we know six decades after their discovery? Gene 398:78–85

Gupta KJ, Hebelstrup KH, Mur LAJ, Igamberdiev AU (2011) Plant hemoglobins: important players at the crossroads between oxygen and nitric oxide. FEBS Lett 585:3843–3849

Halliwell B, Gutteridge JMC (2007) Free radicals in biology and medicine, 4th edn. Oxford University Press, Oxford

Horchani F, Prévot M, Boscari A, Evangelisti E, Meilhoc E, Bruand C, Raymond P, Boncompagni E, Aschi-Smiti S, Puppo A, Brouquisse R (2011) Both plant and bacterial nitrate reductases contribute to nitric oxide production in *Medicago truncatula* nitrogen-fixing nodules. Plant Physiol 155:1023–1036

Hoy JA, Hargrove MS (2008) The structure and function of plant hemoglobins. Plant Physiol Biochem 46:371–379

Hunt PW, Watts RA, Trevaskis B, Llewelyn DJ, Burnell J, Dennis ES, Peacock WJ (2001) Expression and evolution of functionally distinct haemoglobin genes in plants. Plant Mol Biol 47:677–692

Igamberdiev AU, Hill RD (2004) Nitrate, NO and haemoglobin in plant adaptation to hypoxia: an alternative to classic fermentation pathways. J Exp Bot 55:2473–2482

Jacquot JP, Eklund H, Rouhier N, Schürmann P (2009) Structural and evolutionary aspects of thioredoxin reductases in photosynthetic organisms. Trends Plant Sci 14:336–343

Loscos J, Naya L, Ramos J, Clemente MR, Matamoros MA, Becana M (2006) A reassessment of substrate specificity and activation of phytochelatin synthases from model plants by physiologically relevant metals. Plant Physiol 140:1213–1221

Lucas MM, Van de Sype G, Hérouart D, Hernández MJ, Puppo A, de Felipe MR (1998) Immunolocalization of ferritin in determinate and indeterminate legume root nodules. Protoplasma 204:61–70

Marino D, Dunand C, Puppo A, Pauly N (2012) A burst of plant NADPH oxidases. Trends Plant Sci 17:9–15

Matamoros MA, Clemente MR, Sato S, Asamizu E, Tabata S, Ramos J, Moran JF, Stiller J, Gresshoff PM, Becana M (2003) Molecular analysis of the pathway for the synthesis of thiol tripeptides in the model legume *Lotus japonicus*. Mol Plant Microbe Interact 16:1039–1046

Meakin GE, Bueno E, Jepson B, Bedmar EJ, Richardson DJ, Delgado MJ (2007) The contribution of bacteroidal nitrate and nitrite reduction to the formation of nitrosylleghaemoglobin complexes in soybean root nodules. Microbiology 153:411–419

Meyer Y, Buchanan BB, Vignols F, Reichheld JP (2009) Thioredoxins and glutaredoxins: unifying elements in redox biology. Annu Rev Genet 43:335–367

Nagata M, Murakami E, Shimoda Y, Shimoda-Sasakura F, Kucho K, Suzuki A, Abe M, Higashi S, Uchiumi T (2008) Expression of a class 1 hemoglobin gene and production of nitric oxide in response to symbiotic and pathogenic bacteria in *Lotus japonicus*. Mol Plant Microbe Interact 21:1175–1183

Navascués J, Pérez-Rontomé C, Gay M, Marcos M, Yang F, Walker FA, Desbois A, Abián J, Becana M (2012) Leghemoglobin green derivatives with nitrated hemes evidence production of highly reactive nitrogen species during aging of legume nodules. Proc Natl Acad Sci USA 109:2660–2665

Puppo A, Groten K, Bastian F, Carzaniga R, Soussi M, Lucas MM, de Felipe MR, Harrison J, Vanacker H, Foyer CH (2005) Legume nodule senescence: roles for redox and hormone signalling in the orchestration of the natural aging process. New Phytol 165:683–701

Ramos J, Clemente MR, Naya L, Loscos J, Pérez-Rontomé C, Sato S, Tabata S, Becana M (2007) Phytochelatin synthases of the model legume *Lotus japonicus*. A small multigene family with differential response to cadmium and alternatively spliced variants. Plant Physiol 143:1110–1118

Ramos J, Naya L, Gay M, Abián J, Becana M (2008) Functional characterization of an unusual phytochelatin synthase, LjPCS3, of *Lotus japonicus*. Plant Physiol 148:536–545

Ramos J, Matamoros MA, Naya L, James EK, Rouhier N, Sato S, Tabata S, Becana M (2009) The glutathione peroxidase gene family of *Lotus japonicus*: characterization of genomic clones, expression analyses and immunolocalization in legumes. New Phytol 181:103–114

Rouhier N, Jacquot JP (2005) The plant multigenic family of thiol peroxidases. Free Radic Biol Med 38: 1413–1421

Rubio MC, James EK, Clemente MR, Bucciarelli B, Fedorova M, Vance CP, Becana M (2004) Localization of superoxide dismutase and hydrogen peroxide in legume root nodules. Mol Plant Microbe Interact 17:1294–1305

Rubio MC, Becana M, Sato S, James EK, Tabata S, Spaink HP (2007) Characterization of genomic clones and expression analysis of the three types of superoxide dismutases during nodule development in *Lotus japonicus*. Mol Plant Microbe Interact 20:262–275

Santos R, Hérouart D, Sigaud S, Touati D, Puppo A (2001) Oxidative burst in alfalfa-*Sinorhizobium meliloti* symbiotic interaction. Mol Plant Microbe Interact 14:86–89

Scandalios JG, Guan L, Polidoros AN (1997) Catalases in plants: gene structure, properties, regulation, and expression. In: Scandalios JG (ed) Oxidative stress and the molecular biology of antioxidant defenses.

Cold Spring Harbor Laboratory Press, Plain View, New York, pp 343–406

Spínola MC, Pérez-Ruiz JM, Pulido P, Kirchsteiger K, Guinea M, González M, Cejudo FJ (2008) NTRC new ways of using NADPH in the chloroplast. Physiol Plant 133:516–524

Tovar-Méndez A, Matamoros MA, Bustos-Sanmamed P, Dietz KJ, Cejudo FJ, Rouhier N, Sato S, Tabata S, Becana M (2011) Peroxiredoxins and NADPH-dependent thioredoxin systems in the model legume *Lotus japonicus*. Plant Physiol 156:1535–1547

Vieweg MF, Hohnjec N, Küster H (2005) Two genes encoding different truncated hemoglobins are regulated during root nodule and arbuscular mycorrhiza symbioses of *Medicago truncatula*. Planta 220:757–766

Wheeler GL, Jones MA, Smirnoff N (1998) The biosynthetic pathway of vitamin C in higher plants. Nature 393:365–369

Plant-Specialized Metabolism and Its Genomic Organization in Biosynthetic Gene Clusters in *Lotus japonicus*

14

Adam M. Takos and Fred Rook

Abstract

Plants produce a wide spectrum of specialized metabolites that function in plant chemical defense against pathogens and herbivores or have signaling roles in the interaction with other organisms. The plant-specialized metabolites that have received most attention in legumes in general, and in *Lotus japonicus* as a legume model species, are proanthocyanidins, isoflavonoids, cyanogenic and non-cyanogenic hydroxynitrile glucosides, and triterpenoids. Here, we review these four classes of plant-specialized metabolites in terms of the specific compounds produced by *L. japonicus*, the biosynthetic genes responsible, and the genomic organization of the genes. We previously reported that in *L. japonicus,* the non-homologous genes encoding the complete biosynthetic pathway for the cyanogenic glucosides lotaustralin and linamarin are organized in a gene cluster. Here, we additionally describe gene clusters in the *L. japonicus* genome for triterpenoid and isoflavonoid biosynthesis. A model explaining how selection for reduced recombination results in gene cluster formation is presented.

14.1 Plant-Specialized Metabolism in Legumes

Plants species produce a large number and wide diversity of specialized metabolites, also known as secondary metabolites, which function in the interaction between the plant and its environment. Many of these compounds have a role as chemical defense compounds, protecting the plant against herbivores or combatting bacterial and fungal infections. Other compounds serve roles in attracting pollinators, or specifically for flavonoids in legumes, have a signaling role in the interaction with symbiotic bacteria. Four classes of such plant-specialized metabolites have received most attention in legumes: proanthocyanidins, isoflavonoids, triterpenoids, and cyanogenic glucosides (Fig. 14.1). This attention is partly due to the importance of forage legumes to animal nutrition and the effects these compounds have on digestion and animal health. In addition, plant chemical defense and its evolution are a topic of major scientific interest. The genomic organization of these biosynthetic pathways in *Lotus japonicus* reveals

A.M. Takos · F. Rook (✉)
Plant Biochemistry Laboratory, Department of Plant and Environmental Sciences, University of Copenhagen, Thorvaldsensvej 40, 1871 Frederiksberg, Denmark
e-mail: frro@plen.ku.dk

S. Tabata and J. Stougaard (eds.), *The Lotus japonicus Genome*, Compendium of Plant Genomes,
DOI 10.1007/978-3-662-44270-8_14,

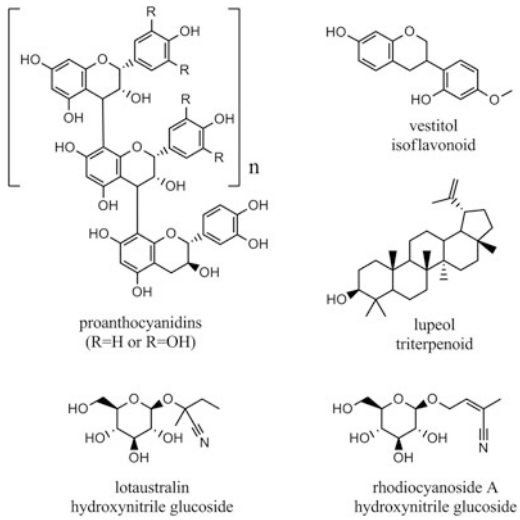


Fig. 14.1 Example compounds of classes of plant-specialized metabolites present in *L. japonicus*. The polymeric proanthocyanidins (a schematic structure is drawn), the isoflavonoid vestitol, the triterpenoid lupeol, and the cyanogenic α-hydroxynitrile glucoside lotaustralin and the non-cyanogenic γ-hydroxynitrile glucoside rhodiocyanoside A

some of the evolutionary dynamics resulting from the interaction between a plant species and its biotic environment. Gene duplication is highly relevant to the evolution of metabolic diversity as it is often followed by functional divergence, or alternatively the formation of pseudogenes. As a result, clusters of functional and non-functional versions of a specific biosynthetic gene are commonly observed in plant genomes, particular in plant-specialized metabolism. However, more remarkable is the organization of several of these metabolic pathways in gene clusters consisting of non-homologous genes encoding different biosynthetic enzymes of the same pathway. Here, we describe three such biosynthetic gene clusters present in the genome of *L. japonicus* and discuss the evolutionary mechanism responsible for their formation.

14.2 Proanthocyanidins

Proanthocyanidins (PAs), also known as condensed tannins, are polyphenolic compounds synthesized by a branch of the phenylpropanoid pathway that also produces anthocyanins and flavonols. They have a polymeric structure consisting of flavan-3-ol units linked in various ways. PAs occur in a wide range of plants where their primary role is defense against herbivores by being toxic to insects and by decreasing protein digestion in vertebrates (Barbehenn and Constabel 2011). There has been significant interest in manipulation of PA levels in forage legume crop species as their ability to precipitate protein reduces pasture bloat and improves animal productivity (Patra and Saxena 2011). PAs are also being studied because of the proposed role of polyphenols in the prevention of cardiovascular disease when included in the human diet (Quinones et al. 2013). Metabolic engineering of PA content in various crop plants through transgenic approaches requires a detailed understanding of the PA biosynthetic genes and their regulation, and considerable progress has been made in this area (Bogs et al. 2007; Dixon et al. 2013).

Leaf PA content of 31 Lotus species and accessions was determined by Gruber et al. (2008) and varied from undetectable in seven species (including *L. japonicus* Gifu B129) to very high in *L. unifoliolatus* (4.1 % FW). In *L. japonicus,* PAs accumulate in floral organs, seeds, and stems, but not in leaves (Skadhauge et al. 1997). There was no relationship between PA content and morphological traits or geographic location. A wide variation in chemical composition, mean degree of polymerization (mDP), and degree of uniformity of PAs was found in twelve *Lotus* spp. by Sivakumaran et al. (2006). The mDP varied from 8 to 97 units, although most species had a mDP of less than 20, with only a high mDP found in *L. pedunculatus* and *L. americanus*. There was considerable variability in the hydroxylation pattern of the B-ring of the flavan-3-ol PA polymer extension units. The dihydroxylated procyanidin (PC) units or trihydroxylated prodelphinidin (PD) units were predominant in some species and approximately equal in others. Regarding the stereochemistry at the C-ring, in all species, the 2,3-*cis* isomers epicatechin and epigallocatechin were predominant as extension units, while the 2,3-*trans* isomer catechin was the typical terminal unit of the polymer.

The monomeric flavan-3-ol units are synthesized on the cytoplasmic side of the endoplasmic reticulum and are believed to be subsequently condensed into polymers in the vacuole (Zhao et al. 2010). The first committed enzyme to flavonoid biosynthesis is a polyketide synthase named chalcone synthase (CHS), which catalyzes the condensation of three malonyl-CoA molecules with one 4-coumaroyl-CoA molecule to from naringenin chalcone. At least thirteen *CHS* genes have been identified in the *L. japonicus* genome, of which eight genes (*CHS4-11*) cluster on chromosome 2 (contig CM0018, Shimada et al. 2007). Phylogenetic analysis suggests that this diversity of *CHS* enzymes was generated after the divergence of the legume clade and that *CHS1* (chr1.CM0104.1140) represents a non-leguminous type. Chalcone isomerase (CHI) catalyzes the stereo-specific cyclization of the chalcones to form the generic flavonoid C6-C3-C6 aromatic ring structure. *L. japonicus* contains four *CHI* genes in a 15-kb region of chromosome 5 (contig CM0180, Shimada et al. 2003, 2007). It is of interest to note that while CHI1, CHI3, and CHI4 are close homologs that are 90 % identical at the amino acid sequence level, CHI2 only showed 50 % amino acid identity with the other three lotus enzymes. CHI2 is similar to non-leguminous CHIs (referred to as type I) and catalyzes the isomerization of 6′hydroxychalcone to the 5′-hydroxyflavone naringenin, which is a precursor of all flavonoids including the flavan-3-ol units of PAs. CHI1 and CHI3 are more similar to legume-specific CHIs (referred to as type II), which in addition to the type I reaction also convert 6′-deoxychalcone to 5′-deoxyflavone, the precursor for isoflavonoid biosynthesis.

The next step is the hydroxylation of the C-ring by the non-heme iron enzyme flavanone 3-hydroxylase (F3H) to give the dihydrokaempferol. The cytochrome P450 enzyme flavonoid 3′-hydroxylase (F3′H) catalyzes the hydroxylation at the 3 position of the B-ring to form dihydroquercetin. This is a precursor for the procyanidin flavan-3-ol units that will be incorporated into PAs. A second cytochrome P450, flavonoid 3′-5′-hydroxylase (F3′5′H), catalyzes the hydroxylation of dihydroquercetin at the 5 position of the B-ring to form dihydromyricetin. Conversely, this is the precursor for the prodelphinidin flavan-3-ol units of PAs. Genes encoding F3H, F3′H, and F3′5′H enzymes have not been investigated in any detail in *L. japonicus*, but candidate genes can be readily identified in its genome sequence. The subsequent reduction of the various B-ring hydroxylated intermediates to leucoanthocyanidins is catalyzed by a reductase–epimerase–dehydrogenase (RED) super family member, the enzyme dihydroflavonol reductase (DFR). This is the first committed enzyme of the anthocyanin and PA branch of the flavonoid pathway. Five DFR genes that form a cluster in a 38-kb region on chromosome 5 (on contig CM0077) were previously characterized (Shimada et al. 2005). The encoded DFRs have alternative substrate preferences for the various B-ring hydroxylated dihyroflavonols, and variation in the expression of these genes in *Lotus* species could explain the differences in hydroxylation patterns of PA units, but additional regulatory mechanisms are also thought to exist (Shimada et al. 2005). It was more recently shown that only the *DFR2* promoter is activated by the transcription factor LjTT2, a known MYB regulator of PA biosynthesis genes (Yoshida et al. 2010).

The leucoanthocyanidins are an intermediate in anthocyanin biosynthesis and also a substrate for the first specific step for PA biosynthesis. It is speculated that they may also be used as extension units of PA polymers (Pang et al. 2007). Their reduction to 2,3-*trans*-flavan-3-ols such as catechin, the typical *L. japonicus* PA terminal unit, is catalyzed by leucoanthocyanidin reductase (LAR). Like DFR, LAR is a member of the RED super family but is more closely related to isoflavone reductase-like proteins than DFR. Two cDNAs encoding putative LARs have been cloned from *L. corniculatus*, both of which are expressed in leaf tissue (Paolocci et al. 2007). When both types of LcLARs were expressed in *E. coli,* they were able to catalyze the formation of catechin. In *L. japonicus,* the gene corresponding to *LcLAR1* is probably Chr2.CM0124.20 and a DNA fragment encoding a similar sequence to *LcLAR2* is also present (LjSGA_076819.1).

The leucoanthocyanidins can be oxidized by anthocyanidin synthase (ANS), also called leucoanthocyanidin dioxygenase (LDOX), a non-heme iron 2-oxoglutarate-dependent oxygenase to form anthocyanidin. The glycosylation of this compound (along with methylation and acylation) forms the stable color pigments anthocyanins. The anthocyanidins are also substrates for a second PA-specific enzyme, anthocyanidin reductase (ANR) that catalyzes their reduction to 2,3-*cis* flavan-3-ols, for example epicatechin, which is the predominant extension unit in PAs of *Lotus* spp. The ANR enzyme is also a member of the RED super family and was first identified in Arabidopsis (Xie et al. 2003). That anthocyanidins are a substrate fits with the observation that ANS is essential for PA biosynthesis in Arabidopsis (Abrahams et al. 2003). Two *ANR* genes have been identified in *L. corniculatus*; *LcANR2* appears to be a pseudogene as it lacks several exons, while *LcANR1* when expressed in *E. coli* was able to catalyze the formation of epicatechin (Paolocci et al. 2007). In *L. japonicus,* the gene corresponding to *LcANR1* is probably Chr4.CM1616.680.

The flavan-3-ol precursors of PAs accumulate on the cytoplasmic side of the ER, but PA polymerization is believed to occur in the vacuole. The process of glycosylation, transport, and polymerization of PAs is still poorly understood, and we only mention it briefly. The first transporter identified in PA biosynthesis was the Arabidopsis TT12, a multidrug and toxic compound extrusion (MATE) family protein (Debeaujon et al. 2001). A *M. truncatula* transporter, MATE1, complements the Arabidopsis *tt12* mutant phenotype and both Arabidopsis TT12 and Medicago MATE1 appear to prefer epicatechin 3′-O-glucoside as substrate (Zhao and Dixon 2009). A glycosyltransferase with activity toward epicatechin, UGT72L1, was identified by transcript profiling of Medicago hairy roots expressing the Arabidopsis transcription factor TT2 (Pang et al. 2008). The closest related gene to *MATE1* in the genome of *L. japonicus* is chr2.LjT36E17.20.

14.3 Isoflavonoids and Their Role in Plant–Microbe Interactions

Isoflavonoids represent a different branch of the phenylpropanoid pathway and are considered characteristic defense compounds of legumes, although they may also act as signaling molecules in symbiotic interactions. Like PAs, isoflavonoids have gained interest because of their proposed health benefits in humans, such as a reduced incidence of hormone-related cancers attributed to their phytoestrogen properties (Cornwell et al. 2004). The estrogenic activity of isoflavonoids in forage legumes can lead to breeding and infertility problems in farm animals, a condition known as clover disease (Mustonen et al. 2006).

Isoflavonoids differ from other flavonoids by the phenolic B-ring being attached to the C-3 of the C-ring instead of to the C-2 like in other flavonoids (Fig. 14.2). Although considered a characteristic class of compounds for the Fabaceae, isoflavonoids are not unique to them and have been reported in plant species from at least 59 other plant families (Reynaud et al. 2005; Lapcik 2007). For example, a wide variety of isoflavones and their glycosides occur in the genus *Iris*, also demonstrating their presence in both monocots and dicots (Wang et al. 2010). It should, however, be noted that some reports of the occurrence of isoflavonoids in non-legume species relate to the presence of trace amounts that may result from side reactions of enzymes with related primary functions. Such enzymatic promiscuity is common in plant-specialized metabolism and provides the evolutionary origin for the selection of alternative enzymatic functionalities and the emergence of new biosynthetic pathways (Khersonsky and Tawfik 2010).

Isoflavonoids function as phytoalexins, and their biosynthesis is induced in response to, for instance, microbial infection or experimental treatments with elicitors such as reduced glutathione. Isoflavonoids known to be produced by *L. japonicus*, and the related *L. corniculatus*, are

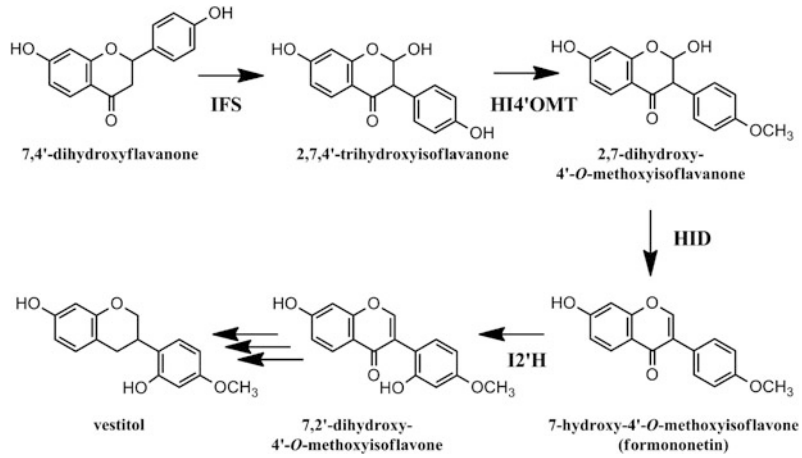


Fig. 14.2 Early steps in the isoflavonoid biosynthesis pathway branching off from general flavonoid metabolism. Enzymatic steps are described in the text

vestitol and sativan (Bonde et al. 1973), but additional isoflavonoid compounds and some derived glycosides are also thought to be produced. Natural variation in isoflavonoid composition seems to exist within the *L. japonicus* species, for instance, we have identified isosativan as a prominent compound in the MG74 accession. In *L. japonicus*, vestitol is the isoflavonoid that has been investigated the most extensively in terms of its biosynthesis and its physiological role in the defense against microbial pathogens. Inoculation of lotus roots with *Calonectoria ilicola*, a fungus pathogenic to soybean but not to lotus, induced the expression of key genes in vestitol biosynthesis and led to vestitol exudation from the roots (Masunaka et al. 2011). In contrast, inoculation with *Trichoderma koningi*, a plant growth promoting fungus able to colonize the roots of *L. japonicus*, only led to brief transient induction of gene expression. Inoculation with the symbiotic bacterium *Mesorhizobium loti* did not upregulate isoflavonoid biosynthetic gene expression or vestitol levels in root exudates. Vestitol also has a defense role in plant–plant interactions as its biosynthesis was highly upregulated in lotus roots challenged with the incompatible parasitic plant *Striga hermonthica* (Ueda and Sugimoto 2010).

Flavonoids and isoflavonoids not only function as antimicrobial defense compounds but also as plant-derived signaling molecules that induce the expression of nodulation genes in symbiotic *Rhizobium* species and are as such an important early factor in the successful establishment of symbiosis (reviewed in Hassan and Mathesius 2012). The specificity for which plant compound acts as an inducer of *nod* gene expression depends on the bacterial strain. This was demonstrated by Kosslak et al. (1987) who monitored *nod* gene expression by measuring the β-galactosidase activity resulting from a *nodABC-lacZ* translational fusion construct. Using the response of *Bradyrhizobium japonicum*, the original source of the *nodABC* operon, two *nod* gene-inducing compounds present in soybean root extracts were identified as the isoflavones daidzein and genistein (Kosslak et al. 1987). Following the transfer of the *nodABC-lacZ* construct into *Rhizobium trifolii,* it showed no induction by soybean root extracts or individual isoflavones, but the construct could be induced by flavones able to induced *R. trifolii nod* genes, such as 4′,7-dihydroxyflavone, apigenin, and luteolin. Surprisingly, for a model system to study symbiosis, the identity of the flavonoid or non-flavonoid compounds produced by *Lotus* species that induce nod factor production in its symbiotic partner *Mesorhizobium loti* have not been identified, but the phytoalexins vestitol and sativan seem not to be involved (Cooper 2007; Rispail et al. 2010).

14.4 Biosynthesis of the Isoflavonoid Vestitol in *L. japonicus*

Shimada et al. (2007) proposed a biosynthetic pathway for vestitol via medicarpin. Starting from the first committed enzymatic step of the isoflavonoid branch of the phenylpropanoid pathway, the aryl migration reaction of the B-ring from C-2 to the C-3 position, this pathway included eight enzymatic steps. Here, we briefly describe the sequential steps in the biosynthesis of vestitol with emphasis on the first half of the pathway. The early steps in vestitol biosynthesis are the most determinative and represent a branch point from general phenylpropanoid metabolism. As vestitol is a 5-deoxyisoflavonoid, its synthesis involves the legume type II CHIs mentioned earlier, to produce 7,4′-dihydroxyflavanone (Shimada et al. 2003). This is the substrate for the key branch point enzyme isoflavone synthase (IFS), a cytochrome P450 of the CYP93C family catalyzing the aryl migration reaction (Fig. 14.2), resulting in the production of 2,7,4′-trihydroxyisoflavanone. Two functional IFS genes (*IFS1,* chr4.CM0432.2900, and *IFS2*, ch4.CM0432.3190) have been identified in *L. japonicus* (Shimada et al. 2000, 2007) and in the section on the genomic organization of the isoflavonoid biosynthetic pathway, we will describe how these genes form the core of a cluster of non-homologous isoflavonoid biosynthetic genes. The acid-labile product of the IFS reaction was shown to be the substrate for the *O*-methyltransferase HI4′OMT (SAM: 2,7,4′-trihydroxyisoflavanone 4′-*O*-methyltransferase, chr4.CM0432.2880), resulting in the formation of 2,7-dihydroxy-4′-*O*-methoxyisoflavanone (Akashi et al. 2003).

The substrate specificity of the HID (2-hydroxyisoflavanone dehydratase, chr5.CM0200.1460) is important for which isoflavonoids are produced by a particular legume species. In soybean, HID is able to act directly on the product of the IFS reaction, converting, for instance, 2,7,4′-trihydroxyisoflavanone into daidzein and leading to the formation of 4′-hydroxy type isoflavonoids. In *L. japonicus* and several other legume species, the 4′methoxyisoflavone formononetin is the intermediate in isoflavonoid production, resulting from HID working sequentially to HI4′OMT. While the HID of licorice (*Glycyrrhiza echinata*) showed substrate specificity for 4′-methoxy substrates, the HID from soybean was more promiscuous and able to dehydrate both 4′-methoxy and 4′-hydroxy substrates (Akashi et al. 2005). Expression of the broad substrate specificity HID from soybean in *L. japonicus* resulted in the production of the isoflavonoids daidzein and genistein, which are not normally observed in lotus, confirming that the endogenous HID from lotus prefers 4′-methoxy substrates (Shimamura et al. 2007). The isoflavone 2′-hydroxylase (I2′H) gene (chr4.CM0026.1220) encodes another cytochrome P450 enzyme of the pathway, this time a member of the CYP81E family (Shimada et al. 2000). I2′H converts formononetin to 7,2′-dihydroxy-4′-*O*-methoxyisoflavone (Fig. 14.2).

Additional potential genes of the vestitol biosynthetic pathway have been assigned on the basis of sequence similarity to the genes found in other legume species such as *Medicago sativa* (Shimada et al. 2007). As these enzymes from *L. japonicus* in the second half of the proposed vestitol biosynthetic pathway are largely uncharacterized, or even unidentified, we will only mention them briefly. Isoflavone reductase (IFR) was first identified in *M. sativa* and converts 7,2′-dihydroxy-4′-*O*-methoxyisoflavone into vestitone (Paiva et al. 1991). In *L. japonicus*, two putative IFR encoding genes, *IFR1* (chr2.CM0249.1380) and *IFR2* (chr2.CM0249.1390), are adjacent on chromosome 2. An IFR-like sequence which shows 60 % amino acid identity with IFR1 and IFR2 is encoded by gene chr2.CM0249.1420 (Shimada et al. 2007). Two genes highly homologous to vestitone reductase (VR) from *M. sativa* have also been identified in the lotus genome (chr1.CM1255.100 and chr1.CM1255.110). VR is a member of the short-chain dehydrogenase/reductase superfamily of enzymes and catalyzes the NADPH-dependent reduction of vestitone to 7,2′-dihydroxy-4′-methoxy-isoflavanol (DMI). The VR from *M. sativa* is highly stereochemically specific,

recognizing only (*3R*)-vestitone as substrate and not (3*S*)-vestitone (Shao et al. 2007). Amino acid residues likely involved in the stereochemical specificity of this VR, such as His227, are also present in the two *L. japonicus* enzymes, suggesting that (3*R*)-vestitone is indeed their substrate. Conversion of DMI to medicarpin, the main phytoalexin of *M. sativa*, is catalyzed by DMI dehydratase (DMID) which introduces an ether ring with the loss of water. This enzyme was partly purified from *M. sativa*, but to our knowledge has not been identified at the gene level (Guo et al. 1994). Conversion of medicarpin to vestitol requires the activity of additional members of the short-chain dehydrogenase/reductase family. From *L. japonicus* EST data, Akashi et al. (2006) were able to identify two pterocarpan reductases, PTR1 (chr3.CM0091.1150) and PTR2 (chr3.CM0091.1170) that converted medicarpin to vestitol with high-specific activity and enantiospecificity to (-)-medicarpin.

14.5 Triterpenoids and Triterpenoid Saponins

Triterpenoids are a diverse group of isoprene-derived specialized metabolites often functioning in plant chemical defense. When triterpenoids are glycosylated, they are referred to as triterpenoid saponins, and these amphiphatic molecules are well known as plant chemical defense compounds affecting membrane integrity (Augustin et al. 2011). Oxidosqualene cyclase (OSC) is the collective name for a group of enzymes that form the branching point between sterol and triterpene biosynthesis, by catalyzing alternative cyclization reactions of the common precursor 2,3-oxidosqualene. OSCs that function as cycloartenol synthases lead to the production of sterols, while other OSCs catalyze cyclization reactions that lead to the formation of different triterpenoid skeletons. Triterpenes suggested to occur in *L. japonicus* include β-amyrin and lupeol (Fig. 14.1) (Iturbe-Ormaetxe et al. 2003; Sawai et al. 2006a; Delis et al. 2011). Saponins that are β-amyrin derived, such as glycosides of soyasapogenol B, have been reported for the *Lotus* genus and other legumes (Tava et al. 2011; Golea et al. 2012).

Triterpenes and their biosynthesis have been investigated in several legume species (Seki et al. 2008; Tava et al. 2011; Carelli et al. 2011). In *L. japonicus*, present knowledge is largely restricted to the OSCs, the characteristic class of branch point enzymes. OSCs from *L. japonicus* were identified by cDNA cloning and isolated from genomic libraries and functionally characterized (Iturbe-Ormaetxe et al. 2003; Sawai et al. 2006a). Eight OSC genes were identified that grouped in two separate contigs located on either chromosome 2 (contig CM0373) or chromosome 3 (contig CM0292). Functional characterization in yeast established that *OSC1* (chr3.CM0292.500; also named *LjAMY1*) encodes a β-amyrin synthase, *OSC3* (chr2.CM0373.870) encodes a lupeol synthase, and *OSC5* (chr2.CM0373.850) encodes a cycloartenol synthase (Iturbe-Ormaetxe et al. 2003; Sawai et al. 2006a). *OSC3* is exclusively expressed in roots and nodules, and its expression was highly induced in plants infected with *Mesorhizobium loti* (Delis et al. 2011). *OSC6* (chr2.CM0373.810) and *OSC7* (chr2.CM0373.800) are closely related but differentially expressed, either in roots and nodules (*OSC6*), or in stem and leaves (*OSC7*). As it was originally thought that sterols in plants are biosynthesized from cycloartenol and those in fungi and animals from lanosterol, it was of particular interest that *OSC7* was able to complement a lanosterol synthase-deficient yeast mutant (Sawai et al. 2006b). Labeling studies in Arabidopsis have supported the existence of a plant biosynthetic pathway for phytosterols via lanosterol, but its contribution to membrane sterols was minor in comparison with the route via cycloartenol (Ohyama et al. 2009). It was suggested that the lanosterol pathway could contribute to the production of steroids in relation to plant defense. *LjAMY2/OSC8* (chr3.CM0292.40) is closely related to the β-amyrin synthase *OSC1*, but encodes a mixed function OSC, able to synthesize both β-amyrin and lupeol (Iturbe-Ormaetxe et al. 2003). A cDNA fragment designated *OSC4* was 99 % identical with *OSC8* but a full-length cDNA sequence could not be obtained, while the

sequence of a genomic clone containing it was unclear (Sawai et al. 2006a). In the present assembly build 2.5 of the *L. japonicus* genome, *OSC4* likely corresponds to the partial gene sequence chr3.CM0292.280, while *OSC2* is a pseudogene located between *OSC4* and *OSC1* (Sawai et al. 2006a). Identification of further triterpenoid biosynthetic genes in *L. japonicus* is proposed to benefit from the existence of a biosynthetic gene cluster in the region containing *OSC1* and *OSC8*, as will be discussed below.

14.6 Hydroxynitrile Glucosides and Cyanogenesis in *L. japonicus*

A prominent class of chemical defense compounds in legumes is that of the hydroxynitrile glucosides, of which the α-hydroxynitrile glucosides are commonly referred to as cyanogenic glycosides. Cyanogenic glucosides are part of a constitutive two-component plant chemical defense system. Upon tissue damage by herbivore feeding, cyanogenic glucosides are hydrolyzed by specific β-glucosidase enzymes, releasing the unstable α-hydroxynitrile aglycone which dissociates into toxic hydrogen cyanide (HCN) and a ketone or aldehyde. The experimental resources available in *L. japonicus* have promoted this species as the genetic model to study hydroxynitrile glucoside metabolism and its evolution, and a large number of *cyanogenesis deficient* mutants (*cyd*) were identified in a genetic screen (Takos et al. 2010, 2011).

L. japonicus contains the cyanogenic glucosides lotaustralin and linamarin (Forslund et al. 2004). In addition, it contains the non-cyanogenic γ-hydroxynitrile glucoside rhodiocyanoside A and the β-hydroxynitrile glucoside rhodiocyanoside D (Fig. 14.1). Lotaustralin and rhodiocyanoside A are the two major hydroxynitrile glucosides in the *L. japonicus* accessions MG20 and Gifu, but intraspecific variation in hydroxynitrile glucoside composition has been documented, for instance, the absence of rhodiocyanosides in the MG74 accession of *L. japonicus* (Bjarnholt et al. 2008). A role for rhodiocyanoside A in chemical defense has been more difficult to establish, but a toxic furanone product is formed following its hydrolysis (Bjarnholt and Møller 2008; Saito et al. 2012).

The biosynthesis of hydroxynitrile glucosides starts from amino acids with linamarin derived from valine and lotaustralin and the rhodiocyanosides derived from isoleucine. The biosynthesis of linamarin and lotaustralin involves three enzymatic steps, catalyzed by two cytochrome P450 enzymes and a glucosyltransferase acting sequentially. The close paralogs *CYP79D3* (chr3.CM0241.700) and *CYP79D4* (chr3.CM0241.310) encode the enzymes for the first biosynthetic step of the pathway, converting isoleucine or valine into their corresponding oximes. The genes differ in their expression pattern, with *CYP79D3* highly expressed in newly developing leaves and involved in leaf cyanogenesis, and *CYP79D4* expressed in roots (Forslund et al. 2004). The oxime produced from isoleucine is not only the first intermediate in the biosynthesis of lotaustralin but also in the synthesis of rhodiocyanoside A and D. The biosynthetic pathway for lotaustralin and rhodiocyanosides diverges at the second enzymatic step, the formation of the hydroxynitrile aglycones. Identification of the enzyme for this second step in *L. japonicus* was made possible by the observation that all biosynthetic genes for the synthesis of cyanogenic glucosides co-localized in the genome (Takos et al. 2011). We also observed such gene clusters for cyanogenic glucoside biosynthesis in the genomes of cassava (*Manihot esculenta*) and sorghum (*Sorghum bicolor*). Unlike in these last two cyanogenic species, the second step in *L. japonicus* did not involve a member of the CYP71 family of cytochrome P450 enzymes, but was catalyzed by the product of the *CYP736A2* gene (chr3.CM0241.850) present in the gene cluster. The gene encoding the final enzyme responsible for glucosylation and stabilization of the hydroxynitrile aglycone, the UDP-glucosyltransferase gene *UGT85K3* (chr3.CM0241.610), is also part of the gene cluster. Transient co-expressing of *CYP79D3*, *CYP736A2,* and *UGT85K3* in *Nicotiana benthamiana* resulted in the production of the cyanogenic glucosides

lotaustralin and linamarin (Takos et al. 2011). The biosynthetic pathway was further supported by genetic evidence. The absence of all hydroxynitrile glucosides in the *cyd1* mutant in *CYP79D3* demonstrated its inability to produce the shared oxime intermediates. A strong reduction in the cyanogenic glucosides lotaustralin and linamarin was observed in the *cyd4* mutant in *CYP736A2*. Due to the specificity of CYP736A2, no rhodiocyanosides were produced in the above-mentioned transient expression in tobacco, and rhodiocyanoside A and D levels were not reduced in the *cyd4* mutant. A genetic locus named *Rho* is responsible for the production of rhodiocyanosides but also contributes to cyanogenic glucoside production. *Rho* is closely linked to the gene cluster but falls outside the presently available sequence of the CM0241 contig (Takos et al. 2011). Biochemical evidence suggested the involvement of a cytochrome P450 enzyme to form 2-methyl-2-butenenitrile as an intermediate in the biosynthesis of rhodiocyanoside A, requiring a subsequent hydroxylation step to form the rhodiocyanoside A aglycone (Saito et al. 2012). The biosynthetic genes responsible for these steps are being identified.

14.7 Biosynthetic Gene Clusters in *L. Japonicus*

It has become apparent that the pathways for several classes of chemical defense compounds are clustering in plant genomes (Chu et al. 2011; Takos and Rook 2012). It is important to stress that these gene clusters are not just repeats of homologous genes, which are more common as illustrated by many of the gene families in phenylpropanoid metabolism described earlier, but consist of non-homologous genes encoding different types of enzymes of the same biosynthetic pathway. Such biosynthetic gene clusters have been described for other plant chemical defense compounds such as for DIBOA (2,4-dihydroxy-1,4-benzoxazin-3-one) in maize (Frey et al. 1997), the triterpenoids avenacin in oat, and thalianol and marneral in Arabidopsis (Qi et al. 2004; Field et al. 2008; Field et al. 2011), diterpenoid phytoalexins such as momilactones and phytocassanes in rice (Wilderman et al. 2004; Shimura et al. 2007; Swaminathan et al. 2009), the alkaloid noscapine in opium poppy (Winzer et al. 2012), and recently for terpene biosynthesis and steroidal glycoalkaloids in solanaceous species such as tomato (Matsuba et al. 2013; Itkin et al. 2013). For *L. japonicus,* we previously reported a gene cluster for the biosynthesis of cyanogenic glucosides (Fig. 14.3) consisting of the genes *CYP79D3*, *CYP79D4*, *CYP736A2,* and *UGT85K3*, also demonstrating that the analysis of a gene cluster can aid gene discovery (Takos et al. 2011).

Several gene clusters have so far been reported for triterpenoid biosynthetic pathways. Their positive identification benefits from the characteristic presence of oxidosqualene cyclase genes. OSCs catalyze the cyclization of the common precursor 2,3-oxidosqualene, representing the first committed step of the pathway (Augustin et al. 2011). Various modifications to the backbone structure are catalyzed by enzymes such as cytochrome P450 s which have a broad functional diversity. Complex glycosylation patterns of triterpenoid saponins are made by the activity of UDP-glycosyltransferases. Both these enzyme classes are encoded by large gene families, and assigning specific metabolic functions to individual genes requires biochemical and genetic evidence. The existence of a putative gene cluster would immediately suggest candidate genes for further analysis. The genomic region containing *OSC1* and *OSC8* on contig CM0292 seems to contain a triterpenoid gene cluster, and a functional characterization of one cytochrome P450 gene in this region was recently reported (Fig. 14.3; Krokida et al. 2013). The *CYP71D353* gene (chr3.CM0292.110) encodes a cytochrome P450 enzyme proposed to catalyze the oxidation of 20-hydroxylupeol to 20-hydroxybetulinic acid. The adjacent chr3.CM0292.120 encodes CYP88D5, and chr3.CM0292.180 is identical to *CYP88D4*. Members of the CYP88D subfamily were previously reported to be involved in triterpene biosynthesis, for example, CYP88D6 functions as a β-amyrin C11-oxidase in *Glycyrrhiza uralensis* (Seki et al. 2008). A complete and

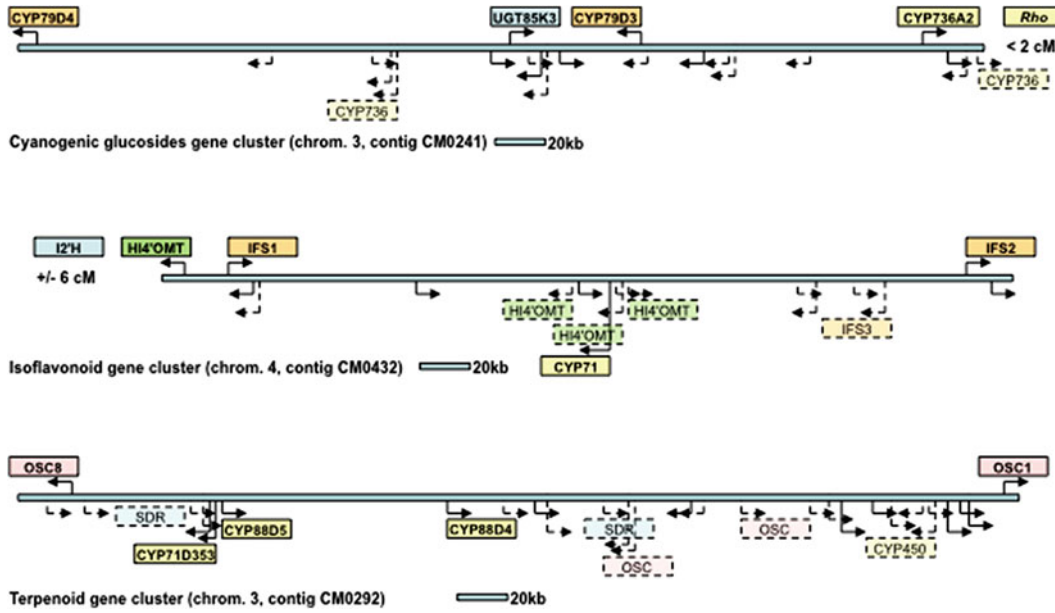


Fig. 14.3 Graphical representation of three regions in the genome of *L. japonicus* which contain biosynthetic gene clusters for cyanogenic glucosides, isoflavonoids, and terpenoids. *Arrows* represent gene sequences as annotated in the Kazusa DNA Research Institute genome assembly build 2.5. Functional genes with a confirmed role in the biosynthetic pathway are indicated above each contig bar. Functional genes with no confirmed role, or partial-/pseudogenes, are indicated below each contig bar. Gene annotations marked as partial- and/or pseudogenes are indicated with hatched *lines* and *boxes*. Biosynthetic genes of interest are named and indicated with *boxes*. The distance to genetic loci that fall outside the available sequence of each contig is indicated in cM. All contigs are drawn to the same scale

extended sequence of contig CM0292 may reveal further candidate genes in this region.

As the perceived absence of gene clusters for some classes of plant-specialized metabolites has been a question of interest, the existence of a gene cluster for a branch of the phenylpropanoid pathway directly involved in plant–microbe interactions is particularly noteworthy. That such a gene cluster for isoflavonoid biosynthesis exists in *L. japonicus*, with *IFS1* and *HI4'OMT* reportedly localized to TAC clone LjT24P23 (Shimada et al. 2007), was not yet recognized in the gene cluster literature. The encoded enzymes catalyze the first two steps in the isoflavonoid-specific branch of flavonoid metabolism. Since the original Shimada et al. (2007) publication making this observation, more sequence data and chromosomal locations for specific biosynthetic genes have become available. It is now clear that *IFS1* (chr4.CM0432.2900) and *HI4'OMT* (chr4.CM0432.2880) are adjacent genes spaced 18.5 kb apart. Moreover, *IFS2* (chr4.CM0432.3190) and the pseudogene *IFS3* (chr4.CM0432.3150) are now included on this same contig, and the distance between *IFS1* and *IFS2* is approximately 290 kb. This region also contains several pseudogenes related to *HI4'OMT*. Similarly, pseudogenes resembling *CYP736A2* were also observed in the cyanogenic glucoside gene cluster, suggesting a dynamic evolution of these regions involving gene duplications, possible functional divergence, and gene loss. The next step of the isoflavonoid pathway is mediated by the enzyme 2-hydroxyiosflavanone dehydratase (HID) which has been positioned on chromosome 5 (chr5.CM0200.1460) and is therefore not part of the gene cluster. The *CYP81E6* gene encoding isoflavone 2′-hydroxylase (I2′H, chr4.CM0026.1220) is, however, genetically linked to the partial cluster at a distance of approximately 6 cM and consequently

would experience a high degree of co-inheritance with the first two enzymes of the pathway. Such a degree of genetic linkage is also observed in the biosynthetic pathways of DIBOA in maize, which spans a distance of 6 cM (Frey et al. 1997), and in the avenacin pathway in oat where a locus for glucosylation is at 3.6 cM from the core gene cluster (Qi et al. 2004). The presence of an uncharacterized cytochrome P450 gene in the region between *IFS1* and *IFS2* should also be mentioned, as a yet unknown functional role in isoflavonoid metabolism cannot be excluded.

14.8 Gene Cluster Formation

It was noted that the terpenoid biosynthetic pathways for avenacin and thalianol, and by implication their biosynthetic gene clusters, had independently evolved from primary metabolism in monocot and dicot species (Field and Osbourn 2008). Also, the three gene clusters for the biosynthesis of cyanogenic glucosides in the genomes of cassava, sorghum, and *L. japonicus* evolved independently (Takos et al. 2011). This suggests that there is a general evolutionary mechanism promoting this remarkable genome organization. We recently proposed that a simple but comprehensive explanation for gene cluster formation can be based on a genetic principle first described by Ronald A. Fisher in his 1930 book "*The Genetical Theory of Natural Selection*" (Fisher 1930; Takos and Rook 2012). Under certain conditions interacting genetic loci on the same chromosome will tend to reduce the recombination frequency between them, resulting in ever closer genetic linkage. The situation under which this occurs requires that alternative gene or allele combinations are maintained within the genome of a species. Each combination of beneficially interacting genes provides a selective advantage, but the effect of each combination is also counteracted by the effect of the alternative gene combinations. Although gene translocations are rare and their direction random, the emergence of a modified chromosome with a reduced recombination frequency between beneficially interacting loci will replace the original one in a population by contributing more of the fitter genotypes to each subsequent generation. A repeated process of gene translocation and selection for reduced recombination leads to ever closer physical linkage over evolutionary time. Such gene translocations leading to the formation of eukaryotic gene clusters have been documented in fungi (Wong and Wolfe 2005; Proctor et al. 2009; Slot and Rokas 2010). We have argued that the Fisher model for selection for reduced recombination applies to the evolution of sex chromosomes in animals and to gene clusters for flower dimorphism and self-incompatibility loci in plants (Takos and Rook 2012). Sex chromosomes evolve under sexually antagonistic selection by selecting for reduced recombination between a sex-determining locus and loci with alleles that are beneficial for the corresponding sex (Bergero and Charlesworth 2008). Self-incompatibility gene clusters in plant sexual reproduction minimally consist of a receptor and its corresponding protein ligand. Such self-incompatibility loci are subjected to frequency-dependent selection, a form of balancing selection, as the rare genotypes in a population have a reproductive advantage. The biosynthetic pathways for many plant chemical defense compounds are also under antagonistic selection pressures, or phrased in population genetic terms under balancing selection with fitness varying in time and space. For example, cyanogenic glucosides in *Phaseolus lunatus* provide protection against herbivores but make the plant more susceptible to fungal infection (Ballhorn et al. 2010). Consequently, chemical defense polymorphisms, including the presence/absence of a biosynthetic pathway, are dynamically maintained in natural populations. Such chemical defense polymorphisms represent the competing beneficial allele combinations that promote gene cluster formation. The study of such natural variation in plant chemical defense on the genomic and ecological level, for which *L. japonicus* provides a very suitable model system, will further contribute to our understanding of how ecological interactions shape eukaryotic genomes.

Acknowledgements We acknowledge financial support for our research program from The Danish Council for Independent Research | Natural Sciences.

References

Abrahams S, Lee E, Walker AR et al (2003) The Arabidopsis TDS4 gene encodes leucoanthocyanidin dioxygenase (LDOX) and is essential for proanthocyanidin synthesis and vacuole development. Plant J 35:624–636

Akashi T, Sawada Y, Shimada N et al (2003) cDNA cloning and biochemical characterization of *S*-adenosyl-L-methionine: 2,7,4′-trihydroxyisoflavanone 4′-*O*-methyltransferase, a critical enzyme of the legume isoflavonoid phytoalexin pathway. Plant Cell Physiol 44:103–112

Akashi T, Aoki T, Ayabe S (2005) Molecular and biochemical characterization of 2-hydroxyisoflavanone dehydratase. Involvement of carboxylesterase-like proteins in leguminous isoflavone biosynthesis. Plant Physiol 137:882–891

Akashi T, Koshimizu S, Aoki T et al (2006) Identification of cDNAs encoding pterocarpan reductase involved in isoflavan phytoalexin biosynthesis in *Lotus japonicus* by EST mining. FEBS Lett 580:5666–5670

Augustin JM, Kuzina V, Andersen et al (2011) Molecular activities, biosynthesis and evolution of triterpenoid saponins. Phytochemistry 72:435–457

Ballhorn DJ, Pietrowski A, Lieberei R (2010) Direct trade-off between cyanogenesis and resistance to a fungal pathogen in lima bean (*Phaseolus lunatus* L.). J Ecol 98:226–236

Barbehenn RV, Constabel CP (2011) Tannins in plant-herbivore interactions. Phytochemistry 72:1551–1565

Bergero R, Charlesworth D (2008) The evolution of restricted recombination in sex chromosomes. Trends Ecol Evol 24:94–102

Bjarnholt N, Rook F, Motawia MS et al (2008) Diversification of an ancient theme: hydroxynitrile glucosides. Phytochemistry 69:1507–1516

Bjarnholt N, Møller BL (2008) Hydroxynitrile glucosides. Phytochemistry 69:1947–1961

Bogs J, Jaffe FW, Takos AM et al (2007) The grapevine transcription factor VvMYBPA1 regulates proanthocyanidin synthesis during fruit development. Plant Physiol 143:1347–1361

Bonde MR, Millar RL, Ingham JL (1973) Induction and identification of sativan and vestitol as two phytoalexins from *Lotus corniculatus*. Phytochemistry 12:2957–2959

Carelli M, Biazzi E, Panara F et al (2011) *Medicago truncatula* CYP716A12 is a multifunctional oxidase involved in the biosynthesis of hemolytic saponins. Plant Cell 23:3070–3081

Chu HY, Wegel E, Osbourn A (2011) From hormones to secondary metabolism: the emergence of metabolic gene clusters in plants. Plant J 66:66–79

Cooper JE (2007) Early interactions between legumes and rhizobia: disclosing complexity in a molecular dialogue. J Appl Microbiol 103:1355–1365

Cornwell T, Cohick W, Raskin I (2004) Dietary phytoestrogens and health. Phytochemistry 65:995–1016

Debeaujon I, Peeters AJM, Leon-Kloosterziel KM et al (2001) The *TRANSPARENT TESTA 12* gene of Arabidopsis encodes a multidrug secondary transporter-like protein required for flavonoid sequestration in vacuoles of the seed coat endothelium. Plant Cell 13:853–871

Delis C, Krokida A, Georgiou S et al (2011) Role of lupeol synthase in *Lotus japonicus* nodule formation. New Phytol 189:335–346

Dixon RA, Liu C, Jun JH (2013) Metabolic engineering of anthocyanins and condensed tannins in plants. Curr Opin Biotech 24:329–335

Field B, Osbourn AE (2008) Metabolic diversification—independent assembly of operon-like gene clusters in different plants. Science 320:543–547

Field B, Fiston-Lavier A-S, Kemen A et al (2011) Formation of plant metabolic gene clusters within dynamic chromosomal regions. Proc Natl Acad Sci USA 108:16116–16121

Fisher RA (1930) The genetical theory of natural selection. Oxford University Press, Oxford

Forslund K, Morant M, Jørgensen B et al (2004) Biosynthesis of the nitrile glucosides rhodiocyanoside A and D and the cyanogenic glucosides lotaustralin and linamarin in *Lotus japonicus*. Plant Physiol 135:71–84

Frey M, Chomet P, Glawischnig E et al (1997) Analysis of a chemical plant defense mechanism in grasses. Science 277:696–699

Golea L, Haba H, Lavaud C et al (2012) Chemical constituents from *Lotus pusillus* Medik. Biochem Syst Ecol 45:12–15

Gruber M, Skadhauge B, Yu M et al (2008) Variation in morphology, plant habit, proanthocyanidins, and flavonoids within a *Lotus* germplasm collection. Can J Plant Sci 88:121–132

Guo L, Dixon RA, Paiva NL (1994) Conversion of vestitone to medicarpin in alfalfa (*Medicago sativa* L.) is catalyzed by two independent enzymes. Identification, purification, and characterization of vestitone reductase and 7,2′-dihydroxy-4′-methoxyisoflavanol dehydratase. J Biol Chem 269:22372–22378

Hassan S, Mathesius U (2012) The role of flavonoids in root-rhizosphere signalling: opportunities and challenges for improving plant-microbe interactions. J Exp Bot 63:3429–3444

Itkin M, Heinig U, Tzfadia O et al (2013) Biosynthesis of antinutritional alkaloids in solanaceous crops is mediated by clustered genes. Science 341:175–179

Iturbe-Ormaetxe I, Haralampidis K, Papadopoulou K et al (2003) Molecular cloning and characterization of triterpene synthases from *Medicago truncatula* and *Lotus japonicus*. Plant Mol Biol 51:731–743

Khersonsky O, Tawfik DS (2010) Enzyme promiscuity: a mechanistic and evolutionary perspective. Annu Rev Biochem 79:471–505

Kosslak RM, Bookland R, Barkei J et al (1987) Induction of *Bradyrhizobium japonicum* common *nod* genes by isoflavones isolated from *Glycine max*. Proc Natl Acad Sci USA 84:7428–7432

Krokida A, Delis C, Geisler K et al (2013) A metabolic gene cluster in *Lotus japonicus* discloses novel enzyme functions and products in triterpene biosynthesis. New Phytol 200:675–690

Lapcik O (2007) Isoflavonoids in non-leguminous taxa: a rarity or a rule? Phytochemistry 68:2909–2916

Masunaka A, Hyakumachi M, Takenaka S (2011) Plant growth-promoting fungus, *Trichoderma koningi* suppresses isoflavonoid phytoalexin vestitol production for colonization on/in the roots of *Lotus japonicus*. Microbes Environ 26:128–134

Matsuba Y, Nguyen TTH, Weigert K (2013) Evolution of a complex locus for terpene biosynthesis in *Solanum*. Plant Cell 25:2022–2036

Mustonen EA, Jokela T, Saastamoinen I et al (2006) High serum *S*-equol content in red clover fed ewes: the classical endocrine disrupter is a single enantiomer. Environ Chem Lett 3:154–159

Ohyama K, Suzuki M, Kikuchi J et al (2009) Dual biosynthetic pathways to phytosterol via cycloartenol and lanosterol in *Arabidopsis*. Proc Natl Acad Sci USA 106:725–730

Paiva NL, Edwards R, Sun Y et al (1991) Stress responses in alfalfa (*Medicago sativa* L.) 11. Molecular cloning and expression of alfalfa isoflavone reductase, a key enzyme of isoflavonoid phytoalexin biosynthesis. Plant Mol Biol 17:653–667

Pang Y, Peel GJ, Wright E et al (2007) Early steps in proanthocyanidin biosynthesis in the model legume *Medicago truncatula*. Plant Physiol 145:601–615

Pang Y, Peel GJ, Sharma SB et al (2008) A transcript profiling approach reveals an epicatechin-specific glucosyltransferase expressed in the seed coat of *Medicago truncatula*. Proc Natl Acad Sci USA 105:14210–14215

Paolocci F, Robbins MP, Madeo L et al (2007) Ectopic expression of a *Basic Helix-Loop-Helix* gene transactivates parallel pathways of proanthocyanidin biosynthesis. Structure, expression analysis, and genetic control of *Leucoanthocyanidin 4-Reductase* and *Anthocyanidin Reductase* genes in *Lotus corniculatus*. Plant Physiol 143:504–516

Patra AK, Saxena J (2011) Exploitation of dietary tannins to improve rumen metabolism and ruminant nutrition. J Sci Food Agric 91:24–37

Proctor RH, McCormick SP, Alexander NJ et al (2009) Evidence that a secondary metabolic biosynthetic gene cluster has grown by gene relocation during evolution of the filamentous fungus *Fusarium*. Mol Microbiol 74:1128–1142

Qi X, Bakht S, Leggett M et al (2004) A gene cluster for secondary metabolism in oat: implications for the evolution of metabolic diversity in plants. Proc Natl Acad Sci USA 101:8233–8238

Quinones M, Miguel M, Aleixandre A (2013) Beneficial effects of polyphenols on cardiovascular disease. Pharmacol Res 68:125–131

Reynaud J, Guilet D, Terreux R et al (2005) Isoflavonoids in non-leguminous families: an update. Nat Prod Rep 22:504–515

Rispail N, Hauck B, Bartholomew B et al (2010) Secondary metabolite profiling of the model legume *Lotus japonicus* during its symbiotic interaction with *Mesorhizobium loti*. Symbiosis 50:119–128

Saito S, Motawia MS, Olsen CE et al (2012) Biosynthesis of rhodiocyanosides in *Lotus japonicus*: Rhodiocyanoside A is synthesized from (Z)-2-methylbutanaloxime via 2-methyl-2-butenenitrile. Phytochemistry 77:260–267

Sawai S, Shindo T, Sato S et al (2006a) Functional and structural analysis of genes encoding oxidosqualene cyclases of *Lotus japonicus*. Plant Sci 170:247–257

Sawai S, Akashi T, Sakurai N et al (2006b) Plant lanosterol synthase : divergence of the sterol and triterpene biosynthetic pathways in eukaryotes. Plant Cell Physiol 47:673–677

Seki H, Ohyama K, Sawai S et al (2008) Licorice β-amyrin 11-oxidase, a cytochrome P450 with a key role in the biosynthesis of the triterpene sweetener glycyrrhizin. Proc Natl Acad Sci USA 105:14204–14209

Shao H, Dixon RA, Wang X (2007) Crystal structure if vestitone reductase from alfalfa (*Medicago sativa* L.). J Mol Biol 369:265–276

Shimada N, Akashi T, Aoki T et al (2000) Induction of isoflavonoid pathway in the model legume *Lotus japonicus*: molecular characterization of enzymes involved in phytoalexin biosynthesis. Plant Sci 160:37–47

Shimada N, Aoki T, Sato S et al (2003) A cluster of genes encodes the two types of chalcone isomerase involved in the biosynthesis of general flavonoids and legume-specific 5-deoxy(iso)flavonoids in *Lotus japonicus*. Plant Physiol 131:941–951

Shimada N, Sasaki R, Sato S et al (2005) A comprehensive analysis of six dihydroflavonol 4-reductases encoded by a gene cluster of the *Lotus japonicus* genome. J Exp Bot 56:2573–2585

Shimada N, Sato S, Akashi T et al (2007) Genome-wide analyses of the structural gene families involved in the legume-specific 5-deoxyisoflavonoid biosynthesis of *Lotus japonicus*. DNA Res 14:25–36

Shimamura M, Akashi T, Sakurai N et al (2007) 2-Hydroxyisoflavanone dehydratase is a critical determinant of isoflavone productivity in hairy root cultures of *Lotus japonicus*. Plant Cell Physiol 48:1652–1657

Shimura K, Okada A, Okada K et al (2007) Identification of a biosynthetic gene cluster in rice for momilactones. J Biol Chem 282:34013–34018

Sivakumaran S, Rumball W, Lane GA (2006) Variation of proanthocyanidins in *Lotus* species. J Chem Ecol 32:1797–1816

Skadhauge B, Gruber MY, Thomsen KK et al (1997) Leucocyanidin reductase activity and accumulation of proanthocyanidins in developing legume tissue. Am J Bot 84:494–503

Slot JC, Rokas A (2010) Multiple *GAL* pathway gene clusters evolved independently and by different mechanisms in fungi. Proc Natl Acad Sci USA 107:10136–10141

Swaminathan S, Morrone D, Wang Q et al (2009) CYP76M7 is an *ent*-cassadiene C11α-hydroxylase defining a second multifunctional diterpenoid biosynthetic gene cluster in rice. Plant Cell 21:3315–3325

Takos M, Lai D, Mikkelsen L et al (2010) Genetic screening identifies cyanogenesis-deficient mutants of *Lotus japonicus* and reveals enzymatic specificity in hydroxynitrile glucoside metabolism. Plant Cell 22:1605–1619

Takos AM, Knudsen C, Lai D et al (2011) Genomic clustering of cyanogenic glucoside biosynthetic genes aids their identification in *Lotus japonicus* and suggests the repeated evolution of this chemical defence pathway. Plant J 68:273–286

Takos AM, Rook F (2012) Why biosynthetic genes for chemical defense compounds cluster. Trends Plant Sci 17:383–388

Tava A, Scotti C, Avato P (2011) Biosynthesis of saponins in the genus *Medicago*. Phytochem Rev 10:459–469

Ueda H, Sugimoto Y (2010) Vestitol as a chemical barrier against intrusion of parasitic plant *Striga hermonthica* into *Lotus japonicus* roots. Biosci Biotechnol Biochem 74:1662–1667

Wang H, Cui Y, Zhao C (2010) Flavonoids of the genus *Iris* (Iridaceae). Mini-Rev Med Chem 10:643–661

Wilderman PR, Xu M, Jin Y et al (2004) Identification of *syn*-pimara-7,15-diene synthase reveals functional clustering of terpene synthases involved in rice phytoalexin/allelochemical biosynthesis. Plant Physiol 135:2098–2105

Winzer T, Gazda V, He Z et al (2012) A *Papaver somniferum* 10-gene cluster for synthesis of the anticancer alkaloid noscapine. Science 336:1704–1708

Wong S, Wolfe KH (2005) Birth of a metabolic gene cluster in yeast by adaptive gene relocation. Nat Genet 37:777–782

Xie D-Y, Sharma SB, Paiva NL et al (2003) Role of anthocyanidin reductase, encoded by *BANYULS* in plant flavonoid biosynthesis. Science 299:396–399

Yoshida K, Iwasaki R, Shimada N et al (2010) Transcriptional control of the dihydroflavonol 4-reductase multigene family in *Lotus japonicus*. J Plant Res 123:801–805

Zhao J, Dixon RA (2009) MATE transporter facilitate vacuolar uptake of epicatechin 3′-*O*-glucoside for proanthocyanidin biosynthesis in *Medicago truncatula* and *Arabidopsis*. Plant Cell 21:2323–2340

Zhao J, Pang Y, Dixon RA (2010) The mysteries of proanthocyanidin transport and polymerization. Plant Physiol 153:437–443

Genes Involved in Pathogenesis and Defense Responses

15

Tomomi Nakagawa, Shin Okazaki, and Naoto Shibuya

Abstract

The legume family includes important grain, forage, and agroforestry species. One of the major constraints for cultivation of these legumes is obviously production loss by disease (Graham and Vance in Plant Physiol 131(3):872–877, 2003). Thus, *Lotus japonicus* is utilized not only as a symbiotic model but also as a research platform for studying serious diseases in legumes. However, most of our knowledge about the defense mechanism in *L. japonicus* comes from the study of legume–rhizobia symbiosis. In this section, we mainly focus on the regulation of defense responses in host symbiotic process, the pathogenic aspect of symbiotic microbial partners, and then illustrate the *Lotus* pathogens.

15.1 General Introduction of Plant Defense Mechanism

In the natural environment, land plants are constantly exposed to the threats of pathogenic microbes. To defend themselves from the infection of these hostile microbes, plants have evolved two layers of inducible defense mechanisms (Jones and Dangl 2006). The first layer of defense mechanisms relies on the detection of microbe-associated molecular patterns (MAMPs), which are common components of various microbes but not present in higher plants/animals, and activates basal defense responses such as generation of reactive oxygen species (ROS), induction of defense genes, and production of antimicrobial compounds such as phytoalexins. Rapid activation of these basal defense responses is generally sufficient to block microbial invasions. The common and invariant nature of the MAMP molecules such as fungal cell wall chitin and bacterial flagellin for various microbes also makes MAMPs-triggered immunity (MTI) effective for the prevention of most pathogenic infections (Monaghan and Zipfel 2012; Segonzac and Zipfel 2011). Recent studies revealed that microbial strategies for successful infection often intend to avoid or perturb the plant recognition for MAMPs (Dodds and Rathjen 2010). For example, many Gram-negative bacteria inject a repertoire of "virulence effector" proteins around or into the host cells through type III or IV secretion system (T3SS or T4SS) and manipulate host immunity to enhance their own pathogenicity (Hueck 1998; Christie and Vogel 2000). To cope

T. Nakagawa (✉) · S. Okazaki · N. Shibuya
Department of Life Sciences, School of Agriculture, Meiji University, 1-1-1 Higashi-Mita, Tama-Ku, Kawasaki, Kanagawa 214-8571, Japan
e-mail: tf11001@meiji.ac.jp

S. Tabata and J. Stougaard (eds.), *The Lotus japonicus Genome*, Compendium of Plant Genomes,
DOI 10.1007/978-3-662-44270-8_15,

with these advanced microbial strategies using effectors, plants evolved the second layer of defense system. Plant nucleotide-binding leucine-rich repeat (NB-LRR) receptors constantly monitor either the existence of microbial effectors or the condition of key signaling components of MTI. Detection of effectors or abnormal modification of signaling components targeted by effectors results in the induction of robust defense responses including localized programmed cell death at infection sites, called effector-triggered immunity (ETI). Thus, these NB-LRR receptors, which have been known as a major group of "resistance proteins" (R proteins), play crucial roles for the specific recognition of pathogens (Heidrich et al. 2012). These highly sophisticated defense mechanisms effectively block the vast majority of hostile microbes; however, a small number of successful pathogens can overcome both of the two layers of the defense system and cause serious damage to the plant growth. Therefore, nature is a grueling battlefield between plants and microbes and it is surprising that mutual plant–microbe symbioses are achieved in such an environment.

15.2 MAMPs-Triggered Immunity in *L. japonicus*

Although legumes accept the infection of rhizobia and establish an endosymbiosis, these plants are assumed to be equipped with the defense mechanism, protecting them from the pathogen's attacks. In *L. japonicus*, treatments with MAMPs, such as chitin oligosaccharides or flagellin epitope, flg22, upregulate the expression of defense-related genes including PR gene homologs, peroxidases, chitinases, ERF and WRKY transcription factors, and the genes involved in the biosynthesis of pterocarpans, which are known to be legume-specific phytoalexins (Shimada et al. 2007; Nakagawa et al. 2011). These results clearly indicate the existence of MTI in legumes. Interestingly, rhizobial symbiotic signal molecules, Nod factors (NFs), consist of a chitin oligosaccharide backbone, in which the non-reducing end is *N*-acylated and the reducing end is decorated with various molecules (Cullimore et al. 2001). Thus, the structure of NFs is closely related to a typical MAMP, chitin oligosaccharide; however, their physiological effects are opposite, i.e., friendly acceptance or rejection by host plants.

Recognition of MAMPs is the crucial step for MTI, and plants have evolved the corresponding receptors for each MAMP molecule. For example, flg22 and chitin oligosaccharides are perceived by FLS2 and CERK1 in *A. thaliana*, respectively (Gomez-Gomez and Boller 2000; Miya et al. 2007). A homologous gene of FLS2, *LjFLS2*, is also found in *L. japonicus*. Expression of *LjFLS2* is observed in both leaves and roots but decreased in nodules (Lopez-Gomez et al. 2012). The latter observation may suggest the presence of downregulation of defense mechanism in nodules to favor the entry of symbiotic partners. On the other hand, the peptide sequences of flg22 in *Rhizobium* or a highly related pathogen, *Agrobacterium*, are very different and the *Rhizobium* peptide could not trigger MTI in both host and non-host plants (Felix et al. 1999; Lopez-Gomez et al. 2012). These results suggested the presence of reciprocal adaptation between host and symbiotic partner.

NF receptor, NFR1, has a domain structure composed of an extracellular LysM domain, a single-pass transmembrane domain and an intracellular kinase domain, which is exactly the same as that of CERK1 in *Arabidopsis* (Radutoiu et al. 2003; Miya et al. 2007; Shimizu et al. 2010). Among the LysM domain-containing receptor-like kinases found in *Arabidopsis* genome, CERK1 is a single best match homolog of NFR1 (Zhang et al. 2007; Zhu et al. 2006). In addition, genomic structures around *Arabidopsis* CERK1 (AtCERK1) and *Lotus* NFR1 showed limited but significant synteny, indicating that these genes are the descendants of a common ancestor (Zhu et al. 2006). Indeed, NFs not only activate symbiosis genes but also transiently activate defense-related genes through NFR1 in *L. japonicus* (Nakagawa et al. 2011). On the other hand, the kinase domain of AtCERK1 is

functionally distinct from that of NFR1 because the chimeric gene consisting of the LysM domain of NFR1 and the kinase domain of AtCERK1 did not rescue the symbiotic defect of *Lotus nfr1* mutants (Nakagawa et al. 2011). Surprisingly, however, only three consecutive amino acid substitutions in the AtCERK1 kinase domain conferred the symbiotic activity to the chimeric receptor. In addition to the structural similarity between chitin oligosaccharides and NFs, these findings suggest the close evolutionary relationships between defense and symbiosis.

15.3 Pathogenic Aspect of the Symbiotic Partner

Rhizobia are beneficial to the host plants; however, their family, *Rhizobiaceae*, also includes a pathogenic microbe, *Agrobacterium*. The pathogenic strains of *Rhizobiaceae* carry virulence genes as plasmids named pTi or pRi that are required for tumorigenic or hairy root-inducing symptoms, respectively. On the other hand, in symbiotic strains, the genera *Rhizobium* and *Sinorhizobium* carry plasmids (pSym) containing *nod* and *nif* genes that are indispensable for the nodulation and nitrogen fixation (see Chap. 6). Therefore, these plasmids are the critical determinants for pathogenic or symbiotic phenotypes. Interestingly, these plasmids are transferred among pathogenic, nonpathogenic, and symbiotic species (Chen et al. 2001; Moulin et al. 2001; Vandamme et al. 2002). In addition, natural strains carrying pTi, pRi, and pSym together were found and confirmed to have the ability for inducing tumors, hairy roots, and nodules (Velazquez et al. 2005).

As mentioned in Chap. 6, rhizobia also carry a set of genes for the T3/4SS in the symbiosis islands or sym plasmids. The genes encoding T3SSs have been found not in all, but in some rhizobia including *Rhizobium* sp. NGR234, *Sinorhizobium fredii*, *Bradyrhizobium japonicum*, *B. elkanii,* and *Mesorhizobium loti* MAFF303099, whereas *M. loti* R7A possesses T4SS (Deakin and Broughton 2009). A distinctive feature of rhizobial secretion systems is that the expression is tightly regulated by flavonoids derived from the host legume. This unique regulation is achieved by the transcriptional regulator NodD which senses the host flavonoids and activates genes downstream of special promoters called *nod* boxes (Spaink 1995) (see Chap. 6). The *nod* boxes are found not only in the upstream of genes for Nod-factor synthesis (e.g., *nod, noe, and nol*), but also in the promoter region of the transcriptional activator *ttsI*. TtsI activates transcription of T3SS genes by binding to specific *cis*-elements (*tts* boxes) found upstream of the rhizobial T3SS genes (Krause et al. 2002). Therefore, the activation of T3SS appears to be coordinated with the biosynthesis of Nod factors, which initiate nodule organogenesis in host legumes.

15.4 Physiological Roles of Rhizobial Type III Secretion System and Effectors

Depending on the host plant species, rhizobial T3SS approaches result in different effects. In *M. loti* MAFF303099, for example, deletion of *tts* genes led to a reduced nodule number with *Lotus corniculatus* subsp. *frondosus* (Okazaki et al. 2010). Likewise, deletion of *B. japonicum tts* genes caused a reduction in nodule number on *Macroptilium atropurpureum* and delayed nodulation with soybean (Krause et al. 2002). Similar nodulation enhancement has also been observed in other T3SS-containing rhizobia. The infection and nodulation processes that the rhizobial T3SS facilitates remain to be elucidated. However, transcriptional studies have shown that rhizobial T3SS is expressed at all stages of infection. The T3SS of *B. japonicum* USDA110 was expressed in infection threads and developing nodules of *Glycine max* (Zehner et al. 2008), and expression of several T3SS genes of *Rhizobium* sp. NGR234 was detected in mature nodules of *Cajanus cajan* and *V. unguiculata* (Perret et al. 1999; Viprey et al. 1998).

In pathogenic microbes, T3/4SS is a delivery system of virulence agents, i.e., effector proteins. In rhizobia, these secreted proteins are designated

as nodulation outer proteins (Nops) (Fig. 15.1). *Rhizobium* sp. NGR234 secretes at least eight Nops via the T3SS that are either components of the T3SS-dependent pilus (NopA and NopB), the putative translocon that forms pores in the plant plasma membrane (NopX), or putative effectors (NopL, NopP, NopJ, NopM, and NopT) (Kambara et al. 2009; Skorpil et al. 2005; Saad et al. 2005; Ausmees et al. 2004; Viprey et al. 1998; Deakin et al. 2005). Among the putative rhizobial effectors, NopL and NopP appear to be specific to rhizobia. Other Nops homologs are widespread in T3SS-containing pathogens. NopJ, for example, belongs to the YopJ family of ubiquitin-like protein proteases found in many plant and animal pathogens. This family includes numerous Avr proteins such as AvrBsT of *Xanthomonas campestris* pv. *vesicatoria*, a pathogen of *Vigna* and soybean (Marie et al. 2003). NopM is homologous to YopM of *Yersinia pestis*. YopM is a nuclear-targeting protein that modulates phosphorylation signaling cascades and diminishes defense responses in animal cells (Viboud and Bliska 2005). NopT has homology with the Avr protein AvrPphB of the phytopathogen *Pseudomonas syringae* as well as YopT of *Yersinia* spp. (Shao et al. 2003). Proteins of this family are cysteine proteases. The proteolytic activity of AvrPphB is essential for autoproteolytic cleavage of an AvrPphB precursor as well as for eliciting the hypersensitive response (HR) in plants. The effector repertoires vary in size and composition between strains, which might determine the host-dependent effect of T3SS.

Host legume gene expression affected by rhizobial T3SS has recently been elucidated. Sánchez et al. reported that the relative expression levels of gibberellin-2 oxidase (GA-2 OXIDASE), flavanone 3-hydroxylase (F3H), and nodule inception (NIN) in *L. japnonicus* were significantly reduced in roots inoculated with the T3SS mutant compared to those inoculated with the wild-type strain of *M. loti* MAFF303099 (Sánchez et al. 2009). The GA-2 OXIDASE and F3H genes are strongly induced in mature nodules (Kouchi et al. 2010), while the NIN gene, after induction by Nod factors (Schauser et al. 1999), maintains and increases its expression during nodulation. Bartsev et al. (2004) reported that ectopic expression of rhizobial effector gene *nopL* in *L. japonicus* suppressed the expression of pathogen-related (PR) genes, chitinase, and glucanase (Bartsev et al. 2004). Biochemical studies revealed that NopL is a substrate for plant

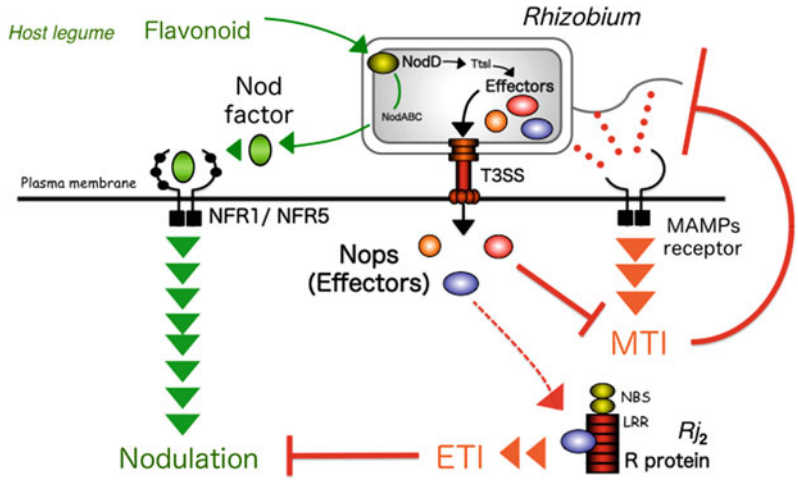


Fig. 15.1 Nodulation signaling and modulation by the rhizobial type III secretion system. A host plant-derived flavonoid induces the production of Nod factors (NFs) in rhizobia. Recognition of NFs by NF receptors (NFRs) triggers a signaling cascade leading to nodulation. The flavonoid also induces rhizobial T3SS, which injects effector proteins into host cells. One effector modulates nodulation signaling toward nodulation, whereas another is recognized by the host defense system, which is capable of aborting the nodulation process

MAP kinase, suggesting that NopL modulates signaling pathways that culminate in the activation of PR genes. Therefore, these effector proteins should play positive roles for the rhizobial infection.

In plant–pathogen interactions, a secretion of the effecter proteins to incompatible host plants is perceived by NB-LRR receptors and results in the induction of ETI and thus has a detrimental effect to the pathogenicity. From these results, effectors were sometimes called "avirulence gene" (Avr). Similar results were also observed in rhizobial infections. In *Lotus halophilus*, for example, the wild-type strain of *M. loti* MAFF303099 almost completely retards nodule formation, while the number of nodules drastically increases when inoculated with a T3SS null mutant. In this interaction, a gene of *M. loti* mlr6361 was identified as the major determinant of the nodulation restriction. The predicted gene product of mlr6361 is a protein of 3,056 amino acids containing 15 repetitions of a sequence motif of 40–45 residues and a shikimate kinase-like domain at its carboxyl terminus. Homologs with similar repeat sequences are present in the hypersensitive response and pathogenicity (Hrp) regions of several plant pathogens, including strains of *P. syringae*, *Ralstonia solanacearum*, and *Xanthomonas* species. These results suggest that *L. halophilus* recognizes Mlr6361 as potentially pathogen derived, and subsequently halts the infection process. Similar negative (Avr-like) effects on the nodulation of legumes have been reported in several other cases such as NopJ and NopT of NGR234 in *P. vulgaris* and *C. juncea*, respectively. In most incompatible interactions, nodulation of the T3SS-harboring strain is completely abolished, indicative of the involvement of a rapid and robust defense reaction, which is reminiscent of ETI. Considering the homology of Nops to Avr proteins of phytopathogens, it is tempting to speculate that these Nops are recognized by the leguminous R genes.

Although the detailed analyses of *Lotus* R genes are missing, the symbiotic function of R gene has recently been elucidated in soybean. In general, soybeans establish a mutualistic symbiosis with wide range of rhizobia including genus *Bradyrhizobium* or *Sinorhizobium*. However, some soybean cultivars belonging to *Rj2*, *Rj3*, *Rj4,* or *Rfg1* restrict the symbiotic partner to the specific strains of *Bradyrhizobium* (Hayashi et al. 2012) and *Sinorhizobium*. Among them, soybean cultivar Hardee, which carries *Rj2,* formed no nodules in the inoculation of USDA122, while other strains such as USDA110 induced formation of fully matured nodules. Yang et al. (2010) revealed that the soybean *Rj2* encodes the Toll-interleukin receptor/nucleotide-binding site/leucine-rich repeat (TIR-NBS-LRR) class of plant R protein. In addition, T3SS mutants of USDA122 that abolish the secretion of typical effector proteins gained the ability to nodulate Hardee (Tsukui et al. 2013). These results suggest the Avr–R interaction governs host specificity of nodulation between rhizobia and legumes including *Lotus* spp. The involvement of legume R genes in the control of genotype-specific nodulation reveals a common recognition mechanism underlying symbiotic and pathogenic host–bacteria interactions.

15.5 Candidates of the Model Pathogen in *L. japonicus*

Although there are some reports of fungal and bacterial diseases (Sisterna and Lori 2005; Alippi 2005), only a few inoculation systems were reported for *L. japonicus* (Schumpp et al. 2007; Takeuchi et al. 2007). For example, Schumpp et al. (2007) have performed the inoculation tests on several *Lotus* species with various kinds of viruses that are known to infect legumes such as *Medicago sativa*, or *Trifolium pratense*. However, these *Lotus* plants were resistant to almost all tested viruses and only specific isolates of *Alfalfa mosaic virus* and *Tobacco ringspot virus* were able to infect occasionally. After the adaptation of these viruses, highly virulent inoculum of *Alfalfa mosaic virus* to *Lotus* plants was obtained (Schumpp et al. 2007). Because the susceptibility to the obtained virus varied among the different *Lotus* accessions or species, this pathosystem seems to be a good model for analyzing the genes involved in the virus resistance in legumes.

Fungal pathogen, *Fusarium solani*, was also reported as a model pathogen of *L. japonicus* (Takeuchi et al. 2007). A fungal isolate, MAFF240020 (deposited in GenBank, National Institute of Agrobiological Sciences), causes the wilt disease by root rot in *L. japonicus* accession Gifu B-129. In addition, MAFF240020 also causes chlorotic to necrotic lesions on leaves after wound inoculation. Thus, *F. solani* MAFF240020 is suitable for studying pathogenic interactions in *L. japonicus*.

15.6 Conclusion

As discussed above, our knowledge about the *Lotus* defense mechanism largely relies on the study of symbiosis and mainly restricted to MTI. Even so, these studies revealed the presence of close relationship between defense and symbiosis and thus highlight strongly the importance of the study of *Lotus*–pathogen interactions. Investigations of the defense mechanisms in *L. japonicus*, especially about ETI, will contribute to the understanding of not only legume defense mechanisms but also the true nature of symbiosis.

References

Alippi A (2005) Bacterial diseases of *Lotus* spp. Lotus Newslett 35(1):17–18

Ausmees N, Kobayashi H, Deakin WJ, Marie C, Krishnan HB, Broughton WJ, Perret X (2004) Characterization of NopP, a type III secreted effector of *Rhizobium* sp. strain NGR234. J Bacteriol 186(14):4774–4780

Bartsev AV, Deakin WJ, Boukli NM, McAlvin CB, Stacey G, Malnoë P, Broughton WJ, Staehelin C (2004) NopL, an effector protein of *Rhizobium* sp. NGR234, thwarts activation of plant defense reactions. Plant Physiol 134(2):871–879

Chen WM, Laevens S, Lee TM, Coenye T, De Vos P, Mergeay M, Vandamme P (2001) Ralstonia taiwanensis sp. nov., isolated from root nodules of Mimosa species and sputum of a cystic fibrosis patient. Int J Syst Evol Microbiol 51(Pt 5):1729–1735

Christie PJ, Vogel JP (2000) Bacterial type IV secretion: conjugation systems adapted to deliver effector molecules to host cells. Trends Microbiol 8(8):354–360

Cullimore JV, Ranjeva R, Bono JJ (2001) Perception of lipo-chitooligosaccharidic Nod factors in legumes. Trends Plant Sci 6(1):24–30

Deakin WJ, Broughton WJ (2009) Symbiotic use of pathogenic strategies: rhizobial protein secretion systems. Nat Rev Microbiol 7(4):312–320

Deakin WJ, Marie C, Saad MM, Krishnan HB, Broughton WJ (2005) NopA is associated with cell surface appendages produced by the type III secretion system of *Rhizobium* sp. strain NGR234. Mol Plant Microbe Interact 18(5):499–507

Dodds PN, Rathjen JP (2010) Plant immunity: towards an integrated view of plant-pathogen interactions. Nat Rev Genet 11(8):539–548

Felix G, Duran JD, Volko S, Boller T (1999) Plants have a sensitive perception system for the most conserved domain of bacterial flagellin. Plant J 18(3):265–276

Gomez-Gomez L, Boller T (2000) FLS2: an LRR receptor-like kinase involved in the perception of the bacterial elicitor flagellin in *Arabidopsis*. Mol Cell 5 (6):1003–1011

Graham PH, Vance CP (2003) Legumes: importance and constraints to greater use. Plant Physiol 131 (3):872–877

Hayashi M, Saeki Y, Haga M, Harada K, Kouchi H, Umehara Y (2012) Rj (rj) genes involved in nitrogen-fixing root nodule formation in soybean. Breed Sci 61 (5):544–553

Heidrich K, Blanvillain-Baufume S, Parker JE (2012) Molecular and spatial constraints on NB-LRR receptor signaling. Curr Opin Plant Biol 15(4):385–391

Hueck CJ (1998) Type III protein secretion systems in bacterial pathogens of animals and plants. Microbiol Mol Biol Rev 62(2):379–433

Jones JD, Dangl JL (2006) The plant immune system. Nature 444(7117):323–329

Kambara K, Ardissone S, Kobayashi H, Saad MM, Schumpp O, Broughton WJ, Deakin WJ (2009) Rhizobia utilize pathogen-like effector proteins during symbiosis. Mol Microbiol 71(1):92–106

Kouchi H, Imaizumi-Anraku H, Hayashi M, Hakoyama T, Nakagawa T, Umehara Y, Suganuma N, Kawaguchi M (2010) How many peas in a pod? Legume genes responsible for mutualistic symbioses underground. Plant Cell Physiol 51(9):1381–1397

Krause A, Doerfel A, Göttfert M (2002) Mutational and transcriptional analysis of the type III secretion system of *Bradyrhizobium japonicum*. Mol Plant Microbe Interact 15(12):1228–1235

Lopez-Gomez M, Sandal N, Stougaard J, Boller T (2012) Interplay of flg22-induced defence responses and nodulation in *Lotus japonicus*. J Exp Bot 63 (1):393–401

Marie C, Deakin WJ, Viprey V, Kopciñska J, Golinowski W, Krishnan HB, Perret X, Broughton WJ (2003) Characterization of Nops, nodulation outer proteins, secreted via the type III secretion system of NGR234. Mol Plant Microbe Interact 16(9):743–751

Miya A, Albert P, Shinya T, Desaki Y, Ichimura K, Shirasu K, Narusaka Y, Kawakami N, Kaku H, Shibuya N (2007) CERK1, a LysM receptor kinase, is essential for chitin elicitor signaling in *Arabidopsis*. Proc Natl Acad Sci U S A 104(49):19613–19618

Monaghan J, Zipfel C (2012) Plant pattern recognition receptor complexes at the plasma membrane. Curr Opin Plant Biol 15(4):349–357

Moulin L, Munive A, Dreyfus B, Boivin-Masson C (2001) Nodulation of legumes by members of the beta-subclass of Proteobacteria. Nature 411 (6840):948–950

Nakagawa T, Kaku H, Shimoda Y, Sugiyama A, Shimamura M, Takanashi K, Yazaki K, Aoki T, Shibuya N, Kouchi H (2011) From defense to symbiosis: limited alterations in the kinase domain of LysM receptor-like kinases are crucial for evolution of legume-Rhizobium symbiosis. Plant J 65 (2):169–180

Okazaki S, Okabe S, Higashi M, Shimoda Y, Sato S, Tabata S, Hashiguchi M, Akashi R, Gottfert M, Saeki K (2010) Identification and functional analysis of type III effector proteins in *Mesorhizobium loti*. Mol Plant Microbe Interact 23(2):223–234

Perret X, Freiberg C, Rosenthal A, Broughton WJ, Fellay R (1999) High-resolution transcriptional analysis of the symbiotic plasmid of *Rhizobium* sp. NGR234. Mol Microbiol 32(2):415–425

Radutoiu S, Madsen LH, Madsen EB, Felle HH, Umehara Y, Gronlund M, Sato S, Nakamura Y, Tabata S, Sandal N, Stougaard J (2003) Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases. Nature 425(6958):585–592

Saad MM, Kobayashi H, Marie C, Brown IR, Mansfield JW, Broughton WJ, Deakin WJ (2005) NopB, a type III secreted protein of *Rhizobium* sp. strain NGR234, is associated with pilus-like surface appendages. J Bacteriol 187(3):1173–1181

Sánchez C, Iannino F, Deakin WJ, Ugalde RA, Lepek VC (2009) Characterization of the *Mesorhizobium loti* MAFF303099 type-three protein secretion system. Mol Plant Microbe Interact 22(5):519–528

Schauser L, Roussis A, Stiller J, Stougaard J (1999) A plant regulator controlling development of symbiotic root nodules. Nature 402(6758):191–195

Schumpp O, Ramel ME, Gugerli P, Broughton WJ, Deakin WJ (2007) Identification of a *Lotus* viral pathogen. J Plant Res 120(5):651–654

Segonzac C, Zipfel C (2011) Activation of plant pattern-recognition receptors by bacteria. Curr Opin Microbiol 14(1):54–61

Shao F, Golstein C, Ade J, Stoutemyer M, Dixon JE, Innes RW (2003) Cleavage of *Arabidopsis* PBS1 by a bacterial type III effector. Sci Signal 301(5637):1230

Shimada N, Sato S, Akashi T, Nakamura Y, Tabata S, Ayabe S, Aoki T (2007) Genome-wide analyses of the structural gene families involved in the legume-specific 5-deoxyisoflavonoid biosynthesis of *Lotus japonicus*. DNA Res 14(1):25–36

Shimizu T, Nakano T, Takamizawa D, Desaki Y, Ishii-Minami N, Nishizawa Y, Minami E, Okada K, Yamane H, Kaku H, Shibuya N (2010) Two LysM receptor molecules, CEBiP and OsCERK1, cooperatively regulate chitin elicitor signaling in rice. Plant J 64(2):204–214

Sisterna M, Lori GA (2005) Fungal diseases on *Lotus* spp in Argentina. Lotus Newslett 35(1):15–16

Skorpil P, Saad MM, Boukli NM, Kobayashi H, Ares-Orpel F, Broughton WJ, Deakin WJ (2005) NopP, a phosphorylated effector of *Rhizobium* sp. strain NGR234, is a major determinant of nodulation of the tropical legumes *Flemingia congesta* and *Tephrosia vogelii*. Mol Microbiol 57(5):1304–1317

Spaink HP (1995) The molecular basis of infection and nodulation by rhizobia: the ins and outs of sympathogenesis. Annu Rev Phytopathol 33(1):345–368

Takeuchi K, Tomioka K, Kouchi H, Nakagawa T, Kaku H (2007) A novel pathosystem to study the interactions between *Lotus japonicus* and *Fusarium solani*. J Gen Plant Pathol 73:336–341

Tsukui T, Eda S, Kaneko T, Sato S, Okazaki S, Kakizaki-Chiba K, Itakura M, Mitsui H, Yamashita A, Terasawa K (2013) The Type III Secretion System of *Bradyrhizobium japonicum* USDA122 mediates symbiotic incompatibility with Rj2 soybean plants. Appl Environ Microbiol 79(3):1048–1051

Vandamme P, Goris J, Chen WM, de Vos P, Willems A (2002) *Burkholderia tuberum* sp. nov. and *Burkholderia phymatum* sp. nov., nodulate the roots of tropical legumes. Syst Appl Microbiol 25(4):507–512

Velazquez E, Peix A, Zurdo-Pineiro JL, Palomo JL, Mateos PF, Rivas R, Munoz-Adelantado E, Toro N, Garcia-Benavides P, Martinez-Molina E (2005) The coexistence of symbiosis and pathogenicity-determining genes in *Rhizobium rhizogenes* strains enables them to induce nodules and tumors or hairy roots in plants. Mol Plant Microbe Interact 18(12):1325–1332

Viboud GI, Bliska JB (2005) Yersinia outer proteins: role in modulation of host cell signaling responses and pathogenesis. Annu Rev Microbiol 59:69–89

Viprey V, Del Greco A, Golinowski W, Broughton WJ, Perret X (1998) Symbiotic implications of type III protein secretion machinery in *Rhizobium*. Mol Microbiol 28(6):1381–1389

Yang S, Tang F, Gao M, Krishnan HB, Zhu H (2010) R gene-controlled host specificity in the legume-rhizobia symbiosis. Proc Natl Acad Sci USA 107 (43):18735–18740

Zehner S, Schober G, Wenzel M, Lang K, Göttfert M (2008) Expression of the *Bradyrhizobium japonicum* type III secretion system in legume nodules and analysis of the associated tts box promoter. Mol Plant Microbe Interact 21(8):1087–1093

Zhang XC, Wu X, Findley S, Wan J, Libault M, Nguyen HT, Cannon SB, Stacey G (2007) Molecular evolution of lysin motif-type receptor-like kinases in plants. Plant Physiol 144(2):623–636

Zhu H, Riely BK, Burns NJ, Ane JM (2006) Tracing nonlegume orthologs of legume genes required for nodulation and arbuscular mycorrhizal symbioses. Genetics 172(4):2491–2499

Metabolomics

16

Yuji Sawada and Toshio Aoki

Abstract

Metabolomics is an "omics" approach for the comprehensive profiling of the small molecules in an organism, which has been made possible by recent advances in mass spectrometry (MS), and is considered to be one of the essential tools for functional genomics and systems biology. Several methodologies have been developed to meet the various types of requirements of metabolomics, such as high throughput, high sensitivity, high selectivity, unbiasedness and reproducibility. This chapter outlines recent advances in plant metabolomics with case studies of *Lotus japonicus*, including targeted analyses, widely targeted analysis and non-targeted analysis, and poses the next challenges for integrated metabolomics.

16.1 Introduction

The last two decades have seen advances in technology that have brought about revolutionary changes in biological research. The level of analysis has shifted from studying the functions and expression of individual genes to studying large numbers of genes and gene products simultaneously. Advances in automated nucleotide sequencing have enabled the accumulation of enormous sets of sequence information for both genomic and complementary DNAs, and genomes of a broad range of organisms continue to be sequenced. mRNA profiling using microarrays and whole transcriptome shotgun sequencing (RNA-seq) allows comprehensive analysis of transcripts (De Luis et al. 2012). Advances in mass spectrometry (MS) have made comprehensive and highly sensitive analyses of proteins and metabolites possible. These technologies have advanced the fields of functional genomics and systems biology, both of which depend on the comprehensive profiling of large numbers of gene expression products. Such profiling is referred to as transcriptomics, proteomics and metabolomics.

Metabolomics is an "omics" approach for the comprehensive detection of the small molecules in an organism, including amino acids, organic

Y. Sawada
RIKEN Center for Sustainable Resource Science, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Y. Sawada
RIKEN Plant Science Center, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

T. Aoki (✉)
Department of Applied Biological Sciences, Nihon University, Fujisawa, Kanagawa 252-0880, Japan
e-mail: aoki.toshio@nihon-u.ac.jp

S. Tabata and J. Stougaard (eds.), *The Lotus japonicus Genome*, Compendium of Plant Genomes,
DOI 10.1007/978-3-662-44270-8_16,

acids, carbohydrates and others. The central dogma of biology states that genetic information is transcribed into individual mRNA; each mRNA contains the program for synthesis of a particular protein. However, a large number of proteins in fact serve as enzymes involved in metabolic pathways whose metabolites largely govern the phenotype. Gene perturbations lead to changes in the transcript that result in changes in the enzyme levels and in turn in the metabolic profile. A comprehensive description of metabolic behaviour is thus a critical piece of information in biology.

Plant metabolomics is of particular interest because plants are characterized by their diversity of so-called secondary metabolites, also referred to as natural products or phytochemicals. To date, more than 100,000 metabolites have been identified, and this number may well be less than 10 % of the total (Wink 1988). The plant metabolome—the complement of metabolites in a plant species—represents enormous chemical diversity owing to the complex set of metabolites produced. Estimates of plant metabolomes vary from 5,000 to 25,000; even in the absence of definite data, the metabolomes of plants seem to be larger than those of prokaryotes or animals. Human beings use plants for their metabolites as dyestuffs, resins, fibres, oils, fats, seasonings, flavourings and pharmaceutical agents. It is well accepted that phytochemicals play critical roles in resistance against pathogens, herbivores and other environmental stresses. Plant metabolomics would provide fascinating data to researchers working on plant metabolism. Plant metabolomics not only contributes to functional genomics and systems biology of plants but can also be exploited in molecular breeding aimed at improving productivity and functionality of crops, incorporating stress tolerance, and producing pharmaceutical materials, functional foods, biomaterials and biofuels (Xu et al. 2013).

At present, there are large gaps in our knowledge of genomics and metabolomics because metabolites have more complex than the elements of the classical central dogma. The extreme complexity of metabolites, especially plant metabolites, lies not only in their great number but also in their chemical diversity. High-throughput, high-sensitivity, high-selectivity and unbiased analytical methodologies that permit better handling of phytochemicals have been required ever since the beginning of plant metabolomics. For instance, untargeted analyses using MS-based technology have been developed to provide the metabolic profiles of known and unknown phytochemicals simultaneously (Sumner et al. 2003; Weckwerth 2003; Schauer and Fernie 2006; Guy et al. 2008; Hall et al. 2008; Saito et al. 2008).

Reproducibility, another requirement of metabolomics, would be ensured by unambiguous identification of each metabolite measured, but in general, it is not easy due to the vast diversity of phytochemicals. MS provides two types of structural information: the molecular mass and the fragment ion profile of the compound, i.e., its mass spectrum. The reproducibility of fragmentation depends on the ionization technique used. Electron ionization (EI), the standard ionization method in gas chromatography (GC)-MS, is accompanied by the cleavage of the compounds, generating a series of fragment ions according to the ionization energy. The resultant MS spectrum is highly reproducible, particularly when obtained at 70 eV, and can be used for the identification of the compound by searching a MS spectra library containing more than 100,000 spectra. In contrast, relatively soft ionization methods, such as electrospray ionization (ESI) and atmospheric pressure chemical ionization (APCI), which are suitable for liquid chromatography (LC)-MS and capillary electrophoresis (CE)-MS, generate only a limited number of fragment ions. In the case of LC- or CE-MS, tandem mass spectrometry (MS/MS) is applied to obtain structural information on the basis of the fragmentation pattern. LC(CE)-ESI-MS and LC(CE)-APCI-MS are highly sensitive techniques that provide information about molecular mass. However, they are not useful for the identification of compounds in a strict sense because their ionization conditions are not standardized and only a few hundreds to thousands of MS and MS/MS spectra obtained by ESI or APCI are available in databases. Except in GC-EI-MS, reproducibility of analysis may seem to be

incompatible with high throughput; conventional targeted analyses ensure reproducibility, but their low throughput does not meet the requirements of metabolomics. However, in recent LC-based untargeted analyses, high repeatability has been achieved together with high throughput, and reproducibility could be obtained in combination with validation of the data for the compounds detected. Sawada et al. (2009) proposed a widely targeted analysis aimed at further compatibility between throughput and reproducibility in LC-based analysis in consideration of the data validation (Sawada et al. 2009). In this review, we introduce an overview of plant metabolomics and selected case studies of *Lotus japonicus*, including targeted analyses, widely targeted analysis, integrated metabolomics, an MS/MS database for phytochemicals, and future aspects of *L. japonicus* metabolomics.

16.2 Targeted Analyses for Biosynthetic Research on Leguminous Phytochemicals

Fabaceae (in the Angiosperm Phylogeny Group system) is the third most diverse family in the flowering plants and is known as a rich source of phytochemicals. Many of the metabolites of members of this family have bioactivity beneficial for human health (Dixon and Sumner 2003). Several Fabaceae genera, such as *Acacia*, *Astragalus*, *Caesalpinia*, *Cassia*, *Glycyrrhiza*, *Lablab*, *Pueraria* and *Sophora,* contain medicinal plants. Leguminous plants are also characterized by symbiotic nitrogen fixation, pod formation, metabolite accumulation in seeds and the formation of compound leaves and bilateral flowers. Leguminous crops are valuable resources for protein and oil; soya bean is one of the principal world crops. Metabolomics of leguminous plants is thus expected to have significant impact on broad fields.

Of particular interest for several reasons are flavonoids (including isoflavonoids) and triterpenoids in leguminous plants. In legume–rhizobia symbiotic systems, (iso)flavonoids and chalcones, another class of phenolic compounds closely related to flavonoids, play essential roles as signal molecules of the host plants towards the symbionts (Hayashi et al. 2000). The isoflavonoid pathway of leguminous plants also produces inducible antibiotic substances, termed phytoalexins, in response to challenges by biotic stresses such as infection by pathogenic microorganisms (Aoki et al. 2000). Although the distribution of isoflavonoids is not exclusive to Fabaceae, 90 % of isoflavonoids reported to date are found in Fabaceae. Therefore, isoflavonoids are considered as characteristic phytochemicals of this family. Since the 1970s, the biosynthesis of isoflavonoid phytoalexins has been studied biochemically, and the identification of 2-hydroxyisoflavanone synthase, the key enzyme in isoflavone formation, from a *Glycyrrhiza* species and soya bean (Akashi et al. 1999a; Steele et al. 1999) has provided important clues for the molecular basis of isoflavonoid biosynthesis.

Triterpenes are a major subgroup of the terpene superfamily, the largest group of phytochemicals. They are derived from the cyclization of 2,3-oxidosqualene by oxidosqualene cyclases (OSCs), whose first identification led to subsequent studies of this pathway (Abe and Prestwich 1995). Triterpenes that are the direct products of OSCs, such as α-amyrin, β-amyrin and lupeol, have been suggested to play roles in rhizobial and mycorrhizal symbioses, and lupeol was shown to have negative effects on nodule formation, as revealed by a transgenic approach using *L. japonicus* (Delis et al. 2010 and references cited therein). Triterpene saponins, which are triterpenes modified by hydroxylation and glycosylation, include various bioactive natural products of legumes (Dixon and Sumner 2003).

In general, research on plant metabolism begins with the analysis of metabolites and the elucidation of the condition that induces the metabolic pathway of interest, together with a search for suitable plant materials. Suspension cell cultures are powerful tools in many cases. Identification of biosynthetic enzymes includes biochemical assays with crude extracts prepared from the plant materials and recombinant cells heterologously expressing cDNAs that putatively

encode the target enzyme. Identification of the metabolites in the plant materials and reaction products of the enzyme assays are essential for metabolic studies including the functional characterization of biosynthetic genes. The techniques that have been used in the field of plant secondary metabolism since its pioneering studies are classical low-throughput techniques, such as LC equipped with a UV or photodiode array detector, infusion EI-MS and nuclear magnetic resonance spectroscopy (Akashi et al. 1998a, b, 1999a, b; Sawada et al. 2002). Until very recently, studies in this field have not benefited from the advances in technology that enabled metabolomics.

Validation of gene functions using a model plant provides reliable clues to the complex metabolism of leguminous plants. In the late 1990s, studies aimed at the comprehensive clarification of structural genes involved in the biosynthesis of isoflavonoid phytoalexins (Shimada et al. 2003, 2005, 2006, 2007), and triterpenes (Sawai et al. 2006a, b) of *L. japonicus* were begun with the aim of establishing a basis for functional genomics and molecular genetic approaches for the model legume. Candidate genes and cDNAs were cloned with the powerful aid of genomics data, such as expression sequence tags and genome sequence information, both of which have been accumulated in the genome project of *L. japonicus* promoted by Kazusa DNA Research Institute (KDRI). However, the analysis of metabolites and elucidation of gene functions by biochemical assays depended on "classical" procedures, i.e., the isolation of metabolites and their structural identification by low-throughput techniques capable of handling only limited numbers of targeted metabolites.

16.3 Untargeted Analysis for Metabolomics

Elucidation of the comprehensive metabolic profile, i.e., untargeted metabolic profiling, requires high-throughput and high-sensitivity/high-selectivity analysis of a broad range of phytochemicals. To meet these requirements, untargeted metabolomics aims to make the best use of MS-based technology combined with chromatographic separation. In transcriptomics, sequence-based gene annotation as well as expression data is essential for biological findings. The analogy with transcriptomics suggests the necessity of metabolite annotation, which can also be implemented by MS. GC-MS has the advantage that its standard ionization method, EI, reproducibly generates fragment ions, and this technology first enabled the high-throughput and unbiased analysis of thousands of metabolites from plants and other organisms (Dixon and Sumner 2003). Desbrosses et al. (2005) characterized the metabolic profiles of *L. japonicus* by GC-MS, measuring relative levels of primary metabolites in nodules, roots, leaves and flowers of symbiotic plants. Discrete metabolic characteristics were revealed by principal component analysis and hierarchical cluster analysis. Prior to measurement, they created mass spectral tag libraries that contained both retention time of GC and fragment ions representing metabolites (Desbrosses et al. 2005).

The recent development of LC-MS and sufficiently accurate high-resolution MS (HR/MS), which allows us to define the elemental compositions of detected ions, has extended the potential of plant metabolomics. In this technique, high-resolution mass data are provided as metabolite annotation. HR/MS of LC-Fourier transfer ion cyclotron resonance-MS (LC-FTICR-MS) is an example of the most advanced metabolomic techniques for untargeted analysis (Marshall and Hendrickson 2008; Glauser et al. 2012).

KDRI is a pioneer in metabolomics with LC-FTICR-MS, which was applied to characterize the metabolic diversity between the two generally used accessions of *L. japonicus*, B-129 Gifu and MG-20 Miyakojima. These accessions differ in stem colour and are accordingly thought to differ in their flavonoid metabolism (Suzuki et al. 2008). Natural metabolic variation among wild-type accessions has important implications for the mechanisms of environmental adaptation by plants. A total of 61 known and unknown

flavonoids were successfully assigned by their elemental compositions. These results clearly showed differences in flavonoid metabolism among accessions and developmental stages (Suzuki et al. 2008). At that time, peak picking and MS analysis were mainly performed by manual procedures. After this leading study, the next metabolomics platforms incorporated automated procedures using newly developed software and supported the comprehensive metabolite annotation of tomato (*Solanum lycopersicum*) (Iijima et al. 2008).

MS/MS is a powerful technique for the structural elucidation of metabolites because it discriminates fragment ions derived from related compounds that belong to the same class of metabolites but are partially modified by plant-specific methyltransferases and glycosidases. MS/MS spectra can provide an indication of putative structures of metabolites via manual interpretation or automated comparison with databases of fragmentation patterns. In conventional untargeted analysis, MS/MS data are acquired for only a limited number of peaks of interest, which are selected by a data mining method, and additional MS/MS analyses are required if other interesting peaks are found by another data mining method. Matsuda et al. (2009) developed a peak annotation procedure on the basis of an MS/MS spectral tag (MS2T) library for untargeted metabolic profiling analysis. MS/MS spectra of most of the LC peaks were acquired and stored in MS2T libraries. The experiments for spectral acquisition were repeated 25 times, altering the mass ranges used to select precursor ions. The libraries were created from MS/MS spectra acquired using the automatic data acquisition function of ultra-performance LC quadrupole-time-of-flight-MS (UPLC-QTOF) in experiments distinct from conventional metabolic profiling analyses (Matsuda et al. 2009). UPLC employing a sub-2-μm column dramatically improves metabolite separation. The narrowed peaks (3–6 s) are detected by high-sensitivity scanning of QTOF-MS (Fig. 16.1a, b). Using this platform, untargeted MS2Ts were collected in several developmental stages of Arabidopsis (Matsuda et al. 2010), and the total of MS2Ts to date have amounted to more than 1 million derived from 10 plant species (*A. thaliana*, *Oryza sativa*, *Triticum aestivum*, *S. lycopersicum*, *Glycine max*, *Chara* sp., *Pyrus communis*, *L. japonicus*, *Allium cepa* and *Brassica napus*). Together with the above-mentioned analytical procedure, an integrated searching system using a reference database of MS2Ts has been established at RIKEN PSC (Akiyama et al. 2008; Lim et al. 2003; Hank et al. 2003).

In untargeted analytical platforms with LC-FTICR/MS and UPLC-QTOF/MS, thousands of known and unknown peaks are detected and characterized by elemental composition analysis of HR/MS and MS/MS similarity searches of MS2Ts (Iijima et al. 2008; Matsuda et al. 2009). Given that these instruments are known as big data generators (approximately 20 GB per 30 min of acquisition in the profile mode), and the analytical samples for a project usually number at least a hundred, a practical strategy of research is needed for high-throughput metabolic profiling.

16.4 Widely Targeted Analysis

On a tandem quadrupole mass spectrometer (TQ-MS), targeted metabolites can be analysed using a non-scanning MS technique of high sensitivity and selectivity, termed selective reaction monitoring (SRM). In SRM, two mass analysers are used as static mass filters to monitor a specified fragment ion of a selected precursor ion (Fig. 16.1c). One of the essential parameters of SRM is the specific pair of m/z values associated with the precursor and fragment ions selected, referred to as a "transition." The instrument continuously records the intensity of the ions that match the selected transition, showing a chromatogram instead of MS spectra. SRM using TQ-MS has been used for the analysis of phytohormones in trace amounts (Kojima et al. 2009). A TQ-MS instrument can rapidly switch the transitions monitored and can record the intensity of each transition as a function of the retention time, allowing the measurement of multiple SRM transitions within the same experiment. A recently introduced high-speed

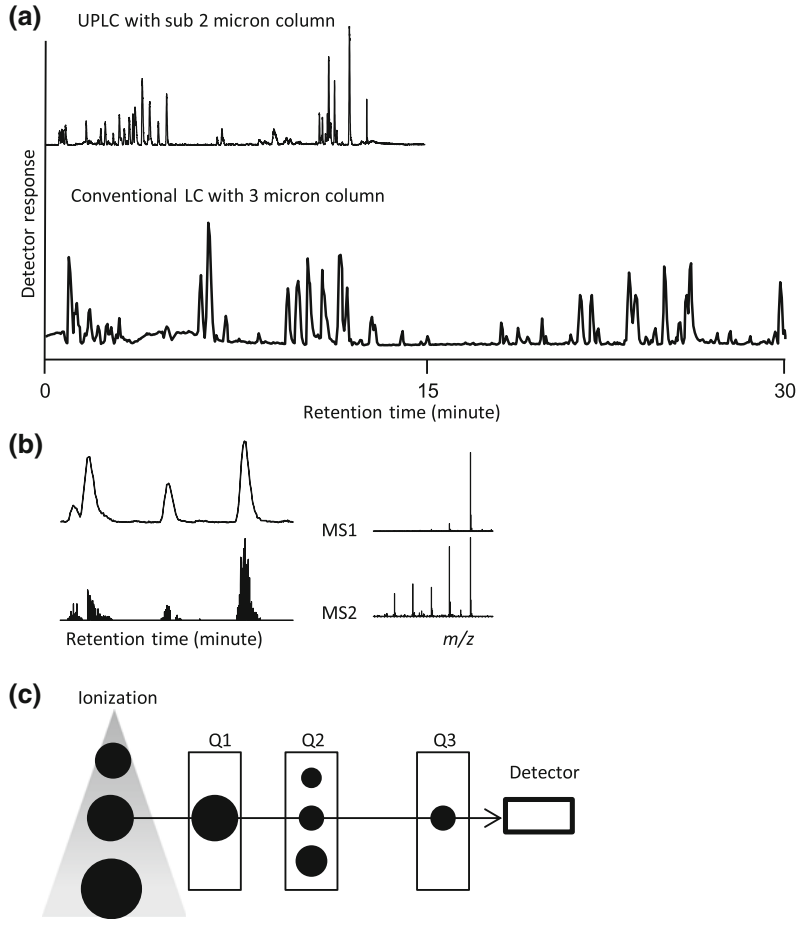


Fig. 16.1 Recent advances in metabolomics for LC separation and MS detection. **a** Sub-2-micron column of UPLC provides peak widths of a few seconds (*upper panel*), the 3-micron column of conventional LC provides peak widths of more than 10 s (*lower panel*). **b** Using high-sensitivity scanning of UPLC-QTOF-MS, almost all LC-separated peaks can be assigned to MS (precursor ions representing molecular weight) and MS/MS data (product ions). **c** In SRM analysis using a triple quadrupole (*Q1*, *Q2* and *Q3*) instrument, the targeted metabolite can be selectively detected by duplicated MS filtrations, the precursor ion in *Q1* and one of the product ions in *Q3* (after fragmentation in *Q2*)

instrument can detect as many as several hundred SRM transitions. The SRM condition set, which comprises retention time, transition, repeller voltage (sampling cone voltage) and ion-accelerating voltage for fragmentation, must be optimized for each metabolite to be analysed. If SRM conditions for a broad range of metabolites are available, large-scale multiple SRM analysis will provide a novel technique for metabolomics. Furthermore, given that SRM is originally a method for targeted analysis, reproducible analysis can be implemented if the analytical conditions and data obtained are validated with authentic compounds and case studies using several plant materials. The SRM conditions for approximately 1,000 authentic compounds were optimized by liquid handling systems and flow injection analysis using TQ-MS and UPLC-TQ-MS (Yamazaki et al. 2013). This analytical system, named widely targeted analysis, is practical for plant metabolomics and offers several advantages including reproducible detection of various (more than 500) metabolites, femtomole-level ultra-high sensitivity and high throughput

enabling the analysis of more than 10^4 samples per year.

Recently, the accelerated increase, in many species, in plant resources for functional genomics, such as wild accessions, insertional knockout/overexpresser lines and recombinant inbred lines (RILs), has made high-throughput plant metabolomics ever more necessary. As a case study, a widely targeted analysis of amino acids and amino acid-derived secondary metabolites was used to screen approximately 3,000 transposon insertional knockout lines and wild accessions of Arabidopsis for over- and under-accumulation mutants of specific metabolites (Nakabayashi et al. 2013a). As a result, one line was shown to accumulate branched-chain amino acids (valine, leucine and isoleucine) in up to 100 times the amount of the wild type of the background ecotype. Interestingly, this mutant phenotype may be specific to that accession; no mutant that overaccumulates branched-chain amino acids has been found in T-DNA insertion lines provided by the Salk Institute for Biological Studies. Characterization of such natural variation is important for elucidating plant metabolic regulation. Moreover, quantitative locus (QTL) mapping using RILs is expected to facilitate the identification of novel genes without reference to sequence homology. Widely targeted metabolic profiling will assist metabolic QTL (mQTL) studies, which are effective in assigning enzymes and regulatory genes to the network of known metabolic reactions (Lisec et al. 2008; Brotman et al. 2011).

16.5 Integrated Metabolomics with Recombinant Inbred Lines of *L. japonicus*

L. japonicus is used for functional genomic approaches in studies of leguminous metabolism because its genomic infrastructure has facilitated the identification and characterization of biosynthetic genes (Shimada et al. 2005, 2007; Forslund et al. 2004; Morant et al. 2008; Saito et al. 2012). Accumulating DNA markers and SNPs of the model legume is making mQTL analysis feasible, in which the metabolic profile is used as a quantitative phenotype for QTL. This analysis is expected to address the gaps between genomics and metabolomics. An integrated metabolomics platform consisting of a combination of widely targeted and untargeted analyses was established with the aim of discovering the comprehensive mQTL set of *L. japonicus*. The putative metabolite structures associated with significant mQTLs were assigned by a MS/MS database search for phytochemicals (Sawada et al. 2012; Sawada and Hirai 2013).

As described above, two UPLC-MS/MS-based platforms for plant metabolomics are available: untargeted analysis using UPLC-QTOF-MS and widely targeted analysis using UPLC-TQ-MS. UPLC-QTOF-MS for untargeted analysis acquires the data of molecular mass and MS spectra for virtually all LC-separated peaks, but they are redundant for each peak. Untargeted analysis thus yields a vast amount of data (a few gigabytes per analysis) and can be applied only with difficulty to large-scale analyses dealing with hundreds to thousands of samples. In contrast, in SRM using UPLC-TQ-MS for widely targeted analysis, no MS spectrum is recorded, and the transition and other parameters for each compound are optimized in advance. Accordingly, SRM using UPLC-TQ-MS yields data sets of practical size (a few megabytes per analysis) and provides a high-throughput metabolomics platform.

Given that high-sensitivity detection by TQ-MS requires optimization of analytical conditions with authentic standard compounds, the next challenge of widely targeted analysis is to detect unidentified metabolites. For this purpose, the SRM conditions for unknown compounds may be best obtained by conversion of the data acquired in an untargeted analysis with UPLC-QTOF-MS, such as MS2Ts. In the new integrated method, data conversion is implemented using a newly developed program, and unidentified compounds corresponding to MS2Ts are analysed using SRM conditions that are converted from MS2Ts, named integrated SRM (iSRM). In the case study in *L. japonicus* seeds (B-129 and MG-20), a total of 80554 MS2Ts were collected in positive and negative ion modes. The SRM conditions derived from the MS2Ts were tested using the MS peak

intensity and analytical threshold, and 342 iSRMs were successfully optimized, 88 of which were significantly different between B-129 and MG-20 (Sawada et al. 2012).

Each iSRM is linked to MS2T data. Thus, in principle, all the detectable metabolites can be annotated on the basis of the putative structures indicated by the MS/MS spectra. The annotation rate of metabolites depends on the availability of MS and MS/MS databases and the similarity of search algorithm (Smith et al. 2005; Fiehn et al. 2005, 2008; Wishart et al. 2007, 2009; Cui et al. 2008; Soh et al. 2003; Mylonas et al. 2009). A plant-specific MS/MS database and a search algorithm established in 2012 will accelerate the annotation rate of plant metabolomics data (http://spectra.psc.riken.jp/).

The new version of widely targeted analysis using iSRM was applied to RILs between MG-20 and B-129 (Hayashi et al. 2001; Kawaguchi et al. 2001; Klein and Grusak 2009; Gondo et al. 2007) to reveal eight mQTLs (LOD score >10). Annotation of the peaks detected by iSRMs was successful using ReSpect search and suggested flavonoid glycosides. These results suggest that widely targeted analysis with iSRM and MS2T-based annotation by MS/MS database search have the potential for effective elucidation of plant metabolism. Sawada and Hirai (2013) developed iSRM for soya bean recombinant inbred lines derived from *G. max* and *G. soja* and found several mQTLs.

16.6 Future Aspects for Metabolomics

As practical metabolomics platforms, iSRM based on MS2T and HR/MS of LC-FTICR-MS have been established for quantitative and qualitative analysis and for elemental composition analysis, respectively (Nakabayashi et al. 2013b). SRM, iSRM, MS2T and HR/MS are now integrated for quantitative and qualitative metabolomics. As an application of the new metabolomics, mQTL analysis of *L. japonicus* has the potential to improve metabolite annotations. If an mQTL is assigned to a biosynthetic gene already characterized, the metabolite annotations could be validated by the gene function. The gene annotation associated with metabolomics information could narrow down the candidate structure of the unknown metabolites, e.g., a cytochrome P450-associated metabolite can be predicted to have hydroxyl groups in its chemical structure. As a source of high-density markers, single nucleotide

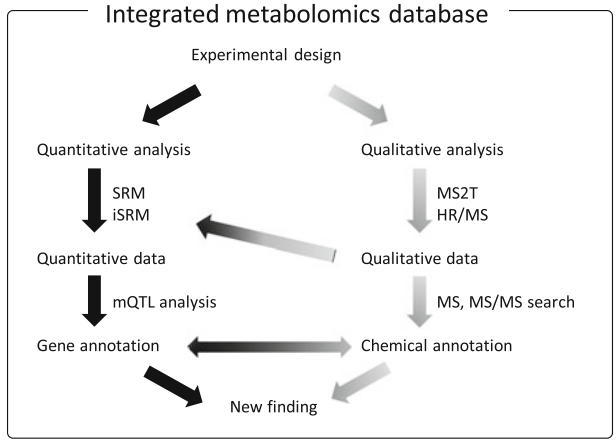


Fig. 16.2 Workflow of integrated metabolomics. On the *left side* (*black arrows*), quantitative data can be obtained by SRM and iSRM analyses. In case of annotation of metabolite biosynthesis genes, data are used for mQTL analysis (linkage mapping and GWAS). On the *right side* (*grey arrows*), qualitative data can be obtained by MS2T and HR/MS analysis. In case of chemical annotation of detectable metabolites, data are used for searching external databases. The goal of the integrated metabolomics database is efficient mining of biological discoveries

polymorphisms (SNPs), which are defined by whole-genome sequencing of accessions, are available (Huang et al. 2010; Han and Huang 2013). Based on a sufficient density of SNPs (a few SNPs per gene) in a few hundred accessions, a genome-wide association study (GWAS) could be conducted. Metabolome GWAS (mGWAS) will dramatically improve the annotation of genes and metabolites (Adamski and Suhre 2013). mGWAS analysis in *L. japonicus* will promote the metabolic breeding of leguminous crops. In the course of these activities, an integrated metabolomics database will be developed, and databases and data resources will be standardized as a practical metabolomics platform for mining biological discoveries (Fig. 16.2).

16.7 Conclusions

This chapter outlines recent advances in plant metabolomics with case studies of *L. japonicus* and poses the next challenges for integrated metabolomics. It is desirable that these activities will become general and conventional practice in model and non-model plants. The public data of *L. japonicus* metabolomics will be especially useful in many studies of leguminous crops of economic importance.

Acknowledgments This work was supported by the RIKEN Plant Science Center, Japan Science and Technology Agency (CREST grant), Japan Advanced Plant Science Research Network and Fuji Foundation for Protein Research. RILs were distributed by the National BioResource Project (Legume Base). We thank Dr. Ryo Nakabayashi (RIKEN), Mrs. Makoto Suzuki (RIKEN), Akane Sakata (RIKEN), Yutaka Yamada (RIKEN), Tetsuya Mori (RIKEN) and Muneo Sato (JST GRENE NC-CARP) for their useful comments and LC-MS/MS analysis of the metabolomics technologies.

References

Abe I, Prestwich GD (1995) Molecular-cloning, characterization, and functional expression of rat oxidosqualene cyclase cDNA. Proc Natl Acad Sci USA 92 (20):9274–9278

Adamski J, Suhre K (2013) Metabolomics platforms for genome wide association studies–linking the genome to the metabolome. Curr Opin Biotech 24(1):39–47. doi:10.1016/j.copbio.2012.10.003

Akashi T, Aoki T, Ayabe S (1998a) CYP81E1, a cytochrome P450 cDNA of licorice (*Glycyrrhiza echinata* L.), encodes isoflavone 2′-hydroxylase. Biochem Biophys Res Commun 251(1):67–70

Akashi T, Aoki T, Ayabe S (1998b) Identification of a cytochrome P450 cDNA encoding (2S)-flavanone 2-hydroxylase of licorice (*Glycyrrhiza echinata* L.; Fabaceae) which represents licodione synthase and flavone synthase II. FEBS Lett 431(2):287–290

Akashi T, Aoki T, Ayabe S (1999a) Cloning and functional expression of a cytochrome P450 cDNA encoding 2-hydroxyisoflavanone synthase involved in biosynthesis of the isoflavonoid skeleton in licorice. Plant Physiol 121(3):821–828

Akashi T, Fukuchi-Mizutani M, Aoki T, Ueyama Y, Yonekura-Sakakibara K, Tanaka Y, Kusumi T, Ayabe S (1999b) Molecular cloning and biochemical characterization of a novel cytochrome P450, flavone synthase II, that catalyzes direct conversion of flavanones to flavones. Plant Cell Physiol 40(11):1182–1186

Akiyama K, Chikayama E, Yuasa H, Shimada Y, Tohge T, Shinozaki K, Hirai MY, Sakurai T, Kikuchi J, Saito K (2008) PRIMe: a Web site that assembles tools for metabolomics and transcriptomics. Silico Biol 8 (3–4):339–345

Aoki T, Akashi T, Ayabe S (2000) Flavonoids of leguminous plants: structure, biological activity, and biosynthesis. J Plant Res 113(1112):475–488

Brotman Y, Riewe D, Lisec J, Meyer RC, Willmitzer L, Altmann T (2011) Identification of enzymatic and regulatory genes of plant metabolism through QTL analysis in Arabidopsis. J Plant Physiol 168 (12):1387–1394. doi:10.1016/j.jplph.2011.03.008

Cui Q, Lewis IA, Hegeman AD, Anderson ME, Li J, Schulte CF, Westler WM, Eghbalnia HR, Sussman MR, Markley JL (2008) Metabolite identification via the madison metabolomics consortium database. Nat Biotechnol 26(2):162–164. doi:10.1038/nbt0208-162)

De Luis A, Markmann K, Cognat V, Holt DB, Charpentier M, Parniske M, Stougaard J, Voinnet O (2012) Two microRNAs linked to nodule infection and nitrogen-fixing ability in the legume *Lotus japonicus*. Plant Physiol 160(4):2137–2154. doi:10.1104/pp.112.204883

Delis C, Krokida A, Georgiou S, Peña Rodríguez LM, Kavroulakis N, Ioannou E, Roussis V, Osbourn AE, Papadopoulou KK (2010) Role of lupeol synthase in *Lotus japonicus* nodule formation. New Phytol. doi:10.1111/j.1469-8137.2010.03463.x

Desbrosses GG, Kopka J, Udvardi MK (2005) *Lotus japonicus* metabolic profiling. Development of gas chromatography-mass spectrometry resources for the study of plant-microbe interactions. Plant Physiol 137 (4):1302–1318. doi:10.1104/pp.104.054957

Dixon RA, Sumner LW (2003) Legume natural products: understanding and manipulating complex pathways for human and animal health. Plant Physiol 131 (3):878–885

Fiehn O, Wohlgemuth G, Scholz M (2005) Setup and annotation of metabolomic experiments by integrating biological and mass spectrometric metadata. Proc Lect Notes Bioinformatics 3615:224–239

Fiehn O, Wohlgemuth G, Scholz M, Kind T, Lee DY, Lu Y, Moon S, Nikolau B (2008) Quality control for plant metabolomics: reporting MSI-compliant studies. Plant J 53(4):691–704. doi:10.1111/j.1365-313X.2007.03387.x

Forslund K, Morant M, Jorgensen B, Olsen CE, Asamizu E, Sato S, Tabata S, Bak S (2004) Biosynthesis of the nitrile glucosides rhodiocyanoside A and D and the cyanogenic glucosides lotaustralin and linamarin in *Lotus japonicus*. Plant Physiol 135(1):71–84. doi:10.1104/pp.103.038059

Glauser G, Veyrat N, Rochat B, Wolfender JL, Turlings TC (2012) Ultra-high pressure liquid chromatography-mass spectrometry for plant metabolomics: a systematic comparison of high-resolution quadrupole-time-of-flight and single stage Orbitrap mass spectrometers. J Chromatogr A. doi:10.1016/j.chroma.2012.12.009

Gondo T, Sato S, Okumura K, Tabata S, Akashi R, Isobe S (2007) Quantitative trait locus analysis of multiple agronomic traits in the model legume *Lotus japonicus*. Genome 50(7):627–637. doi:10.1139/g07-040

Guy C, Kopka J, Moritz T (2008) Plant metabolomics coming of age. Physiol Plant 132(2):113–116. doi:10.1111/j.1399-3054.2007.01020.x

Hall RD, Brouwer ID, Fitzgerald MA (2008) Plant metabolomics and its potential application for human nutrition. Physiol Plant 132(2):162–175. doi:10.1111/j.1399-3054.2007.00989.x

Han B, Huang X (2013) Sequencing-based genome-wide association study in rice. Curr Opin Plant Biol. doi:10.1016/j.pbi.2013.03.006

Hank JA, Surfus JE, Gan J, Ostendorf A, Gillies SD, Sondel PM (2003) Determination of peak serum levels and immune response to the humanized anti-ganglioside antibody-interleukin-2 immunocytokine. Methods Mol Med 85:123–131. doi:10.1385/1-59259-380-1:123

Hayashi M, Imaizumi-Anraku H, Akao S, Kawaguchi M (2000) Nodule organogenesis in *Lotus japonicus*. J Plant Res 113(1112):489–495

Hayashi M, Miyahara A, Sato S, Kato T, Yoshikawa M, Taketa M, Pedrosa A, Onda R, Imaizumi-Anraku H, Bachmair A, Sandal N, Stougaard J, Murooka Y, Tabata S, Kawasaki S, Kawaguchi M, Harada K (2001) Construction of a genetic linkage map of the model legume *Lotus japonicus* using an intraspecific F2 population. DNA Res 8(6):301–310

Huang X, Wei X, Sang T, Zhao Q, Feng Q, Zhao Y, Li C, Zhu C, Lu T, Zhang Z, Li M, Fan D, Guo Y, Wang A, Wang L, Deng L, Li W, Lu Y, Weng Q, Liu K, Huang T, Zhou T, Jing Y, Lin Z, Buckler ES, Qian Q, Zhang QF, Li J, Han B (2010) Genome-wide association studies of 14 agronomic traits in rice landraces. Nat Genet 42(11):961–967. doi:10.1038/ng.695

Iijima Y, Nakamura Y, Ogata Y, Tanaka K, Sakurai N, Suda K, Suzuki T, Suzuki H, Okazaki K, Kitayama M, Kanaya S, Aoki K, Shibata D (2008) Metabolite annotations based on the integration of mass spectral information. Plant J 54(5):949–962. doi:10.1111/j.1365-313X.2008.03434.x

Kawaguchi M, Motomura T, Imaizumi-Anraku H, Akao S, Kawasaki S (2001) Providing the basis for genomics in *Lotus japonicus*: the accessions Miyakojima and Gifu are appropriate crossing partners for genetic analyses. Mol Genet Genomics 266(2):157–166

Klein MA, Grusak MA (2009) Identification of nutrient and physical seed trait QTL in the model legume *Lotus japonicus*. Genome 52(8):677–691. doi:10.1139/g09-039

Kojima M, Kamada-Nobusada T, Komatsu H, Takei K, Kuroha T, Mizutani M, Ashikari M, Ueguchi-Tanaka M, Matsuoka M, Suzuki K, Sakakibara H (2009) Highly sensitive and high-throughput analysis of plant hormones using MS-probe modification and liquid chromatography-tandem mass spectrometry: an application for hormone profiling in *Oryza sativa*. Plant Cell Physiol 50(7):1201–1214. doi:10.1093/pcp/pcp057

Lim SY, Jang JH, Surh YJ (2003) Induction of cyclooxygenase-2 and peroxisome proliferator-activated receptor-gamma during nitric oxide-induced apoptotic PC12 cell death. Ann NY Acad Sci 1010:648–658

Lisec J, Meyer RC, Steinfath M, Redestig H, Becher M, Witucka-Wall H, Fiehn O, Torjek O, Selbig J, Altmann T, Willmitzer L (2008) Identification of metabolic and biomass QTL in *Arabidopsis thaliana* in a parallel analysis of RIL and IL populations. Plant J 53(6):960–972. doi:10.1111/j.1365-313X.2007.03383.x

Marshall AG, Hendrickson CL (2008) High-resolution mass spectrometers. Annu Rev Anal Chem 1:579–599. doi:10.1146/annurev.anchem.1.031207.112945

Matsuda F, Hirai MY, Sasaki E, Akiyama K, Yonekura-Sakakibara K, Provart NJ, Sakurai T, Shimada Y, Saito K (2010) AtMetExpress development: a phytochemical atlas of Arabidopsis development. Plant Physiol 152(2):566–578. doi:10.1104/pp.109.148031

Matsuda F, Yonekura-Sakakibara K, Niida R, Kuromori T, Shinozaki K, Saito K (2009) MS/MS spectral tag-based annotation of non-targeted profile of plant secondary metabolites. Plant J 57(3):555–577. doi:10.1111/j.1365-313X.2008.03705.x)

Morant AV, Bjarnholt N, Kragh ME, Kjaergaard CH, Jorgensen K, Paquette SM, Piotrowski M, Imberty A, Olsen CE, Moller BL, Bak S (2008) The beta-glucosidases responsible for bioactivation of hydroxynitrile glucosides in *Lotus japonicus*. Plant Physiol 147(3):1072–1091. doi:10.1104/pp.107.109512

Mylonas R, Mauron Y, Masselot A, Binz PA, Budin N, Fathi M, Viette V, Hochstrasser DF, Lisacek F (2009) X-Rank: a robust algorithm for small molecule identification using tandem mass spectrometry. Anal Chem 81(18):7604–7610. doi:10.1021/ac900954d

Nakabayashi A, Kamei N, Sunagawa T, Suzuki O, Ohkawa S, Kodama A, Kamei G, Ochi M (2013a)

In vivo bioluminescence imaging of magnetically targeted bone marrow-derived mesenchymal stem cells in skeletal muscle injury model. J Orthop Res 31(5):754–759. doi:10.1002/jor.22282

Nakabayashi R, Sawada Y, Yamada Y, Suzuki M, Hirai MY, Sakurai T, Saito K (2013b) Combination of liquid chromatography-Fourier transform ion cyclotron resonance-mass spectrometry with 13C-labeling for chemical assignment of sulfur-containing metabolites in onion bulbs. Anal Chem 85(3):1310–1315. doi:10.1021/ac302733c

Saito K, Hirai MY, Yonekura-Sakakibara K (2008) Decoding genes with coexpression networks and metabolomics - 'majority report by precogs'. Trends Plant Sci 13(1):36–43. doi:10.1016/j.tplants.2007.10.006

Saito S, Motawia MS, Olsen CE, Moller BL, Bak S (2012) Biosynthesis of rhodiocyanosides in *Lotus japonicus*: rhodiocyanoside A is synthesized from (Z)-2-methylbutanaloxime via 2-methyl-2-butenenitrile. Phytochemistry 77:260–267. doi:10.1016/j.phytochem.2012.01.020

Sawada Y, Akiyama K, Sakata A, Kuwahara A, Otsuki H, Sakurai T, Saito K, Hirai MY (2009) Widely targeted metabolomics based on large-scale MS/MS data for elucidating metabolite accumulation patterns in plants. Plant Cell Physiol 50(1):37–47. doi:10.1093/pcp/pcn183

Sawada Y, Hirai MY (2013) Integrated LC-MS/MS system for plant metabolomics. Comput Struct Biotechnol J 4(5):e201301011. doi:http://dx.doi.org/10.5936/csbj.201301011

Sawada Y, Kinoshita K, Akashi T, Aoki T, Ayabe S (2002) Key amino acid residues required for aryl migration catalysed by the cytochrome P450 2-hydroxyisoflavanone synthase. Plant J 31(5):555–564

Sawada Y, Nakabayashi R, Yamada Y, Suzuki M, Sato M, Sakata A, Akiyama K, Sakurai T, Matsuda F, Aoki T, Hirai MY, Saito K (2012) RIKEN tandem mass spectral database (ReSpect) for phytochemicals: a plant-specific MS/MS-based data resource and database. Phytochemistry 82:38–45. doi:10.1016/j.phytochem.2012.07.007

Sawai S, Akashi T, Sakurai N, Suzuki H, Shibata D, Ayabe SI, Aoki T (2006a) Plant lanosterol synthase: divergence of the sterol and triterpene biosynthetic pathways in eukaryotes. Plant Cell Physiol 47(5):673–677

Sawai S, Shindo T, Sato S, Kaneko T, Tabata S, Ayabe S, Aoki T (2006b) Functional and structural analysis of genes encoding oxidosqualene cyclases of *Lotus japonicus*. Plant Sci 170(2):247–257

Schauer N, Fernie AR (2006) Plant metabolomics: towards biological function and mechanism. Trends Plant Sci 11(10):508–516. doi:10.1016/j.tplants.2006.08.007

Shimada N, Aoki T, Sato S, Nakamura Y, Tabata S, Ayabe S (2003) A cluster of genes encodes the two types of chalcone isomerase involved in the biosynthesis of general flavonoids and legume-specific 5-deoxy(iso)flavonoids in *Lotus japonicus*. Plant Physiol 131(3):941–951. doi:10.1104/pp.004820

Shimada N, Nakatsuka T, Nishihara M, Yamamura S, Ayabe S, Aoki T (2006) Isolation and characterization of a cDNA encoding polyketide reductase in *Lotus japonicus*. Plant Biotechnol 23(5):509–513

Shimada N, Sasaki R, Sato S, Kaneko T, Tabata S, Aoki T, Ayabe S (2005) A comprehensive analysis of six dihydroflavonol 4-reductases encoded by a gene cluster of the *Lotus japonicus* genome. J Exp Bot 56(419):2573–2585. doi:10.1093/jxb/eri251

Shimada N, Sato S, Akashi T, Nakamura Y, Tabata S, Ayabe S, Aoki T (2007) Genome-wide analyses of the structural gene families involved in the legume-specific 5-deoxyisoflavonoid biosynthesis of Lotus *japonicus*. DNA Res 14(1):25–36. doi:10.1093/dnares/dsm004

Smith CA, O'Maille G, Want EJ, Qin C, Trauger SA, Brandon TR, Custodio DE, Abagyan R, Siuzdak G (2005) METLIN: a metabolite mass spectral database. Ther Drug Monit 27(6):747–751

Soh Y, Shin MH, Lee JS, Jang JH, Kim OH, Kang H, Surh YJ (2003) Oxidative DNA damage and glioma cell death induced by tetrahydropapaveroline. Mutat Res 544(2–3):129–142

Steele CL, Gijzen M, Qutob D, Dixon RA (1999) Molecular characterization of the enzyme catalyzing the aryl migration reaction of isoflavonoid biosynthesis in soybean. Arch Biochem Biophys 367(1):146–150. doi:10.1006/abbi.1999.1238

Sumner LW, Mendes P, Dixon RA (2003) Plant metabolomics: large-scale phytochemistry in the functional genomics era. Phytochemistry 62(6):817–836

Suzuki H, Sasaki R, Ogata Y, Nakamura Y, Sakurai N, Kitajima M, Takayama H, Kanaya S, Aoki K, Shibata D, Saito K (2008) Metabolic profiling of flavonoids in *Lotus japonicus* using liquid chromatography fourier transform ion cyclotron resonance mass spectrometry. Phytochemistry 69(1):99–111. doi:10.1016/j.phytochem.2007.06.017

Weckwerth W (2003) Metabolomics in systems biology. Annu Rev Plant Biol 54:669–689. doi:10.1146/annurev.arplant.54.031902.135014

Wink M (1988) Plant-breeding—importance of plant secondary metabolites for protection against pathogens and herbivores. Theor Appl Genet 75(2):225–233

Wishart DS, Knox C, Guo AC, Eisner R, Young N, Gautam B, Hau DD, Psychogios N, Dong E, Bouatra S, Mandal R, Sinelnikov I, Xia J, Jia L, Cruz JA, Lim E, Sobsey CA, Shrivastava S, Huang P, Liu P, Fang L, Peng J, Fradette R, Cheng D, Tzur D, Clements M, Lewis A, De Souza A, Zuniga A, Dawe M, Xiong Y, Clive D, Greiner R, Nazyrova A, Shaykhutdinov R, Li L, Vogel HJ, Forsythe I (2009) HMDB: a knowledgebase for the human metabolome. Nucleic Acids Res 37(Database issue):D603–D610. doi:10.1093/nar/gkn810

Wishart DS, Tzur D, Knox C, Eisner R, Guo AC, Young N, Cheng D, Jewell K, Arndt D, Sawhney S, Fung C,

Nikolai L, Lewis M, Coutouly MA, Forsythe I, Tang P, Shrivastava S, Jeroncic K, Stothard P, Amegbey G, Block D, Hau DD, Wagner J, Miniaci J, Clements M, Gebremedhin M, Guo N, Zhang Y, Duggan GE, Macinnis GD, Weljie AM, Dowlatabadi R, Bamforth F, Clive D, Greiner R, Li L, Marrie T, Sykes BD, Vogel HJ, Querengesser L (2007) HMDB: the Human Metabolome Database. Nucleic Acids Res 35(Database issue):D521–D526. doi:10.1093/nar/gkl923

Xu S, Luo Y, Cai Z, Cao X, Hu X, Yang J, Luo D (2013) Functional diversity of CYCLOIDEA-like TCP genes in the control of zygomorphic flower development in *Lotus japonicus*. J Integr Plant Biol 55(3):221–231. doi:10.1111/j.1744-7909.2012.01169.x

Yamazaki M, Mochida K, Asano T, Nakabayashi R, Chiba M, Nirin U, Yamazaki Y, Goodenowe DB, Sankawa U, Yoshida T, Toyoda A, Totoki Y, Sakaki Y, Gongora-Castillo E, Buell CR, Sakurai T, Saito K (2013) Coupling deep transcriptome analysis with untargeted metabolic profiling in ophiorrhiza pumila to further the understanding of the biosynthesis of the anti-cancer alkaloid camptothecin and anthraquinones. Plant Cell Physiol. doi:10.1093/pcp/pct040

A Tutorial on *Lotus japonicus* Transcriptomic Tools

17

Jerome Verdier, Kaustav Bandyopadhyay, and Michael Udvardi

Abstract

Lotus japonicus is widely used as a model species for legume biology. The rapid growth of transcriptomic data available for this species represents an asset for the understanding of molecular mechanisms underlying different processes of legume development. In this chapter, we review the history of *L. japonicus* transcriptomic studies before focusing on two tools developed recently to store, visualize, and analyze transcriptomic data: the *L. japonicus* gene expression atlas (LjGEA) and the legume gene regulatory network Web server (LegumeGRN). A description of the features available on these Web servers is provided with a tutorial describing their uses. These tools are already connected to available transcriptomic data from two other model legumes, *Medicago truncatula* and *Glycine max*, which enables comparative genomic studies.

17.1 Early Transcriptomic Studies

Transcriptomic studies in Lotus were initiated during the first half of the last decade. The earliest approach was identification of expressed sequence tags (ESTs) (Endo et al. 2000, 2002), which enabled development of cDNA arrays in nylon filters. These arrays were used to quantify labeled cDNAs derived from RNA of organs of interest, via hybridization. Being a model for nitrogen fixation research, it is not surprising that the first efforts were focused on nodule and root tissues. In 2002, Colebatch et al. (2002) identified changes in the expression of 83 genes between root and nodule organs in an array of about 2,300 cDNA clones. By 2004, the number of arrayed probes had increased to around 5,000 cDNA clones derived from nodules (Colebatch et al. 2004) and

J. Verdier (✉) · K. Bandyopadhyay · M. Udvardi
Plant Biology Division, The Samuel Roberts Noble Foundation, 2510 Sam Noble Parkway, Ardmore, OK 73401, USA
e-mail: javerdier@icloud.com

K. Bandyopadhyay
e-mail: kbandyopadhyay@noble.org

M. Udvardi
e-mail: mudvardi@noble.org

J. Verdier
Shanghai Center for Plant Stress Biology (PSC), Shanghai Institutes of Biological Sciences, Chinese Academy of Sciences, 3888 Chenhua Road, Shanghai 201602, People's Republic of China

S. Tabata and J. Stougaard (eds.), *The Lotus japonicus Genome*, Compendium of Plant Genomes,
DOI 10.1007/978-3-662-44270-8_17,

18,144 cDNA clones derived from different organs (Kouchi et al. 2004), which allowed identification of 860 and 1,076 genes, respectively, that were differentially expressed between uninfected roots and *Mesorhizobium loti*-infected roots or nodules. In 2005, a transcriptome analysis comparing uninfected and nodulating roots was performed by serial analysis of gene expression (SAGE) (Asamizu et al. 2005). SAGE libraries from these two tissues contained, respectively, 85,482 and 80,233 SAGE tags, which are short cDNA fragments unique to each transcript used to determine transcript level based on their frequency. This analysis revealed more than 800 genes differentially expressed between infected and uninfected roots. In 2008, a draft genome sequence of Lotus was published (Sato et al. 2008), which enabled production of a new generation of gene chip containing probes for most genes.

17.2 The Affymetrix *L. japonicus* Genechip

The *Lotus japonicus* GeneChip is an Affymetrix custom array, called A-AFFY-90. It contains 52,749 *Lotus* and 8,710 *Mesorhizobium loti* (*M. loti*)-derived probe sets, each representing a known or predicted open reading frame (ORF). Each probe set has a unique identifier consisting of a gene or clone name, with one of the following suffixes: "_at" meaning antisense target, "_st" for sense target (i.e., controls), "_s_at" for probe set matching multiple transcripts, "_a_at" for genes coding potential alternatively spliced transcripts and "_x_at" when it was not possible to design unique probes for the transcript (i.e., this probe set is not unique for the gene).

To estimate the coverage of the Affymetrix GeneChip, we mapped the 47,486 gene sequences from the *L. japonicus* Gene Index (LjGI version 6) representing tentative transcripts (TC) and singleton expressed sequence tags (ESTs) (http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/gimain.pl?gudb=l_japonicus) to probe sets present on the chip. This revealed that 77 % (36,599/47,486) of the genes encoding putative proteins in LjGI were represented on the Affymetrix GeneChip.

17.2.1 Annotations

Mapping of Affymetrix probe set sequences to sequences obtained from the LjGI v6 and from the annotated genome (Sato et al. 2008) via reciprocal BLAST enabled us to re-annotate probe set IDs using reciprocal BLAST (Verdier et al. 2013). Each Lotus gene/probe set was assigned to a different bincode classification of the Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa and Goto 2000) and the gene ontology (GO) (Ashburner et al. 2000) consortium (Verdier et al. 2013). KEGG and GO annotations provide information about genes, gene products, and metabolic pathways that may be useful to describe putative gene function. For across-species studies, identification of soybean and Medicago orthologs for each probe set was performed using reciprocal BLASTP and the best hits method combined with soft filtering and Smith-Waterman alignment options (Moreno-Hagelsieb and Latimer 2008).

17.2.2 TFs, Transporters, and LSGs

After refining probe set annotations, 1,489 probe sets were assigned to putative transcription factor genes (from a total of 1,616 TFs identified in the Lotus genome) and 597 probe sets were assigned to putative transporter genes (from a total of 1,087 genes in the Lotus genome). In the same way that TF and transporter genes can provide insight into special features of particular organs, genes that are specific to a family of plants can provide insight into unique biological aspects of that family. Previous analysis of the Lotus genome identified 1,190 putative legume-specific genes (LSGs) (Sato et al. 2008), which matched 729 probe sets present on the chip.

17.3 Update of Published "Affymetrix" Transcriptomes

Recent years of research on *L. japonicus* have been marked by a rapid growth of publicly available transcriptomic data. Databases of gene expression, such as Array Express (Parkinson et al. 2005), provide raw data from sets of transcriptomic experiments covering all organs at various stages of plant development (E-MEXP-1726; Verdier et al. 2013), salt acclimatization experiments on Gifu (E-MEXP-1204; Sanchez et al. 2008) and on other Lotus species (E-MEXP-2344; Sanchez et al. 2011), analysis of a plastidic glutamine synthetase mutant in drought experiments (E-MEXP-2690; Díaz et al. 2010), analysis of arbuscular mycorrhizal symbiosis (Guether et al. 2009), a detailed dissection of nodulation mechanisms using treatments on wild-type and mutant plants (E-TABM-715; Høgslund et al. 2009), analysis of cellular stress following water deprivation (E-MEXP-3710; Betti et al. 2012) and regulation of photorespiration in wild-type and *Ljgln2-2* mutant lines (E-MEXP-3603; Pérez-Delgado et al. 2013). Details about most of the above-mentioned experiments are provided in Table 17.1.

A subset of 24 experiments covering all major plant organs, i.e., stem, petiole, leaf, flower, nodule (four developmental stages), root (eight conditions with or without rizhobia inoculation at various stages), pod (three developmental stages), and seed (five developmental stages), were selected to identify 2,949 genes/probe sets specific or preferentially expressed in each organs (Verdier et al. 2013). Genes that are expressed specifically or preferentially in a particular organ can provide insight into specialized processes in these organs. Previous findings in Arabidopsis showed a high correlation between organ identity and gene expression, where transcript modulation represents a transcriptional signature specific to each individual organ (Aceituno et al. 2008). Organ-specific genes represent good candidates for functional genomic study of developmental process. Interestingly, approximately 39 % of organ-specific genes were nodule-specific, 21 % flower-specific, and 16 % seed-specific, with fewer genes specific to other organs. This may reflect a higher degree of functional specialization in these organs. It has also been reported that a large proportion of TFs and transporters are differentially regulated across all tissues and that their maxima of expression is generally associated with a specific organ, reflecting their specialization in certain organs (Verdier et al. 2013). The gene sets showing organ-specific expression are valuable information for biotechnological use.

Promoters of organ-specific genes can be used to direct expression of transgenes in specific organs and/or at certain developmental stages, which may be crucial for the regulation of genes of interest for research or biotechnological applications. Using the same subset of experiments described above, Verdier et al. (2013) identified 71 stably expressed genes across the different organs, which represent potential reference genes for normalization of expression data from qRT-PCR or probe-hybridization approaches.

17.4 *Lotus japonicus* Gene Expression Atlas (LjGEA) Web Server

To make the most of transcriptomic data, it is important that the research community has free access to all the published data and to a platform upon which they can compare between experiments and explore, analyze, and visualize the data. To meet these needs, the LjGEA Web server was developed. It is a centralized platform for analyzing Lotus transcriptomic data. This open-access server is hosted by the Samuel Roberts Noble Foundation (http://ljgea.noble.org/). Expanding beyond the number of experiments described above, the Web server currently hosts gene expression data from 237 GeneChips from 83 different experiments, covering a broad range of developmental and environmental conditions. The LjGEA utilizes the architecture and tools of the *Medicago truncatula* Gene Expression Atlas (MtGEA) server (He et al. 2009). To normalize

Table 17.1 Detailed description of Lotus published transcriptomes present in LjGEA Web server

	LjGEA ID	Description	Replicate	Organ	Type of analysis	Treatment	Genotype	Culture system	Reference
Plant organ	Fl	Flower	3	Flower	Standard	Fully open flower	*Lotus japonicus MG22*	Turface +Clay	Verdier et al. (2013)
Plant organ	Leaf	Leaf	3	Leaf	Standard	28-day-old plants	*Lotus japonicus MG22*	Turface +Clay	Verdier et al. (2013)
Plant organ	Root0 h	Root	3	Nodule susceptible zone	Standard	Low nitrogen (0.5 mM KNO_3) collected pre-inoculation with *M. loti* MAFF1021	*Lotus japonicus MG22*	Turface +Clay	Verdier et al. (2013)
Plant organ	Nod21	Nod 21dpi	3	Nodule	Standard	Low nitrogen (0.5 mM KNO_3) collected 21 days post-inoculation with *M. loti* MAFF1021	*Lotus japonicus MG22*	Turface +Clay	Verdier et al. (2013)
Plant organ	Pt	Petiole	3	Petiole	Standard	28-day-old plants	*Lotus japonicus MG22*	Turface +Clay	Verdier et al. (2013)
Plant organ	Root	Root	3	Root	Standard	28-day-old plants	*Lotus japonicus MG22*	Turface +Clay	Verdier et al. (2013)
Plant organ	Stem	Stem	3	stem	Standard	28-day-old plants	*Lotus japonicus MG22*	Turface +Clay	Verdier et al. (2013)
Pod development	Pod10	Pod 10dap	3	Pod + seed	Standard	Pod development time course	*Lotus japonicus MG22*	Turface +Clay	Verdier et al. (2013)
Pod development	Pod14	Pod 14dap	3	Pod + seed	Standard	Pod development time course	*Lotus japonicus MG22*	Turface +Clay	Verdier et al. (2013)
Pod development	Pod20	Pod 20dap	3	pod + seed	Standard	Pod development time course	*Lotus japonicus MG22*	Turface +Clay	Verdier et al. (2013)
Seed development	Seed10d	Seed 10dap	3	Seed	Standard	Seed development time course	*Lotus japonicus MG22*	Turface +Clay	Verdier et al. (2013)
Seed development	Seed12d	Seed 12dap	3	Seed	Standard	Seed development time course	*Lotus japonicus MG22*	Turface +Clay	Verdier et al. (2013)
Seed development	Seed14d	Seed 14dap	3	Seed	Standard	Seed development time course	*Lotus japonicus MG22*	Turface +Clay	Verdier et al. (2013)
Seed development	Seed16d	Seed 16dap	3	Seed	Standard	Seed development time course	*Lotus japonicus MG22*	Turface +Clay	Verdier et al. (2013)
Seed development	Seed20d	Seed 20dap	3	Seed	Standard	Seed development time course	*Lotus japonicus MG22*	Turface +Clay	Verdier et al. (2013)
Species	L.burttii_Ctrol	L.burttii_Ctrol	3	Shoot	Ecotype	32-day-old plant—whole shoot	*Lotus burttii*	Soil	Sanchez et al. (2011)
Species	L.corniculatus_Ctrol	L.corniculatus_Ctrol	3	Shoot	Ecotype	32-day-old plant—whole shoot	*Lotus corniculatus*	Soil	Sanchez et al. (2011)

(continued)

Table 17.1 (continued)

	LjGEA ID	Description	Replicate	Organ	Type of analysis	Treatment	Genotype	Culture system	Reference
Species	L.filicaulis_Ctrol	L.filicaulis_Ctrol	3	shoot	Ecotype	32-day-old plant—whole shoot	*Lotus filicaulis*	Soil	Sanchez et al. (2011)
Species	L.glaber_Ctrol	L.glaber_Ctrol	3	Shoot	Ecotype	32-day-old plant—whole shoot	*Lotus glaber*	Soil	Sanchez et al. (2011)
Species	L.japonicus_Gifu_Ctrol	L.japonicus Gifu_Ctrol	3	Shoot	Ecotype	32-day-old plant—whole shoot	*Lotus japonicus Gifu*	Soil	Sanchez et al. (2011)
Species	L. japonicus_MG20_Ctrol	L.japonicus MG20_Ctrol	3	Shoot	Ecotype	32-day-old plant—whole shoot	*Lotus japonicus MG20*	Soil	Sanchez et al. (2011)
Species	L.uliginosus_Ctrol	L.uliginosus_Ctrol	3	Shoot	Ecotype	32-day-old plant—whole shoot	*Lotus uliginosus*	Soil	Sanchez et al. (2011)
Drought	L.burttii_Salt	L.burttii_Salt	3	Shoot	Ecotype—treatment	32-day-old plant—whole shoot—NaCl gradual acclimatation	*Lotus burttii*	Soil	Sanchez et al. (2011)
Drought	L.corniculatus_Salt	L.corniculatus_Salt	3	Shoot	Ecotype—treatment	32-day-old plant—whole shoot—NaCl gradual acclimatation	*Lotus corniculatus*	Soil	Sanchez et al. (2011)
Drought	L.filicaulis_Salt	L.filicaulis_Salt	3	Shoot	Ecotype—treatment	32-day-old plant—whole shoot—NaCl gradual acclimatation	*Lotus filicaulis*	Soil	Sanchez et al. (2011)
Drought	L.glaber_Salt	L.glaber_Salt	3	Shoot	Ecotype—treatment	32-day-old plant—whole shoot—NaCl gradual acclimatation	*Lotus glaber*	Soil	Sanchez et al. (2011)
Drought	L.japonicus_Gifu_Salt	L.japonicus Gifu_Salt	3	Shoot	Ecotype—treatment	32-day-old plant—whole shoot—NaCl gradual acclimatation	*Lotus japonicus Gifu*	Soil	Sanchez et al. (2011)
Drought	L.japonicus_MG20_Salt	L.japonicus MG20_Salt	3	Shoot	Ecotype—treatment	32-day-old plant—whole shoot—NaCl gradual acclimatation	*Lotus japonicus MG23*	Soil	Sanchez et al. (2011)
Drought	L.uliginosus_Salt	L.uliginosus_Salt	3	shoot	Ecotype—treatment	32-day-old plant—whole shoot—NaCl gradual acclimatation	*Lotus uliginosus*	Soil	Sanchez et al. (2011)
Drought	Ljgln2_2_Control	Ljgln2-2 Control1	3	Shoot	transgenics	35-day-old plant—whole shoot—control	*Gifu— gln2-2*	Vermiculite + Sand	Díaz et al. (2010)
Drought	Ljgln2_2_Drought	Ljgln2-2 Drought1	2	Shoot	transgenics	39-day-old plant—whole shoot—4 days of drought conditions	*Gifu— gln2-2*	Vermiculite + Sand	Díaz et al. (2010)
Drought	WT_control	WT control1	3	Shoot	Treatment	35-day-old plant—whole shoot—control	*Lotus japonicus Gifu*	Vermiculite + Sand	Díaz et al. (2010)

(continued)

Table 17.1 (continued)

	LjGEA ID	Description	Replicate	Organ	Type of analysis	Treatment	Genotype	Culture system	Reference
Drought	WT_Drought	WT Drought1	2	Shoot	Treatment	39-day-old plant - whole shoot—4 days of drought conditions	*Lotus japonicus Gifu*	Vermiculite + Sand	Díaz et al. (2010)
Drought	Shoot_0 mM_ sodiumChloride	Shoot 0mM sodiumChloride	3	Shoot	Treatment	32-day-old plant—whole shoot—sodium treatment	*Lotus japonicus Gifu*	Soil	Sanchez et al. (2008)
Drought	Shoot_100 mM_ sodiumChloride_ Gradual	Shoot 100mM sodiumChloride Gradual	3	Shoot	Treatment	32-day-old plant—whole shoot—sodium treatment	*Lotus japonicus Gifu*	Soil	Sanchez et al. (2008)
Drought	Shoot_150 mM_ sodiumChloride_ Gradual	Shoot 150mM sodiumChloride Gradual	3	Shoot	Treatment	32-day-old plant—whole shoot—sodium treatment	*Lotus japonicus Gifu*	Soil	Sanchez et al. (2008)
Drought	Shoot_25 mM_sodium Chloride_Initial	Shoot 25mM sodiumChloride Initial	3	Shoot	Treatment	32-day-old plant—whole shoot—sodium treatment	*Lotus japonicus Gifu*	Soil	Sanchez et al. (2008)
Drought	Shoot_50 mM_sodium Chloride_Gradual	Shoot 50mM sodiumChloride Gradual	3	Shoot	Treatment	32-day-old plant—whole shoot—sodium treatment	*Lotus japonicus Gifu*	Soil	Sanchez et al. (2008)
Drought	Shoot_50 mM_sodium Chloride_Initial	Shoot 50mM sodiumChloride Initial	3	Shoot	Treatment	32-day-old plant—whole shoot—sodium treatment	*Lotus japonicus Gifu*	Soil	Sanchez et al. (2008)
Drought	Shoot_75 mM_sodium Chloride_Initial	Shoot 75mM sodiumChloride Initial	3	Shoot	Treatment	32-day-old plant—whole shoot—sodium treatment	*Lotus japonicus Gifu*	Soil	Sanchez et al. (2008)
Nodulation	cyclops_root_3w_uninocul	cyclops root 3w uninocul	1	Root	Mutant	3-week-old uninoculated roots	*Gifu—cyclops*	Sterilized Leca	Høgslund et al. (2009)
Nodulation	cyclops_root_3w_inocul	Cyclops root 3w inocul	1	Root	Mutant	3-week-old inoculated roots	*Gifu—cyclops*	Sterilized Leca	Høgslund et al. (2009)
Nodulation	cyclops_root_nodule_ 3w_inocul21	Cyclops root + nodule 3w inocul21	1	Root + nodule	Mutant	3-week-old uninoculated roots—21 days post-R7A inoculation	*Gifu—cyclops*	Sterilized Leca	Høgslund et al. (2009)
Nodulation	har1_root_3w_inocul3	har1 root 3w inocul3	2	Root	Mutant	3-week-old uninoculated roots—3 days post-R7A inoculation	*Gifu—har1*	Sterilized Leca	Høgslund et al. (2009)
Nodulation	har1_root_3w_uninocul	har1 root 3w uninocul	2	Root	Mutant	3-week-old uninoculated roots	*Gifu—har3*	Sterilized Leca	Høgslund et al. (2009)

(continued)

Table 17.1 (continued)

	LjGEA ID	Description	Replicate	Organ	Type of analysis	Treatment	Genotype	Culture system	Reference
Nodulation	har1_shoot_3w_inocul3	har1 shoot 3w inocul3	3	Shoot	Mutant	3-week-old uninoculated roots—3 days post-R7A inoculation	*Gifu—har5*	Sterilized Leca	Høgslund et al. (2009)
Nodulation	har1_shoot_3w_uninocul	har1 shoot 3w uninocul	2	Shoot	Mutant	3-week-old uninoculated roots	*Gifu—har8*	Sterilized Leca	Høgslund et al. (2009)
Nodulation	nfr1_rootSZ_3w_inocul1	nfr1 rootSZ 3w inocul1	3	Root susceptible zone	Mutant	3-week-old uninoculated roots—1 day post-R7A inoculation	*Gifu—nfr1*	Sterilized Leca	Høgslund et al. (2009)
Nodulation	nfr1_rootSZ_3w_uninocul	nfr1 rootSZ 3w uninocul	2	Root susceptible zone	Mutant	3-week-old uninoculated roots	*Gifu—nfr1*	Sterilized Leca	Høgslund et al. (2009)
Nodulation	nfr5_rootSZ_3w_inocul1	nfr5 rootSZ 3w inocul1	2	Root susceptible zone	Mutant	3-week-old uninoculated roots—1 day post-R7A inoculation	*Gifu—nfr5*	Sterilized Leca	Høgslund et al. (2009)
Nodulation	nfr5_rootSZ_3w_uninocul	nfr5 rootSZ 3w uninocul1	3	Root susceptible zone	Mutant	3-week-old uninoculated roots—1 day post-R7A inoculation	*Gifu—nfr5*	Sterilized Leca	Høgslund et al. (2009)
Nodulation	nin_rootSZ_3w_inocul1	nin rootSZ 3w inocul1	3	Root susceptible zone	Mutant	3-week-old uninoculated roots—1 day post-R7A inoculation	*Gifu—nin*	Sterilized Leca	Høgslund et al. (2009)
Nodulation	nin_rootSZ_3w_uninocul	nin rootSZ 3w uninocul	3	Root susceptible zone	Mutant	3-week-old uninoculated roots	*Gifu—nin*	Sterilized Leca	Høgslund et al. (2009)
Nodulation	nup133_rootSZ_ 3w_inocul1	nup133 rootSZ 3w inocul1	3	Root susceptible zone	Mutant	3-week-old uninoculated roots—1 day post-R7A inoculation	*Gifu -nup133*	Sterilized Leca	Høgslund et al. (2009)
Nodulation	nup133_rootSZ_ 3w_uninocul	nup133 rootSZ 3w uninocul	3	Root susceptible zone	Mutant	3wk-old uninoculated roots	*Gifu -nup133*	Sterilized Leca	Høgslund et al. (2009)
Nodulation	sen1_nodule_3w_inocul21	sen1 nodule 3w inocul21	3	Nodule	Mutant	3-week-old uninoculated roots—21 days post-R7A inoculation	*Gifu -sen1*	Sterilized Leca	Høgslund et al. (2009)
Nodulation	sen1_root_3w_uninocul	sen1 root 3w uninocul	3	Root	Mutant	3-week-old uninoculated roots	*Gifu -sen1*	Sterilized Leca	Høgslund et al. (2009)

(continued)

Table 17.1 (continued)

	LjGEA ID	Description	Replicate	Organ	Type of analysis	Treatment	Genotype	Culture system	Reference
Nodulation	sst1_nodule_3w_inocul21	sst1 nodule 3w inocul21	3	Nodule	Mutant	3-week-old uninoculated roots—21 days post-R7A inoculation	*Gifu -sst1*	Sterilized Leca	Høgslund et al. (2009)
Nodulation	sst1_root_3w_uninocul	sst1 root 3w uninocul	2	Root	Mutant	3-week-old uninoculated roots	*Gifu -sst1*	Sterilized Leca	Høgslund et al. (2009)
Nodulation	WT_root_3w_uninocul	WT root 3w uninocul	3	Root	Standard	3-week-old uninoculated roots	*Lotus japonicus Gifu*	Sterilized Leca	Høgslund et al. (2009)
Nodulation	WT_root_tip_3w_uninocul	WT root tip 3w uninocul	3	Root tip	Standard	3-week-old uninoculated roots	*Lotus japonicus Gifu*	Sterilized Leca	Høgslund et al. (2009)
Nodulation	WT_rootSZ_3w_uninocul	WT rootSZ 3w uninocul	3	Root susceptible zone	Standard	3-week-old uninoculated roots	*Lotus japonicus Gifu*	Sterilized Leca	Høgslund et al. (2009)
Nodulation	WT_shoot_3w_uninocul	WT shoot 3w uninocul	3	Shoot	Standard	3-week-old uninoculated roots	*Lotus japonicus Gifu*	Sterilized Leca	Høgslund et al. (2009)
Nodulation	WT_flower_13w_ 5 mM_nitrate	WT flower 13w 5 mM nitrate	3	Flower	Treatment	13-week-old plant—flower—nitrate treatment	*Lotus japonicus Gifu*	Sterilized Leca	Høgslund et al. (2009)
Nodulation	WT_leaf_6w_ 5 mM_nitrate	WT leaf 6w 5 mM nitrate	3	Leaf	Treatment	6-week-old plant—leaf—nitrate treatment	*Lotus japonicus Gifu*	Sterilized Leca	Høgslund et al. (2009)
Nodulation	WT_nodule_3w_inocul14	WT nodule 3w inocul14	3	Nodule	Treatment	3-week-old uninoculated roots—14 days post-R7A inoculation	*Lotus japonicus Gifu*	Sterilized Leca	Høgslund et al. (20099)
Nodulation	WT_nodule_3w_inocul21	WT nodule 3w inocul21	3	Nodule	Treatment	3-week-old uninoculated roots—21 days post-R7A inoculation	*Lotus japonicus Gifu*	Sterilized Leca	Høgslund et al. (2009)
Nodulation	WT_root_3w_ 5 mM_nitrate	WT root 3w 5 mM nitrate	3	Root	Treatment	3-week-old plants—nitrate treatment	*Lotus japonicus Gifu*	Sterilized Leca	Høgslund et al. (2009)
Nodulation	WT_root_3w_inocul1	WT root 3w inocul1	3	Root	Treatment	3-week-old uninoculated roots—1 day post-R7A inoculation	*Lotus japonicus Gifu*	Sterilized Leca	Høgslund et al. (2009)
Nodulation	WT_root_3w_inocul3	WT root 3w inocul3	2	Root	Treatment	3-week-old uninoculated roots—3 days post-R7A inoculation	*Lotus japonicus Gifu*	Sterilized Leca	Høgslund et al. (2009)

(continued)

Table 17.1 (continued)

	LjGEA ID	Description	Replicate	Organ	Type of analysis	Treatment	Genotype	Culture system	Reference
Nodulation	WT_root_3w_ nodC_inocul1	WT Root 3w nodC inocul1	3	Root	Treatment	3-week-old uninoculated roots—3 days post-R7A inoculation	*Lotus japonicus Gifu*	Sterilized Leca	Høgslund et al. (2009)
Nodulation	WT_root_6w_5mM_nitrate	WT root 6w 5mM nitrate	3	Root	Treatment	6-week-old plants—nitrate treatment	*Lotus japonicus Gifu*	Sterilized Leca	Høgslund et al. (2009)
Nodulation	WT_root_nodule_ 3w_inocul21	WT root + nodule 3w inocul21	3	Root + nodule	Treatment	3-week-old uninoculated roots—21 days post-R7A inoculation	*Lotus japonicus Gifu*	Sterilized Leca	Høgslund et al. (2009)
Nodulation	WT_root_nodule_ 3w_inocul7	WT root + nodule 3w inocul7	3	Root + nodule	Treatment	3-week-old uninoculated roots—7 days post-R7A inoculation	*Lotus japonicus Gifu*	Sterilized Leca	Høgslund et al. (2009)
Nodulation	WT_rootSZ_3w_inocul1	WT rootSZ 3w inocul1	3	Root	Treatment	3-week-old uninoculated roots—1 day post-R7A inoculation	*Lotus japonicus Gifu*	Sterilized Leca	Høgslund et al. (2009)
Nodulation	WT_rootSZ_ 3w_Nod_inocul1	WT rootSZ 3w Nod inocul1	3	Root susceptible zone	Treatment	3-week-old uninoculated roots—1 day post-R7A inoculation	*Lotus japonicus Gifu*	Sterilized Leca	Høgslund et al. (2009)
Nodulation	WT_shoot_3w_ 5mM_nitrate	WT shoot 3w 5mM nitrate	3	Shoot	Treatment	3-week-old plants—nitrate treatment	*Lotus japonicus Gifu*	Sterilized Leca	Høgslund et al. (2009)
Nodulation	WT_shoot_3w_inocul3	WT shoot 3w inocul3	3	Shoot	Treatment	3-week-old uninoculated roots—3 days post-R7A inoculation	*Lotus japonicus Gifu*	Sterilized Leca	Høgslund et al. (2009)
Nodulation	WT_stem_6w_ 5mM_nitrate	WT stem 6w 5mM Nitrate	3	Stem	Treatment	6-week-old plants—nitrate treatment	*Lotus japonicus Gifu*	Sterilized Leca	Høgslund et al. (2009)
Mycorrhization	root_4dpi control	root 4dpi control	4	Root	Treatment	No inoculation	*Lotus japonicus Gifu*	Sand	Guether et al. (2009)
Mycorrhization	root_4dpi mycorrhized	root 4dpi mycorrhized	4	Root	Treatment	Mycorrhization by inoculation with Gigaspora margarita	*Lotus japonicus Gifu*	Sand	Guether et al. (2009)
Mycorrhization	root_28dpi control	root 28dpi control	3	Root	Treatment	No inoculation	*Lotus japonicus Gifu*	Sand	Guether et al. (2009)
Mycorrhization	root_28dpi mycorrhized	root 28dpi mycorrhized	4	Root	Treatment	Mycorrhization by inoculation with Gigaspora margarita	*Lotus japonicus Gifu*	Sand	Guether et al. (2009)

Abbreviations Dap Day after pollination, *dpi* day post-inoculation, *SZ* susceptible zone, *WT* wild-type, *Ctrol* control, *w* weeks

raw data collected from different experiments, the quantile method with Robust Multichip Average (RMA) was used (Irizarry 2003). The raw data (.CEL files) were also imported into dCHIP software (Li and Wong 2001) to assign presence/absence calls for each probe set using the software's default settings.

17.4.1 Search Criteria

To make LjGEA user-friendly, many different options are given so that even partial information about a given gene is enough to serve as a search criteria and the corresponding transcript can be retrieved (Fig. 17.1). The three main types of queries that can be used to find expression data for a specific gene are as follows:

A. The probe set ID can be used. It is the quickest and the most accurate way of retrieving expression data, given the probe set is specific to the gene. The limitation, on the other hand is that the user has to know the exact probe set ID, which is often not possible with just a sequence in hand.
B. The gene ID or gene annotation: If known, a gene ID (according to the *L. japonicus* Genome Database or LjGDB) can be used as a query. Alternatively, a gene description (e.g., ABC transporter), gene annotation or gene ontology (GO) (e.g., GO: 0004757 for sepiapterin reductase), and KEGG gene bincodes (e.g., 1.1.1.7, which stands for plastidic aldolase) can also be used. The query must be given in specific formats, which are shown as examples on the server.
C. BLAST search: This can be used when only the sequence of the gene or a part of the sequence of the transcript or cDNA sequence is known. This can also be useful for an unknown gene or a gene of which the ortholog from other species (e.g., Medicago)

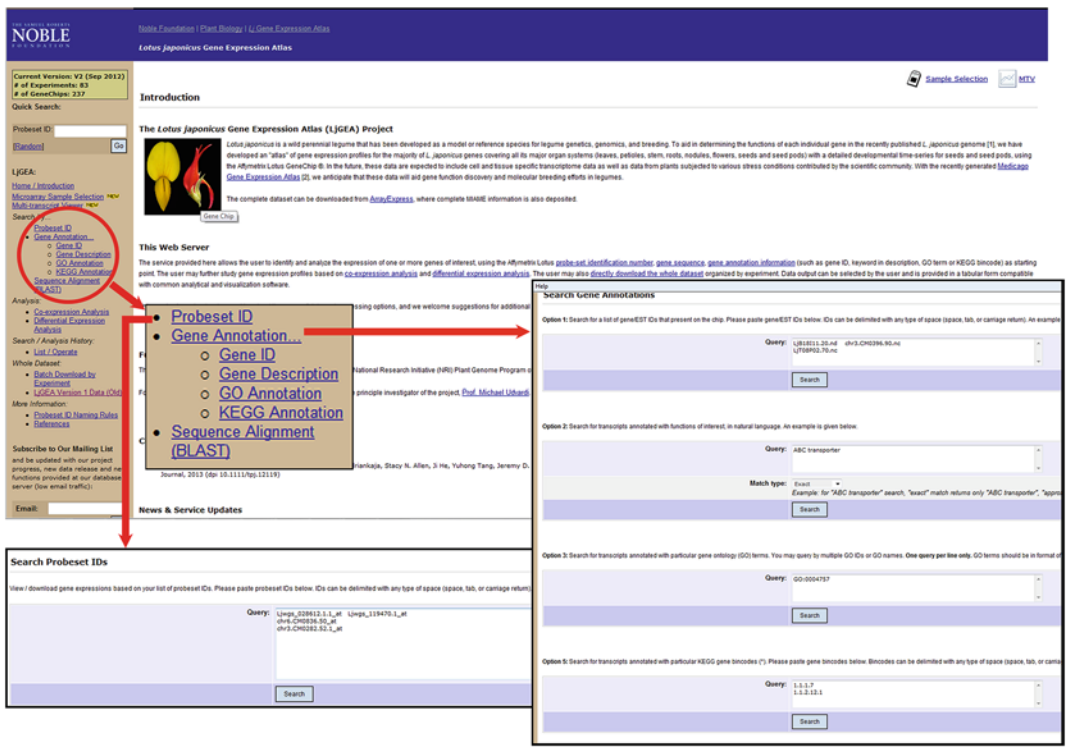


Fig. 17.1 The Lotus GEA Web server: On the *left panel*, clickable links open search pages. The search criteria can be probe set IDs, gene IDs, gene description, GO terms, or KEGG terms. All the search pages contain examples of the correct format for inputting queries

is known. Partial sequence of genomic DNA is not recommended since the program may not find a sequence from an intron. There are two options for target database in the drop down menu, the Affymetrix probe target sequence (the whole transcript) and the Affymetrix probe consensus sequence (the parts of the transcript to which the probe sets bind). Users of the BLAST search option should always check the alignment (which comes as an additional output, Fig. 17.2) to be sure that the given transcript really corresponds to the sequence of interest.

17.4.2 The Result Output Page

There are several layers of result output (Fig. 17.3). The first layer gives the corresponding probe set ID and the values of transcript abundance. There is a clickable button, which will provide the graphic representation of the transcript level in all the tissues tested. Users can download batch results in the case of multiple queries or more generalized queries such as GO or gene description. There is also an 'add to transcript viewer' feature that helps to view multiple transcripts on the same graph so that users can compare between

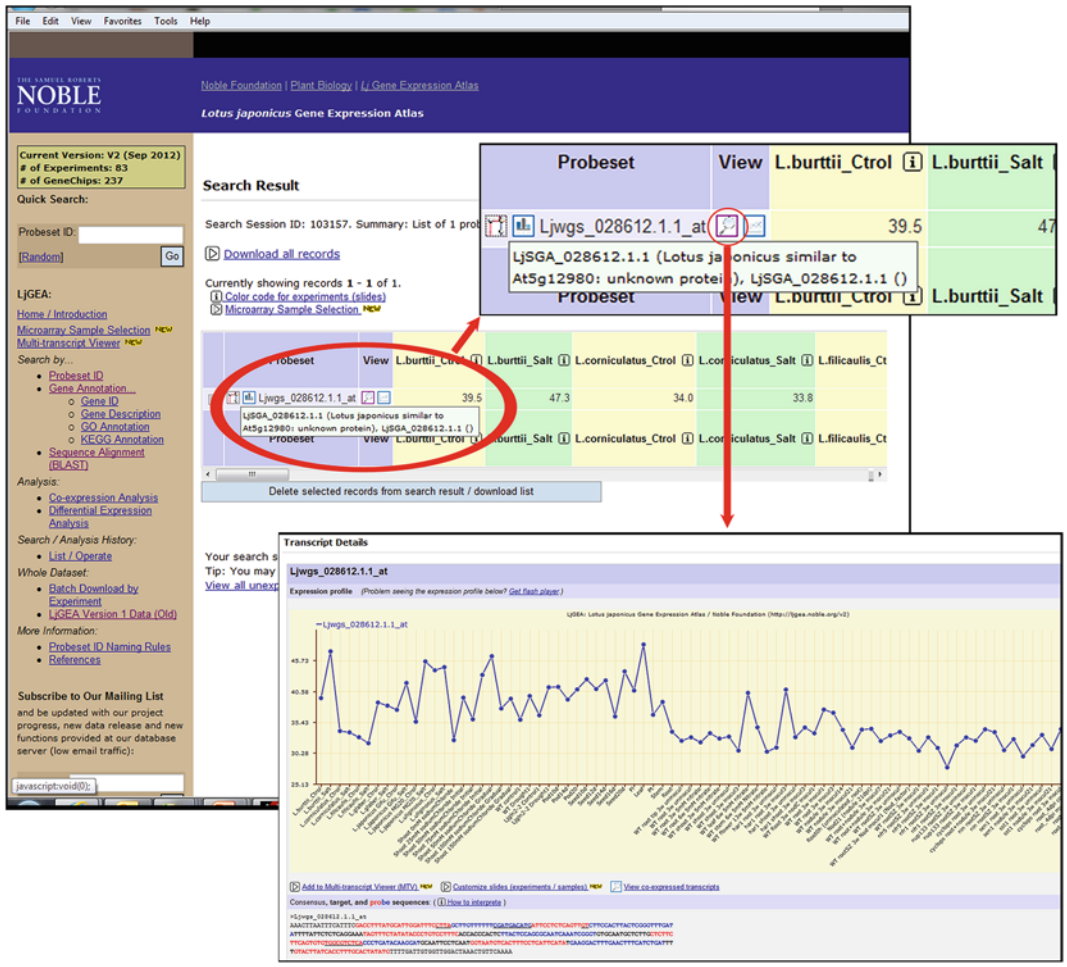


Fig. 17.2 Using the BLAST option: A nucleotide sequence can be used as a search query in the BLAST option. The result page comes with an additional link, which will open the alignment (*inset* on the *lower right*). The bases aligned are marked in green. The individual probes are indicated by alternative *red* and *blue* colors, while the overlapping portions between two consecutive probes are *underlined*

different genes. Alignments are also given in the case of BLAST searches.

17.4.3 Final Analyses

The LjGEA Web server not only provides a platform for gene expression data retrieval, but also serves as a tool for further analysis. The server indicates whether the gene is legume-specific, whether it falls in a known common pathway, or whether it has a homolog in other related legumes (i.e., Medicago and soybean). Two other very useful tools are co- and differential expression analysis. The server enables batch retrieval of genes, whose expression profiles are highly correlated with that of a chosen gene. This can be performed for a subset of experiments or the entire data set from all experiments. For each co-expression analysis session, users customize the co-expression calculation method (currently Pearson's correlation coefficient or cosine correlation; Rodgers and Nicewander 1988) and set a correlation threshold and the maximum number of transcripts to be returned. As described earlier, the multi-transcript viewer enables visual comparison of expression profiles. On the other hand, the differential expression tool allows the user to retrieve all the genes whose expression is changed between two user-chosen conditions. The user can also determine the threshold of change (fold-change). Altogether, this server can be very helpful for both forward and reverse genetic studies.

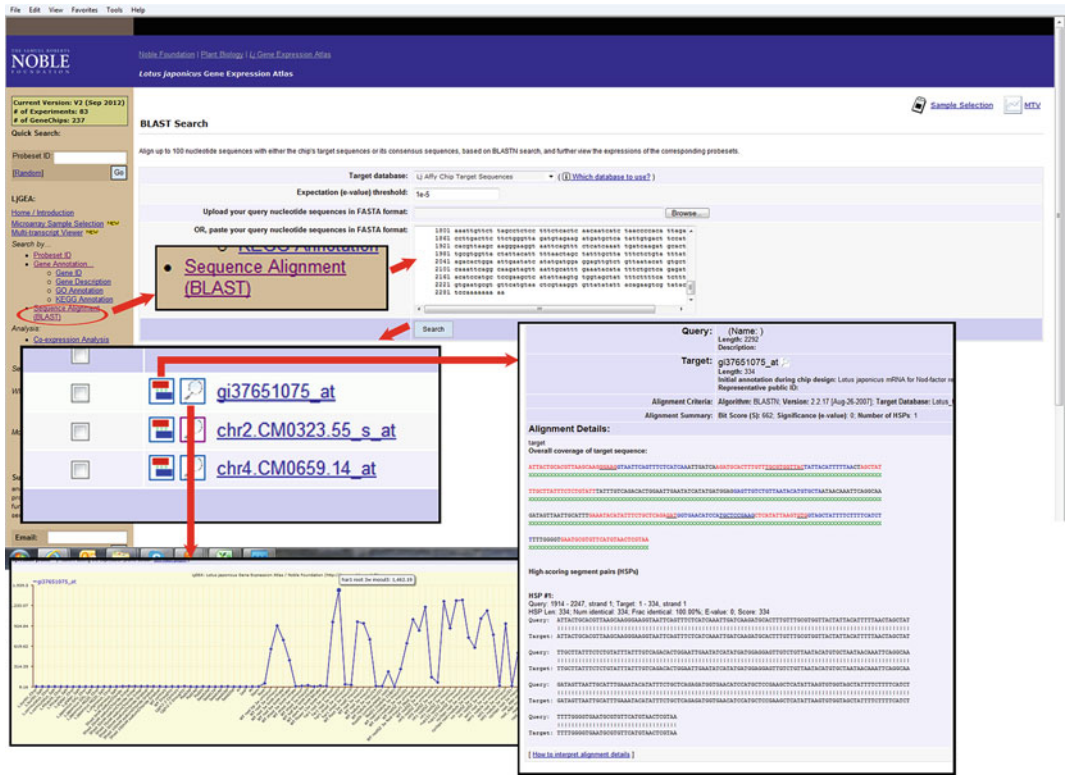


Fig. 17.3 The result page: the first page of result contains the probe set number and the values of transcript level in each condition. The clickable 'magnifying glass' symbol opens another window, where the transcript levels under different conditions are represented graphically

17.5 Gene Regulatory Networks: The LegumeGRN Web Server

Gene expression datasets provide snapshots of the transcriptomes of plant organs with/without treatments and under various experimental conditions. Genes and gene products interact with each other in complex structured regulatory networks. Transcriptomes are valuable data to uncover these complex regulatory interactions. To predict gene interactions, several algorithms have been developed using statistical and computational tools. A Web-based computational service was developed to build, test, analyze, and visualize gene regulatory networks (GRNs) (Wang et al. 2013). This Web server, called LegumeGRN (http://legumegrn.noble.org), is preloaded with Affymetrix GeneChip-based transcriptomic data from Medicago, soybean, and Lotus. The LegumeGRN Web server hosts the 83 Lotus transcriptomic experiments with the same 237 Affymetrix chips stored in the LjGEA and described in 21.4.

17.5.1 Lotus GRN Homepage

When users log in, they open the Lotus GRN homepage by selecting Lotus in the "Submit Gene Network prediction" tab. Then, users are invited to (i) provide a list of probe sets/genes that will be used to build the GRN, (ii) select general options, (iii) select all or a subset of preloaded transcriptomic data, and (iv) select one or several GRN prediction algorithms (Fig. 17.4a).

The primary input of the LegumeGRN is a list of Lotus probe sets/genes, which will serve to build the gene regulatory network. This list of genes has to be selected according to the purpose of the study or according to the selected predictive algorithm. As an example, according to the purpose of the study, users may select a limited number of genes such as organ-specific genes to build a network related to a seed-specific mechanism by checking the specific experiments to use for the GRN prediction. According to the predictive algorithm, users may also select a large set of genes when using "relevance network" algorithm (i.e., co-expression network) but are restricted to a limited number of genes associated with a large number of transcriptomic experiment when using graphical Gaussian model for reasons of specificity and/or RAM memory requirements specific to each algorithm. Users will have access to six different robust algorithms to predict GRNs (Marbach et al. 2012): relevance network based on Pearson's or Spearman correlation (Stuart et al. 2003), graphical Guassian model (GGM, Schäfer and Strimmer 2005), GENIE3 (Huynh-Thu et al. 2010), TIGRESS (Haury et al. 2012), CLR (Faith et al. 2007) and parallel low-order PC algorithm (Wang et al. 2010). Description of algorithms is provided with their specificity by clicking on the 'question mark' icons, and default settings are proposed to users (Fig. 17.4a). Finally, users may change general options or keep default settings concerning using transcription factors as main connectors and the number of connections. By default, immediate connections will be calculated for transcription factors only (i.e., when "yes" is checked for "using transcription factors" option) and the number of edges is unlimited (i.e., −1 is selected as a cutoff). The drop down menu "submit gene network prediction" allows users to enter their own in-house transcriptomic data that are not contained in the Web server.

17.5.2 Result Panel

After calculation, GRN prediction results are saved into LegumeGRN Web server users' accounts, which allow users to store and keep track of their analyses and results. Users have the choice between downloading network results or visualizing and analyzing them using an intuitive Web-based GRN viewer (Fig. 17.4b). The visualization module consists of two parts: a graphical output on the left panel and gene annotations on the right panel (Fig. 17.4b). The graphical

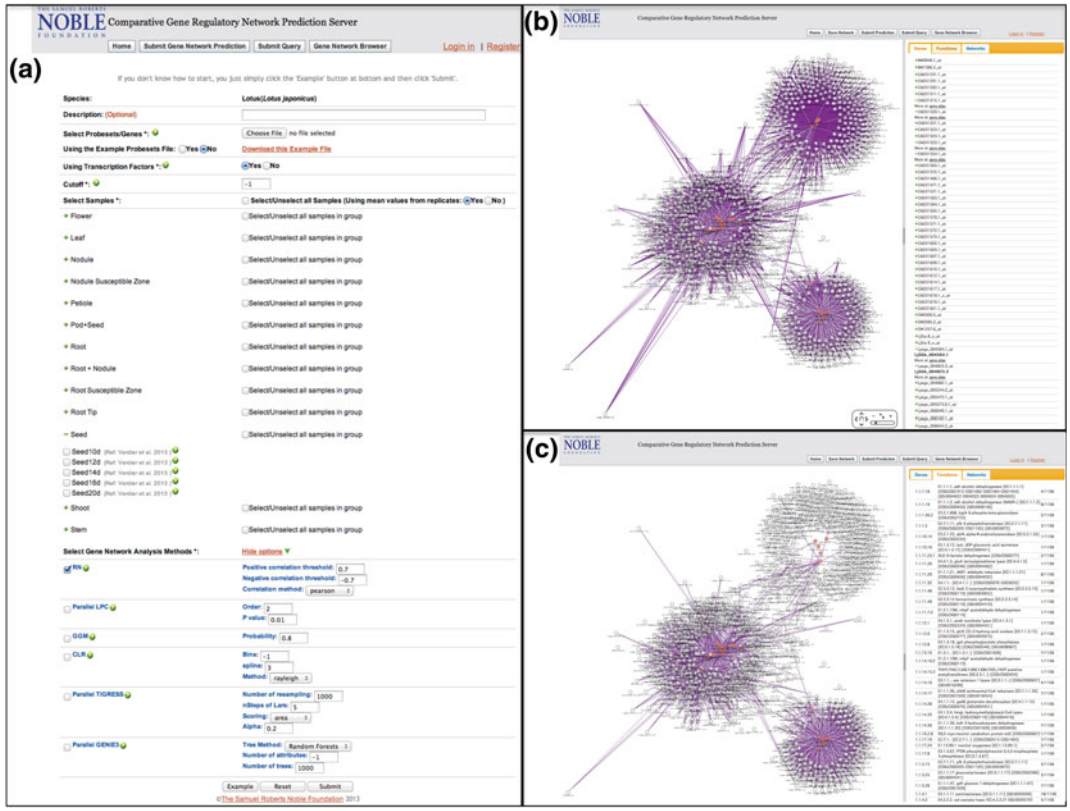


Fig. 17.4 The LegumeGRN Web server. **a** Homepage of the Web server with tabs to select species, tools, and gene network browser on the *top*. This homepage displays the four parts that will be used to build GRN: list of probe sets/gene input, general options, selection of transcriptomic data and predictive algorithms. **b** Example of a GRN visualized using the Web-based GRN viewer. This visualization module consists of a graphical output on the *left panel*, where TFs (*orange nodes*) are connected to putative target genes (*gray nodes*) and an annotation panel on the *right side*. **c** Visualization of connections from a GRN built using two different algorithms: connections from relevance network are identified by *purple lines* and from the GGM algorithm with *gray lines*. This network highlights similarities between connections predicted from the two different predictive algorithms

display has several features such as zoom in/zoom out and move options. Each node with its connection strength and prediction approaches is shown by line width and color, respectively. All circles represent genes (TFs are in orange and putative target genes in open circles) and are clickable to link them to their probe set ID, gene ID, tentative annotation and GO term on the right panel (Fig. 17.4b). Clicking on the edge displays the connection strength value calculated by the corresponding GRN prediction algorithm and a link to show the gene expression profiles for the related gene pair in the LjGEA Web server.

17.5.3 Analysis

Features of GRN analysis include comparison and integration of multiple networks predicted by different algorithms. Users can choose to overlay multiple GRNs and construct a composite network in order to highlight similarities and differences between predictive algorithms (Fig. 17.4c). Another feature is "sub-network query." Usually, GRNs are too large and too complex to be displayed or analyzed. LegumeGRN allows users to extract sub-networks by showing only immediate connections of a specified gene list or by selecting

the most relevant connections according to the confidence ranking generated by the algorithm. Finally, a module to identify significant enrichment in GO terms of each (sub-) network was added to identify molecular functions in which most of genes are involved.

17.5.4 Example of Multiple Algorithm Calculation

As an example, we uploaded the list of 1,190 tissue-specific genes identified in Verdier et al. (2013), we selected the transcriptomic data of the major plant organs (i.e., leaf, root, nodule, seed, pod, petiole, and flower) and chose to predict networks using relevant network (i.e., RN) and GGM algorithms with default settings (Fig. 17.4a). After submitting this first task, we merged both network predictions and extracted the immediate connections of all TFs, options available in the "Submit query" menu. Figure 17.4b represents the predicted network with nodes (i.e., genes) connected by purple lines determined by RN algorithm and gray lines determined by GGM. In Fig. 17.4c, we chose to visualize similarities between these two networks by highlighting connections identified by both methods, which presumably represents a more robust prediction.

17.6 Outlook

In this chapter, we described new tools that are available to store, visualize, and analyze *L. japonicus* transcriptomic data. Most of these transcriptomic data were obtained from the Affymetrix GeneChip. However, the development of the Agilent NimbleGen custom gene expression array for *L. japonicus* (named 4 × 44_Kazusa-001 or A-GEOD-14826) has recently increased the amount of transcriptomic data. This array contains 44,000 probe sets of 60-mer oligonucleotides, which correspond to 21,495 genes. This array has already been used in various experiments, such as cell-type-specific transcriptome analysis in nodules (E-GEOD-34753; Takanashi et al. 2012), constitutive over-expression of *LjMyb14* (E-GEOD-31739; Shelton et al. 2012), and glutathione-induced elicitation of defense responses (E-GEOD-31240) and should provide more data in the next years. In parallel to microarrays, RNA-seq technology is rapidly evolving and will, no doubt, produce much transcriptome data for Lotus in the near future. All novel transcriptomic data generated for *L. japonicus* using different array or RNA-seq technologies can, in principal, be integrated into the LjGEA, using recently developed normalization strategies such as the one described in Battke and Nieselt (2011) to provide a more comprehensive understanding of Lotus biology.

Lotus is closely related to major grain legumes such as lentil, fava, pea, chickpea, common bean, mung bean, soybean, and pigeon pea. By identification of putative orthologs of Lotus genes in the two other model legumes, Medicago and Soybean, the LjGEA provides a useful tool for translational genomics, which can be useful for basic and applied research in many crop legumes. Moreover, the LegumeGRN Web server contains a feature for "across-species comparison," which takes into account ortholog genes from Medicago and Soybean. Indeed, regulatory networks governing basic cellular functions are conserved in diverse species (Stuart et al. 2003). The composite network generated from different species will display color-coded edges representing the network connections for each different species. This multi-species GRN comparison will enable detection of evolutionary conserved gene regulatory (sub-) networks and help to place GRNs in a phylogenetic context. This feature might be extended in the future to other legumes as new genomic data become available.

References

Aceituno F, Moseyko N, Rhee S, Gutierrez R (2008) The rules of gene expression in plants: organ identity and gene body methylation are key factors for regulation of gene expression in *Arabidopsis thaliana*. BMC Genom 9(1):438

Asamizu E, Nakamura Y, Sato S, Tabata S (2005) Comparison of the transcript profiles from the root and the nodulating root of the model legume *Lotus japonicus* by serial analysis of gene expression. Mol Plant Microbe Interact 18(5):487–498

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G (2000) Gene ontology: tool for the unification of biology. Nat Genet 25(1):25–29

Battke F, Nieselt K (2011) Mayday SeaSight: combined analysis of deep sequencing and microarray data. PLoS One 6:e16345. PMID: 21305015. doi:10.1371/journal.pone.0016345

Betti M, Pérez-Delgado C, García-Calderón M, Díaz P, Monza J, Márquez A (2012) Cellular stress following water deprivation in the model legume *Lotus japonicus*. Cells 1(4):1089–1106

Colebatch G, Desbrosses G, Ott T, Krusell L, Montanari O, Kloska S, Kopka J, Udvardi MK (2004) Global changes in transcription orchestrate metabolic differentiation during symbiotic nitrogen fixation in *Lotus japonicus*. Plant J 39(4):487–512

Colebatch G, Kloska S, Trevaskis B, Freund S, Altmann T, Udvardi MK (2002) Novel aspects of symbiotic nitrogen fixation uncovered by transcript profiling with cDNA arrays. Mol Plant Microbe Interact 15 (5):411–420

Díaz P, Betti M, Sánchez DH, Udvardi MK, Monza J, Márquez AJ (2010) Deficiency in plastidic glutamine synthetase alters proline metabolism and transcriptomic response in *Lotus japonicus* under drought stress. New Phytol 188:1001–1013

Endo M, Hakozaki H, Kokubun T, Masuko H, Takahata Y, Tsuchiya T, Higashitani A, Tabata S, Watanabe M (2002) Generation of 919 expressed sequence tags from immature flower buds and gene expression analysis using expressed sequence tags in the model plant *Lotus japonicus*. Genes Genetic Syst 77 (4):277–282

Endo M, Kokubun T, Takahata Y, Higashitani A, Tabata S, Watanabe M (2000) Analysis of expressed sequence tags of flower buds in *Lotus japonicus*. DNA Res 7(3):213–216

Faith JJ, Hayete B, Thaden JT, Mogno I, Wierzbowski J, Cottarel G, Kasif S, Collins JJ, Gardner TS (2007) Large-scale mapping and validation of *Escherichia coli*. Transcriptional regulation from a compendium of expression profiles. PLoS Biol 5(1):e8

Guether M, Balestrini R, Hannah M, He J, Udvardi MK, Bonfante P (2009) Genome-wide reprogramming of regulatory networks, transport, cell wall and membrane biogenesis during arbuscular mycorrhizal symbiosis in *Lotus japonicus*. New Phytol 182 (1):200–212

Haury A-C, Mordelet F, Vera-Licona P, Vert J-P (2012) TIGRESS: trustful inference of gene regulation using stability selection. BMC Syst Biol 6(1):145

He J, Benedito VA, Wang M, Murray JD, Zhao PX, Tang Y, Udvardi MK (2009) The *Medicago truncatula* gene expression atlas web server. BMC Bioinform 10:441

Høgslund N, Radutoiu S, Krusell L, Voroshilova V, Hannah Ma, Goffard N, Sanchez DH, Lippold F, Ott T, Sato S, Tabata S, Liboriussen P, Lohmann GV, Schauser L, Weiller GF, Udvardi MK, Stougaard J (2009) Dissection of symbiosis and organ development by integrated transcriptome analysis of *Lotus japonicus* mutant and wild-type plants. PloS One 4: e6556

Huynh-Thu VA, Irrthum A, Wehenkel L, Geurts P (2010) Inferring regulatory networks from expression data using tree-based methods. PLoS ONE 5(9):e12776

Irizarry RA (2003) Summaries of Affymetrix GeneChip probe level data. Nucleic Acids Res 31:e15

Kanehisa M, Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 28(1):27–30

Kouchi H, Shimomura K, Hata S, Hirota A, Wu G-J, Kumagai H, Tajima S, Suganuma N, Suzuki A, Aoki T, Hayashi M, Yokoyama T, Ohyama T, Asamizu E, Kuwata C, Shibata D, Tabata S (2004) Large-scale analysis of gene expression profiles during early stages of root nodule formation in a model legume, *Lotus japonicus*. DNA Res 11(4):263–274

Li C, Wong WH (2001) Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. Proc Natl Acad Sci USA 98:31–36

Marbach D, Costello JC, Kuffner R, Vega NM, Prill RJ, Camacho DM, Allison KR, Kellis M, Collins JJ, Stolovitzky G (2012) Wisdom of crowds for robust gene network inference. Nat Meth 9(8):796–804

Moreno-Hagelsieb G, Latimer K (2008) Choosing BLAST options for better detection of orthologs as reciprocal best hits. Bioinformatics 24:319–324

Parkinson H, Sarkans U, Shojatalab M, Abeygunawardena N, Contrino S, Coulson R, Farne A, Garcia Lara G, Holloway E, Kapushesky M, Lilja P, Mukherjee G, Oezcimen A, Rayner T, Rocca-Serra P, Sharma A, Sansone S, Brazma A (2005) ArrayExpress—a public repository for microarray gene expression data at the EBI. Nucleic Acids Res 33(suppl 1):D553–D555

Pérez-Delgado CM, García-Calderón M, Sanchez DH, Udvardi M, Kopka J, Marquez AJ, Betti M (2013) Transcriptomic and metabolic changes associated to photorespiratory ammonium accumulation in the model legume *Lotus japonicus*. Plant Physiol. doi:10.1104/pp.113.217216

Rodgers J, Nicewander WA (1988) Thirteen ways to look at the correlation coefficient. Am Stat 42(1):59–66

Sanchez DH, Lippold F, Redestig H, Hannah Ma, Erban A, Krämer U, Kopka J, Udvardi MK (2008) Integrative functional genomics of salt acclimatization in the model legume *Lotus japonicus*. Plant J 53:973–987

Sanchez DH, Pieckenstain FL, Szymanski J, Erban A, Bromke M, Hannah Ma, Kraemer U, Kopka J, Udvardi MK (2011) Comparative functional genomics of salt stress in related model and cultivated plants identifies and overcomes limitations to translational genomics. PloS One 6:e17094

Sato S, Nakamura Y, Kaneko T, Asamizu E, Kato T, Nakao M, Sasamoto S, Watanabe A, Ono A, Kawashima K, Fujishiro T, Katoh M, Kohara M, Kishida Y, Minami C, Nakayama S, Nakazaki N, Shimizu Y, Shinpo S, Takahashi C, Wada T, Yamada M, Ohmido N, Hayashi M, Fukui K, Baba T, Nakamichi T, Mori

H, Tabata S (2008) Genome structure of the legume, *Lotus japonicus*. DNA Res 15:227–239

Schäfer J, Strimmer K (2005) An empirical Bayes approach to inferring large-scale gene association networks. Bioinformatics 21(6):754–764

Shelton D, Stranne M, Mikklesen L, Pakseresht N, Welham T, Hiraka H, Tabata S, Sato S, Paquette S, Wang T, Martin C, Bailey P (2012) Transcription factors of *Lotus japonicus*: regulation of isoflavonoid biosynthesis requires coordinated changes in transcription factor activity. Plant Physiol. doi:10.1104/pp.112.194753

Stuart JM, Segal E, Koller D, Kim SK (2003) A gene-coexpression network for global discovery of conserved genetic modules. Science 302:249–255

Takanashi K, Takahashi H, Sakurai N, Sugiyama A, Suzuki H, Shibata D, Nakazono M, Yazaki K (2012) Tissue-specific transcriptome analysis in nodules of *Lotus japonicus*. Mol Plant Microbe Interact 25 (7):869–876

Verdier J, Torres-Jerez I, Wang M, Andriankaja A, Allen SN, He J, Tang Y, Murray JD, Udvardi MK (2013) Establishment of the *Lotus japonicus* gene expression atlas (LjGEA) and its use to explore legume seed maturation. Plant J 74(2):351–362

Wang M, Augusto Benedito V, Xuechun Zhao P, Udvardi M (2010) Inferring large-scale gene regulatory networks using a low-order constraint-based algorithm. Mol BioSyst 6:988–998

Wang M, Verdier J, Benedito V, Tang Y, Murray J, Ge Y, Becker J, Carvalho H, Rogers C, Udvardi M, He J (2013) LegumeGRN: a web application to predict gene regulatory network for functional and comparative studies. PLOS One 8:e67434

Proteomics

18

Svend Dam and Jens Stougaard

Abstract

Proteomics is an efficient tool to identify proteins present in specific tissues, cell types, or organelles. The resulting proteome reference maps and/or comparative analyses provide overviews of regulated proteins between wild type and mutants or between different conditions together with a comprehensive list of proteins. Post translation modifications (PTMs), such as glycosylation and phosphorylation, are pivotal for protein stability and function. Several strategies for enrichment of PTMs have been developed where targeted proteomic approaches are used to identify these PTMs. The sequenced and annotated *Lotus japonicus* (*Lotus*) genome has been essential for obtaining high-quality protein identifications from proteomics studies. Furthermore, additional genomics and transcriptomics studies from several *Lotus* species/ecotypes support putative gene structures and these can be further supported using proteomics data. Two characteristics of legumes are the high seed protein level and the nitrogen fixing symbiosis. Thus, the majority of the proteomics studies in *Lotus* have been performed on seed/pod and nodule/root tissues in order to create proteome reference maps and to enable comparative analyses within *Lotus* tissues or toward similar tissues from other legume species. More recently, *N*-glycan structures and compositions have been determined from mature *Lotus* seeds using glycomics and glycoproteomics, and finally, phosphoproteomics has been employed.

18.1 Extraction and Separation of Proteins/Peptides

The proteome defines the proteins expressed in an organism, tissue, organ, cell, or subcellular component (Wasinger et al. 1995). Proteomics is a large scale study of the proteome and four major steps, i.e., protein extraction, protein/peptide separation, mass spectrometry, and data

S. Dam (✉) · J. Stougaard
Centre for Carbohydrate Recognition and Signalling, Department of Molecular Biology and Genetics, Aarhus University, Gustav Wieds Vej 10, 8000 Aarhus C, Denmark
e-mail: svd@mb.au.dk

S. Tabata and J. Stougaard (eds.), *The Lotus japonicus Genome*, Compendium of Plant Genomes,
DOI 10.1007/978-3-662-44270-8_18,

searches, are required prior to data analysis. There is no universal proteomics setup and each of the four steps is crucial in order to obtain a useful dataset. For plants, protein extraction is challenging due to interfering compounds such as polyphenols, terpenes, and organic acids that are abundant in green tissues and, thus, several protein extraction protocols have been developed (Wang et al. 2003; Xie et al. 2007; Jellouli et al. 2010). For *Lotus* tissues, we have successfully used a SDS/phenol protein extraction followed by ammonium acetate precipitation and further removal of interfering compounds by acetone washes (Dam et al. 2009; Nautrup-Pedersen et al. 2010; Dam et al. 2014).

After extraction, the protein fraction is redissolved and proteins are either separated/digested or digested/separated. For the separation/digestion methodology, proteins are solubilized and separated using 1D or 2D gels followed by excision of bands/spots, digestion, and desalting prior to MS. Alternatively, the peptides can be further separated using liquid chromatography (LC) prior to MS which is frequently used for 1D separation. The 2D gel technology is a two-step separation of up to 5,000 protein spots (O'Farrell 1975) and these can be visualized/quantified down to 1 ng (Weiss and Görg 2007; Görg et al. 2004). For the digestion/peptide separation methodology, the protein fraction is solubilized, in-solution digested, and peptides are separated using LC. Prior to LC, peptides can be labeling with, for example, iTRAQ used for relative quantification. In *Lotus*, the protein separation/digestion strategy was used to create proteome reference maps of seed, pod, nodule, and root using 2D gels (Nautrup-Pedersen et al. 2010; Dam et al. 2014) together with subcellular protein extraction of seed globulin-, plant cytosolic-, and symbiosome membrane proteins (Dam et al. 2009, 2013, 2014; Nautrup-Pedersen et al. 2010; Credali et al. 2013; Wienkoop and Saalbach 2003) and, furthermore, the digestion/peptide methodology has been initiated for nodules and roots. Currently, the majority of proteomics studies are performed with the digestion/peptide methodology. However, the 2D gel methology has the advantage of visualizing PTMs given that each protein isoform is sufficiently separated in the 2D gel and, thus, can be analyzed individually to identify the variety of PTMs for the protein (Rogowska-Wrzesinska et al. 2013).

18.2 Mass Spectrometry Techniques and Data Search

In the last decade, the MS instruments used for MS scan and MS/MS analysis have developed rapidly both in sensitivity and in speed. Mass spectrometers have three major functional parts, i.e., ion source, mass analyzer, and detector. These are developed as independently functional entities and can be combined differently. For a more comprehensive description of mass spectrometers, see Parker et al. (2010).

For *Lotus* proteomics, the matrix-assisted laser desorption/ionization (MALDI) and electrospray ionization (ESI) were the two ion sources used. For MALDI, the tryptic desalted peptides were eluted on a target plate together with a strong absorption matrix. The mass analyzer, the time of flight (TOF), detected the mass/charge (m/z) of peptide ions and TOF/TOF was used for fragmentation of selected peptide ions. For the ESI ion source, the desalted peptides were separated using a C-18 column to reduce the complexity of peptide ions analyzed. The peptide ions were sprayed into the mass spectrometer and the mass analyzer Q-TOF, ion trap, and recently orbitrap were used for *Lotus* proteomics.

The Mascot software was used for *Lotus* protein identifications. The proteomics data was searched against all predicted *Lotus* protein coding genes, and for validation, the data was analyzed manually using the novel available MS data miner (MDM) software (Dyrlund et al. 2012).

All four major steps in a proteome analysis are crucial and have to be optimized for each experimental setup. In the following, proteomics of different *Lotus* tissues together with some of the biological interpretations are discussed.

18.3 Proteomics of Lotus Seeds and Pods

Legumes have a high level of nutrients in mature seeds, however, the molecular mechanisms behind the accumulation of nutrients have not been revealed. Model legumes with sequenced genomes and the availability of mutant lines are useful to obtain a more detailed knowledge of the molecular mechanisms. In *Lotus*, seeds from five developmental stages and pods from three developmental stages were analyzed using a transcriptomics approach focusing on transcription factors controlling/regulating the seed development (Verdier et al. 2013). Furthermore, all *Lotus* Affymetrix transcriptomics data sets from different mutants and tissues are combined and searchable at the http://ljgea.noble.org/ homepage (Verdier et al. 2013). Thus, together with the sequenced *Lotus* genome, several transcriptomics, and microRNA studies (Verdier et al. 2013; Sato et al. 2008; Høgslund et al. 2009; De Luis et al. 2012), proteomics may provide new knowledge about protein networks essential for the high nutrient value in *Lotus* seeds.

Prior to proteomics of mature *Lotus* seed, protein, starch, lipids, phytic acid, and ash levels were determined (Dam et al. 2009). The mature *Lotus* seeds contain approximately 43 % protein and less than 1 % starch even though starch granulates are visible during the seed filling phase (Dam et al. 2009). This protein and starch pattern is also seen for soybean and *Medicago*, whereas mature pea seeds have a lower protein level at ~25 % but ~50 % starch (Wilson et al. 1978; Prakash and Misra 1988; Djemel et al. 2005). The storage globulin fraction, which is insoluble in H_2O but soluble in 1 M of salt, can be up to 80 % of the protein fraction in mature legume seeds. The globulins were enriched and separated using 2D gels. Spots corresponding to legumin and convicilin, which are two of the major globulins, were identified with an approximate volume ratio 9:1 in mature wild-type seeds (Dam et al. 2009). In *Lotus*, 80 % of the nitrogen transport from root to shoot via xylem is linked to asparagine (Waterhouse et al. 1996). In line with this observation, transcripts of the K^+-dependent NSE1 asparaginase, that catalyzes the hydrolysis of asparagine to aspartic acid, was abundant in *Lotus* sink tissues (Credali et al. 2011). Further, functional analysis supports a role for the NSE1 asparaginase during seed development and/or seed filling. Seeds from the NSE1 mutant have lower seed weight than wild type together with a higher number of abnormal seeds in the pods (Credali et al. 2013). For the most severe mutant, *nse1-3*, the globulin fraction is 20 times less abundant compared with wild-type seeds, whereas the legumin/convicilin ratio was not changed (Credali et al. 2013). In conclusion, proteomics of the *Lotus* seed globulin fraction is useful for composition and abundance analysis of storage globulins from mutants affected in the mature seed protein composition.

For a more elaborate proteomics study of different developmental stages of *Lotus* seed development, the switches between embryogenesis, seed filling, and desiccation phases were determined by calculating the water content together with histological sections of seeds from defined developmental stages (between 7 and 43 days after flowering with a three days interval) (Dam et al. 2009). Proteomics was initially performed from two developmental stages corresponding to the seed filling [between 19 and 25 days after flowering (green seeds)] and desiccation [more than 43 days after flowering (mature seeds)] phases using GeLC-MS/MS with 920 and 264 proteins identified from the two developmental stages, respectively. The lower number of proteins identified from mature seeds represent probably the high abundant legumins and convicilins more than that a lower number of proteins present in mature seeds (Dam et al. 2009). Additionally, a quantitative and a more systematic proteomics study of *Lotus* seed development, using 2D gels within two pH ranges; pH 4-7 and pH 6-11, was performed. Five developmental stages, i.e., two stages of embryogenesis and seed filling phases together with one desiccation phase, were analyzed to determine the differences in the level of specific

proteins during seed development. For a comparative analysis, pods without seeds for the five corresponding stages were included (Nautrup-Pedersen et al. 2010). In total, 604 and 965 protein spots were identified for pods and seeds where the pod proteins correspond to 567 different gene accessions and 263 of those were not identified in seeds, indicating differences in the level and composition of proteins between the two tissues. The identification of different putative enzymes in the urea cycle pathway between pods and seeds suggests that the ammonium of degraded urea can be assimilated into amino acids in the seed and, thus, a possible factor for the high protein level in mature legume seeds (Nautrup-Pedersen et al. 2010). All obtained proteomics data from *Lotus* seed and pod are stored and available at http://www.cbs.dtu.dk/cgi-bin/lotus/db.cgi (Dam et al. 2009; Nautrup-Pedersen et al. 2010). For the 2D gel analysis, a master gel for pods and seeds is uploaded and by clicking on the master gel, protein accessions and spot numbers are visible. Furthermore, GO annotations, identified peptides and quantitative data are available for each of the five developmental stages analyzed.

The overall seed development for the two model legumes *Lotus* and *Medicago* together with soybean is similar with a transiently accumulation of starch during seed development, whereas in the mature seed the starch level is lower than 1 % together with high protein level. Thus, all proteomics data from these three species were merged and available at http://bioinfoserver.rsbs.anu.edu.au/utils/PathExpress/pathexpress4legumes.php . This is useful for a more broad analysis of seed development between species to identify similarities and differences in pathways important, for example, for the accumulation of the high level of proteins in the mature legume seed (Dam et al. 2009; Nautrup-Pedersen et al. 2010; Gallardo et al. 2003, 2007; Hajduch et al. 2005; Agrawal et al. 2008).

18.4 Proteomics of Lotus Nodules and Roots

The majority of legumes, including *Lotus*, have the ability to form symbiosis with rhizobia in specialized root nodule organs. *Lotus* with the sequenced diploid genome and *Lotus* retrotransposon 1 (LORE1) mutant population, currently with more than 80,000 lines, is an excellent model plant to study symbiosis (Sato et al. 2008; Urbanski et al. 2012). Using forward and reverse genetics, several genes essential for symbiosis have been identified in *Lotus*, *Medicago*, soybean, and pea (Madsen et al. 2003; Radutoiu et al. 2003; Limpens et al. 2003; Kouchi et al. 2010). Currently, proteomics is less used to study symbiosis; however, comparative proteomics of wild-type nodules and nodulation mutants can be essential to identify proteins affected/delayed in the functional nodule formation. In *Lotus*, 2D proteome reference maps of cytosolic nodule and root proteins were obtained. In total, 780 and 790 spots were identified from nodule and root corresponding to more than 800 different *Lotus* gene accessions with approximately 45 % intersection. Nodule and root master gels together with obtained data for each spot are available and can be further examined at https://www.cbs.dtu.dk/cgi-bin/lotus2_5/db.cgi (Dam et al. 2014). Furthermore, proteomics of the soybean cytosolic nodule fraction together with the plant and bacterial fractions of *Medicago* nodules were performed (Oehrle et al. 2008; Larrainzar et al. 2007) and the intersection of homologous proteins identified between nodule proteomics studies was determined (Dam et al. 2014).

To identify proteins important for transferring nutrients between the plant and symbiont, proteomics of enriched symbiosome membranes from *Lotus*, soybean, and pea was performed (Wienkoop and Saalbach 2003; Panter et al. 2000; Saalbach et al. 2002). One of the proteins identified from the *Lotus* symbiosome membrane was a predicted sulfate transporter and, subsequently,

the *Lotus symbiotic sulfate transporter* (*sst1*) mutant coding for that predicted sulfate transporter was identified to be crucial for nitrogen fixation (Wienkoop and Saalbach 2003; Krusell et al. 2005). *sst1* plants form ineffective nodules and cannot fix nitrogen efficiently, whereas the mutant grows normally under non-symbiotic conditions (Krusell et al. 2005). Thus, the proteomics data lead to the hypothesis that the plant sulfate transporter is localized on the symbiosome membrane and essential for transporting sulfate from plant cytoplasma to the rhizobia (Krusell et al. 2005).

Currently, most focus on comparative root proteomics has been on stress condition such as high salt, flooding, and temperature to identify regulated proteins in the defense pathways (Nanjo et al. 2012; Salavati et al. 2012; Dumont et al. 2011; Ahsan et al. 2010; Rodriguez-Celma et al. 2011), whereas approaches focusing on the symbiotic initiation of infection and organogenesis in the root infection zone are less developed. However, with the more sensitive proteomics methods, this can be used to elucidate novel knowledge at the protein level for initiating the infection and organogenesis pathways.

18.5 Proteomics of Post Translational Modifications in Lotus

The majority of proteins/enzymes have at least one PTM that can be essential for protein function, activity, stability, or degradation. One of the most common PTMs is *N*-glycosylation for which the functionality can be dependent or even independent of the protein carrier (Sumer-Bayraktar et al. 2011; Anthony and Ravetch 2010; Ohtsubo and Marth 2006). In *Lotus*, a glycomics study of the mature seed globulin fraction displayed a total of 19 different *N*-glycan structures including high mannosidic, pauci-mannosidic, and complex structures (Dam et al. 2013). The glycoproteomics data indicates that the high mannosidic structures are mainly linked to *Lotus* convicilin protein 2 (LCP2), pauci-mannosidic structures to a predicted lectin, and complex structures to a predicted peptidase (Dam et al. 2013 and unpublished data).

Protein phosphorylation/dephosphorylation is a common activation/deactivation mechanism for signaling cascades such as initiation of nodule formation together with a prompt response to pathogens. Thus, within the last decade, several plant phosphoproteomics studies were performed (Nakagami et al. 2010; Yang et al. 2013a, b; Wang et al. 2013). Proteomics and phosphoproteomics datasets of *Lotus* nodules, spontaneous nodules, and roots from wild-type and *spontaneous nodule formation 1* (*snf1*) plants have been generated and is currently being analyzed for proteins and phosphorylations needed for the infection and organogenesis pathways (unpublished data). Furthermore, in *Lotus* seedlings, 721 and 931 phosphopeptides were identified from the cotyledon and hypocotyl, respectively, with an overrepresentation of the GO term "RNA processing" in the hypocotyl (Ino et al. 2013).

18.6 Perspectives

In *Lotus*, proteome reference maps are available for seeds, pods, nodules, and roots together with relative quantification between developmental stages of all tissues analyzed. With use of the LORE1 resource (Urbanski et al. 2012; Fukai et al. 2012), containing thousands of mutant lines, comparative proteomics of wild type and different mutants of interest is now possible. This combined with the large number of transcriptomics data available for *Lotus* can reveal essential proteins for nodulation, infection, and the high protein level in the seed. At the PTM level, glycomics and glycoproteomics from LORE1 seeds which have a LORE1 insert in enzymes catalyze the *N*-glycan maturation has been initiated. In conclusion, all types of comparative proteomics analysis between wild-type and LORE1 mutants can be performed to identify difference at the proteome and/or PTM levels to increase the particular knowledge in areas of interest.

References

Agrawal GK, Hajduch M, Graham K, Thelen JJ (2008) In-depth investigation of the soybean seed-filling proteome and comparison with a parallel study of rapeseed. Plant Physiol 148(1):504–518

Ahsan N, Donnart T, Nouri MZ, Komatsu S (2010) Tissue-specific defense and thermo-adaptive mechanisms of soybean seedlings under heat stress revealed by proteomic approach. J Proteome Res 9 (8):4189–4204

Anthony RM, Ravetch JV (2010) A novel role for the IgG Fc glycan: the anti-inflammatory activity of sialylated IgG Fcs. J Clin Immunol 30:S9–S14

Credali A, Diaz-Quintana A, Garcia-Calderon M, De la Rosa MA, Marquez AJ, Vega JM (2011) Structural analysis of K^+ dependence in L-asparaginases from *Lotus japonicus*. Planta 234(1):109–122

Credali A, Garcia-Calderón M, Dam S, Perry J, Diaz-Quintana A, Parniske M, Wang TL, Stougaard J, Vega JM, Márquez AJ (2013) The K^+-dependent asparaginase, NSE1, is crucial for plant growth and seed production in *Lotus japonicus*. Plant Cell Physiol 54 (1):107–118

Dam S, Dyrlund TF, Ussatjuk A, Jochimsen B, Nielsen K, Goffard N, Ventosa M, Lorentzen A, Gupta V, Andersen SU, Enghild JJ, Ronson CW, Roepstorff P, Stougaard J (2014) Proteome reference maps of the *Lotus japonicus* nodule and root. Proteomics 14:230–240

Dam S, Laursen BS, Ørnfelt JH, Jochimsen B, Staerfeldt HH, Friis C, Nielsen K, Goffard N, Besenbacher S, Krusell L, Sato S, Tabata S, Thøgersen IB, Enghild JJ, Stougaard J (2009) The proteome of seed development in the model legume *Lotus japonicus*. Plant Physiol 149(3):1325–1340

Dam S, Thaysen-Andersen M, Stenkjaer E, Lorentzen A, Roepstorff P, Packer NH, Stougaard J (2013) Combined *N*-glycome and *N*-glycoproteome analysis of the *Lotus japonicus* seed globulin fraction shows conservation of protein structure and glycosylation in legumes. J Proteome Res 12:3383–3392

De Luis A, Markmann K, Cognat V, Holt DB, Charpentier M, Parniske M, Stougaard J, Voinnet O (2012) Two microRNAs linked to nodule infection and nitrogen-fixing ability in the legume *Lotus japonicus*. Plant Physiol 160(4):2137–2154

Djemel N, Guedon D, Lechevalier A, Salon C, Miquel M, Prosperi JM, Rochat C, Boutin JP (2005) Development and composition of the seeds of nine genotypes of the *Medicago truncatula* species complex. Plant Physiol Biochem 43(6):557–566 (PPB/Societe francaise de physiologie vegetale)

Dumont E, Bahrman N, Goulas E, Valot B, Sellier H, Hilbert JL, Vuylsteker C, Lejeune-Hénaut I, Delbreil B (2011) A proteomic approach to decipher chilling response from cold acclimation in pea (*Pisum sativum* L.). Plant Sci Int J Experim Plant Biol 180(1):86–98

Dyrlund TF, Poulsen ET, Scavenius C, Sanggaard KW, Enghild JJ (2012) MS data miner: a web-based software tool to analyze, compare, and share mass spectrometry protein identifications. Proteomics 12 (18):2792–2796

Fukai E, Soyano T, Umehara Y, Nakayama S, Hirakawa H, Tabata S, Sato S, Hayashi M (2012) Establishment of a *Lotus japonicus* gene tagging population using the exon-targeting endogenous retrotransposon *LORE1*. Plant J 69:720–730

Gallardo K, Firnhaber C, Zuber H, Hericher D, Belghazi M, Henry C, Kuster H, Thompson R (2007) A combined proteome and transcriptome analysis of developing *Medicago truncatula* seeds: evidence for metabolic specialization of maternal and filial tissues. Mol Cell Proteomics (MCP) 6(12):2165–2179

Gallardo K, Le Signor C, Vandekerckhove J, Thompson RD, Burstin J (2003) Proteomics of *Medicago truncatula* seed development establishes the time frame of diverse metabolic processes related to reserve accumulation. Plant Physiol 133(2):664–682

Görg A, Weiss W, Dunn MJ (2004) Current two-dimensional electrophoresis technology for proteomics. Proteomics 4(12):3665–3685

Hajduch M, Ganapathy A, Stein JW, Thelen JJ (2005) A systematic proteomic study of seed filling in soybean. Establishment of high-resolution two-dimensional reference maps, expression profiles, and an interactive proteome database. Plant Physiol 137(4):1397–1419

Høgslund N, Radutoiu S, Krusell L, Voroshilova V, Hannah MA, Goffard N, Sanchez DH, Lippold F, Ott T, Sato S, Tabata S, Liboriussen P, Lohmann GV, Schauser L, Weiller GF, Udvardi MK, Stougaard J (2009) Dissection of symbiosis and organ development by integrated transcriptome analysis of *Lotus japonicus* mutant and wild-type plants. PLoS ONE 4 (8):e6556

Ino Y, Ishikawa A, Nomura A, Kajiwara H, Harada K, Hirano H (2013) Phosphoproteome analysis of *Lotus japonicus* seeds. Proteomics 14(1):116

Jellouli N, Salem AB, Ghorbel A, Jouira HB (2010) Evaluation of protein extraction methods for *Vitis vinifera* leaf and root proteome analysis by two-dimensional electrophoresis. J Integr Plant Biol 52 (10):933–940

Kouchi H, Imaizumi-Anraku H, Hayashi M, Hakoyama T, Nakagawa T, Umehara Y, Suganuma N, Kawaguchi M (2010) How many peas in a pod? Legume genes responsible for mutualistic symbioses underground. Plant Cell Physiol 51(9):1381–1397

Krusell L, Krause K, Ott T, Desbrosses G, Krämer U, Sato S, Nakamura Y, Tabata S, James EK, Sandal N, Stougaard J, Kawaguchi M, Miyamoto A, Suganuma N, Udvardi MK (2005) The sulfate transporter SST1 is crucial for symbiotic nitrogen fixation in *Lotus japonicus* root nodules. Plant Cell 17(5):1625–1636

Larrainzar E, Wienkoop S, Weckwerth W, Ladrera R, Arrese-Igor C, Gonzalez EM (2007) *Medicago truncatula* root nodule proteome analysis reveals

differential plant and bacteroid responses to drought stress. Plant Physiol 144(3):1495–1507

Limpens E, Franken C, Smit P, Willemse J, Bisseling T, Geurts R (2003) LysM domain receptor kinases regulating rhizobial Nod factor-induced infection. Science 302(5645):630–633

Madsen EB, Madsen LH, Radutoiu S, Olbryt M, Rakwalska M, Szczyglowski K, Sato S, Kaneko T, Tabata S, Sandal N, Stougaard J (2003) A receptor kinase gene of the LysM type is involved in legume perception of rhizobial signals. Nature 425 (6958):637–640

Nakagami H, Sugiyama N, Mochida K, Daudi A, Yoshida Y, Toyoda T, Tomita M, Ishihama Y, Shirasu K (2010) Large-scale comparative phosphoproteomics identifies conserved phosphorylation sites in plants. Plant Physiol 153(3):1161–1174

Nanjo Y, Skultety L, Uvackova L, Klubicova K, Hajduch M, Komatsu S (2012) Mass spectrometry-based analysis of proteomic changes in the root tips of flooded soybean seedlings. J Proteome Res 11 (1):372–385

Nautrup-Pedersen G, Dam S, Laursen BS, Siegumfeldt AL, Nielsen K, Goffard N, Staerfeldt HH, Friis C, Sato S, Tabata S, Lorentzen A, Roepstorff P, Stougaard J (2010) Proteome analysis of pod and seed development in the model legume *Lotus japonicus*. J Proteome Res 9(11):5715–5726

Oehrle NW, Sarma AD, Waters JK, Emerich DW (2008) Proteomic analysis of soybean nodule cytosol. Phytochemistry 69(13):2426–2438

O'Farrell PH (1975) High resolution two-dimensional electrophoresis of proteins. J Biol Chem 250 (10):4007–4021

Ohtsubo K, Marth JD (2006) Glycosylation in cellular mechanisms of health and disease. Cell 126 (5):855–867

Panter S, Thomson R, de Bruxelles G, Laver D, Trevaskis B, Udvardi M (2000) Identification with proteomics of novel proteins associated with the peribacteroid membrane of soybean root nodules. Mol Plant Microbe Interact (MPMI) 13(3):325–333

Parker CE, Warren MR, Mocanu V (2010) Mass spectrometry for proteomics. In: Alzate O (ed) Neuroproteomics. Taylor and Francis Group, Boca Raton

Prakash D, Misra PS (1988) Protein content and amino acid profile of some wild leguminous seeds. Plant Foods Human Nutr 38(1):61–65

Radutoiu S, Madsen LH, Madsen EB, Felle HH, Umehara Y, Gronlund M, Sato S, Nakamura Y, Tabata S, Sandal N, Stougaard J (2003) Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases. Nature 425(6958):585–592

Rodriguez-Celma J, Lattanzio G, Grusak MA, Abadia A, Abadia J, Lopez-Millan AF (2011) Root responses of *Medicago truncatula* plants grown in two different iron deficiency conditions: changes in root protein profile and riboflavin biosynthesis. J Proteome Res 10 (5):2590–2601

Rogowska-Wrzesinska A, Le Bihan MC, Thaysen-Andersen M, Roepstorff P (2013) 2D gels still have a niche in proteomics. J Proteomics 88:4–13

Saalbach G, Erik P, Wienkoop S (2002) Characterisation by proteomics of peribacteroid space and peribacteroid membrane preparations from pea (*Pisum sativum*) symbiosomes. Proteomics 2(3):325–337

Salavati A, Khatoon A, Nanjo Y, Komatsu S (2012) Analysis of proteomic changes in roots of soybean seedlings during recovery after flooding. J Proteomics 75(3):878–893

Sato S, Nakamura Y, Kaneko T, Asamizu E, Kato T, Nakao M, Sasamoto S, Watanabe A, Ono A, Kawashima K, Fujishiro T, Katoh M, Kohara M, Kishida Y, Minami C, Nakayama S, Nakazaki N, Shimizu Y, Shinpo S, Takahashi C, Wada T, Yamada M, Ohmido N, Hayashi M, Fukui K, Baba T, Nakamichi T, Mori H, Tabata S (2008) Genome structure of the legume, *Lotus japonicus*. DNA Res 15(4):227–239

Sumer-Bayraktar Z, Kolarich D, Campbell MP, Ali S, Packer NH, Thaysen-Andersen M (2011) *N*-glycans modulate the function of human corticosteroid-binding globulin. Mol Cell Proteomics (MCP) 10(8): M111.009100-1–M111.009100-14

Urbanski D, Malolepszy A, Stougaard J, Andersen S (2012) Genome-wide *LORE1* retrotransposon mutagenesis and high-throughput insertion detection in *Lotus japonicus*. Plant J 69:731–741

Verdier J, Torres-Jerez I, Wang M, Andriankaja A, Allen SN, He J, Tang Y, Murray JD, Udvardi MK (2013) Establishment of the *Lotus japonicus* gene expression atlas (LjGEA) and its use to explore legume seed maturation. Plant J Cell Mol Biol 74(2):351–362

Wang P, Xue L, Batelli G, Lee S, Hou YJ, Van Oosten MJ, Zhang H, Tao WA, Zhu JK (2013) Quantitative phosphoproteomics identifies SnRK2 protein kinase substrates and reveals the effectors of abscisic acid action. Proc Natl Acad Sci USA 110 (27):11205–11210

Wang W, Scali M, Vignani R, Spadafora A, Sensi E, Mazzuca S, Cresti M (2003) Protein extraction for two-dimensional electrophoresis from olive leaf, a plant tissue containing high levels of interfering compounds. Electrophoresis 24(14):2369–2375

Wasinger VC, Cordwell SJ, Cerpa-Poljak A, Yan JX, Gooley AA, Wilkins MR, Duncan MW, Harris R, Williams KL, Humphery-Smith I (1995) Progress with gene-product mapping of the mollicutes: *Mycoplasma genitalium*. Electrophoresis 16(1):1090–1094

Waterhouse RN, Smyth AJ, Massonneau A, Prosser IM, Clarkson DT (1996) Molecular cloning and characterisation of asparagine synthetase from *Lotus japonicus*: dynamics of asparagine synthesis in N-sufficient conditions. Plant Mol Biol 30(5):883–897

Weiss W, Görg A (2007) Two-dimensional electrophoresis for plant proteomics. Methods Mol Biol 355:121–143

Wienkoop S, Saalbach G (2003) Proteome analysis. Novel proteins identified at the peribacteroid

membrane from *Lotus japonicus* root nodules. Plant Physiol 131(3):1080–1090

Wilson LA, Birmingham VA, Moon DP, Snyder HE (1978) Isolation and characterization of starch from mature soybeans. Cereal Chem 55(5):661–670

Xie H, Pan S, Liu S, Ye K, Huo K (2007) A novel method of protein extraction from perennial *Bupleurum* root for 2-DE. Electrophoresis 28(5):871–875

Yang F, Melo-Braga MN, Larsen MR, Jørgensen HJ, Palmisano G (2013a) Battle through signaling between wheat and the fungal pathogen *Septoria tritici* revealed by proteomics and phosphoproteomics. Mol Cell Proteom (MCP) 12(9):2497–2508

Yang Z, Guo G, Zhang M, Liu CY, Hu Q, Lam H, Cheng H, Xue Y, Li J, Li N (2013b) Stable isotope metabolic labeling-based quantitative phosphoproteomic analysis of Arabidopsis mutants reveals ethylene-regulated time-dependent phosphoproteins and putative substrates of constitutive triple response 1 kinase. Mol Cell Proteom (MCP) 12(12):3559–3582

Part IV
Resources

Wild Accessions and Mutant Resources 19

Masayoshi Kawaguchi and Niels Sandal

Abstract

Lotus japonicus, *Lotus burttii,* and *Lotus filicaulis* are species of *Lotus* genus that are utilized for molecular genetic analysis such as the construction of a linkage map and QTL analysis. Among them, a number of mutants have been isolated from two wild accessions: *L. japonicus* Gifu B-129 and Miyakojima MG-20. Here, we show the wild accessions and a list of all mutants isolated so far.

19.1 Wild Accessions and Isolated Mutants

Lotus japonicus (Regel) Larsen is widely distributed in Eastern Asia, China, Korea, Taiwan, and Japan (Grant et al. 1962). *L. japonicus* wild-type accession Gifu B-129 is most frequently used for genetic analysis and was originally collected on a riverbank in Gifu, a prefecture located in the Chubu region of central Japan. A number of mutants have been isolated from Gifu as shown in the next paragraph. Another accession Miyakojima MG-20 is often used as a crossing partner with Gifu to produce F2 mapping populations (Hayashi et al. 2001). MG-20 is an early flowering ecotype suitable for indoor cultivation and genetic analysis (Kawaguchi 2000). From a comparative anatomical study, it has been proposed that MG-20 be described as a separate species, *Lotus miyakojimae* (Kramina) (Barykina and Kramina 2006). To allow map-based cloning and QTL analysis, two related diploid species *Lotus filicaulis* B-37, originating from Algeria, and *Lotus burttii* B-303, from West Pakistan, have been used as a crossing partner (Grant et al. 1962; Borsos et al. 1972; Sandal et al. 2002; Kawaguchi et al. 2005). Compared with MG-20 and *L. filicaulis*, *L. burttii* shows an intermediate level of polymorphism with respect to Gifu (Sandal et al. 2012).

These wild-type accessions, with the exception of *L. filicaulis,* are available from the National Bio Resource Project (NBRP) *Lotus* and *Glycine.* To assist with GWAS analysis in the future, many more wild accessions of *L. japonicus* will need to be collected from all over the world.

In order to unveil characteristics of legumes at a molecular level, a large number of *L. japonicus* mutants, including those created by transposon

M. Kawaguchi (✉)
Division of Symbiotic Systems, National Institute for Basic Biology, Okazaki, Aichi 444-8585, Japan
e-mail: masayosi@nibb.ac.jp

N. Sandal
Centre for Carbohydrate Recognition and Signalling, Department of Molecular Biology and Genetics, Aarhus University, Gustav Wieds Vej 10, 8000 Aarhus C, Denmark

S. Tabata and J. Stougaard (eds.), *The Lotus japonicus Genome*, Compendium of Plant Genomes,
DOI 10.1007/978-3-662-44270-8_19,

Table 19.1 *Lotus japonicus* mutants

Locus	Allele	Previous name	Nodulation phenotype	AM phenotype	References
Symbiotic mutants					
Nfr1	*nfr1-1, 2*	*sym1-1, 2*	−	+	Schauser et al. (1998), Radutoiu et al. (2003)
	nfr1-3	*sym83*	−	+	Kawaguchi et al. (2002), Sandal et al. (2006)
	nfr1-4	*sym106*			Sandal et al. (2006)
	nfr1-5 to *−13*				Perry et al. (2009)
Nfr5	*nfr5-1*	*sym5*	−	+	Schauser et al. (1998), Madsen et al. (2003)
	nfr5-2		−	+	Madsen et al. (2003), Sandal et al. (2006)
	nfr5-3	*sym25*	−	+	Szczyglowski et al. (1998), Madsen et al. (2003), Sandal et al. (2006)
	nfr5-4 to *6*				Murray et al. (2006), Sandal et al. (2006)
	nfr5-7 to *10*				Perry et al. (2009)
SymRK	*symrk-1*	*sym2-1*	−	−	Schauser et al. (1998), Stracke et al. (2002)
	symrk-2		−	−	Schauser et al. (1998), Stracke et al. (2002)
	symrk-3	*sym2*	−	−	Stracke et al. (2002), Kistner et al. (2005a, b)
	symrk-4, 5		−	−	Sandal et al. (2006)
	symrk-6, 7	*sym21-1, 2*	−	n.d.	Szczyglowski et al. (1998), Stracke et al. (2002)
	symrk-8 to *11*				Perry et al. (2003)
	symrk-12	*sym84*	−	−	Kawaguchi et al. (2002), Sandal et al. (2006)
	symrk-13, 14		−	−	Murray et al. (2006), Kosuta et al. (2011)
	symrk-15 to *60*				Perry et al. (2009)
Castor	*castor-1*	*sym4-1*	−	+	Schauser et al. (1998), Bonfante et al. (2000), Kistner et al. (2005a, b)
	castor-2	*sym4-2*	−	+	Bonfante et al. (2000), Kistner et al. (2005a, b)
	castor-3	*sym22-1*	−	+	Szczyglowski et al. (1998), Kistner et al. (2005a, b)
	castor-4, 5	*sym71-1, 2*	−	+	Senoo et al. (2000), Kawaguchi et al. (2002), Imaizumi-Anraku et al. (2005)
	castor-6 to *17*				Imaizumi-Anraku et al. (2005)
	castor-18 to *20*				Sandal et al. (2006)
	castor-21 to *25*				Murray et al. (2006)
	castor-26 to *36*				Perry et al. (2009)

(continued)

Table 19.1 (continued)

Locus	Allele	Previous name	Nodulation phenotype	AM phenotype	References
Pollux	*pollux-1, 2*	*sym23-1, 2*	−	−	Szczyglowski et al. (1998), Imaizumi-Anraku et al. (2005)
	pollux-3	*sym86*	−	−	Kawaguchi et al. (2002), Imaizumi-Anraku et al. (2005)
	pollux-4 to *10*		−	−	Imaizumi-Anraku et al. (2005)
	pollux-11 to *−15*				Murray et al. (2006)
	pollux-16 to *−30*				Perry et al. (2009)
Nup133	*nup133-1, 2*	*sym3-1,2*	−	−	Schauser et al. (1998), Kanamori et al. (2006)
	nup133-3	*sym3-3*	−	n.d.	Kanamori et al. (2006)
	nup133-4	*sym26*	−	−	Szczyglowski et al. (1998), Sandal et al. (2006)
	nup133-5 to *−9*				Murray et al. (2006)
	nup133-10 to *−21*				Perry et al. (2009)
Nup85	*nup85-1*	*sym24*	−	−	Szczyglowski et al. (1998), Saito et al. (2007)
	nup85-2	*sym73*	−	∓	Kawaguchi et al. (2002), Saito et al. (2007)
	nup85-3	*sym85*	−	−	Kawaguchi et al. (2002), Saito et al. (2007)
	nup85-4		−	−	Murray et al. (2007), Saito et al. (2007)
	nup85-5		−	−	Kistner et al. (2005a, b), Perry et al. (2009)
	nup85-6 to *−10*				Perry et al. (2009)
Nena	*nena-1, 2*		−, t.s.	∓, t.s.	Groth et al. (2010)
	nena-3		∓, t.s.	n.d.	Groth et al. (2010)
	nena-4, 5		+	n.d.	Groth et al. (2010)
	nena-6[a]		−	−	Groth et al. (2010)
CCaMK	*ccamk-1,2*	*sym15-1,2*	−	−	Schauser et al. (1998), Tirichine et al. (2006a)
	ccamk-3	*sym72*	−	−	Senoo et al. (2000), Kawaguchi et al. (2002), Tirichine et al. (2006a)
	ccamk-4				Tirichine et al. (2006a)
	ccamk-5,6				Murray et al. (2006)
	ccamk-7 to *−13*				Perry et al. (2009)
Cyclops	*cyclops-1, 2*	*sym6-1,2*	−	−	Schauser et al. (1998), Kistner et al. (2005a, b), Yano et al. (2008)
	cyclops-3	*sym30*	−	−	Szczyglowski et al. (1998), Kistner et al. (2005a, b), Yano et al. (2008)
	cyclops-4	*sym82*	∓	−	Kawaguchi et al. (2002), Yano et al. (2008)
	cyclops-5		∓		Yano et al. (2008), Perry et al. (2009)
	cyclops-6, −7				Perry et al. (2009)
	cyclops-6[a]		∓		Suzaki et al. (2012)

(continued)

Table 19.1 (continued)

Locus	Allele	Previous name	Nodulation phenotype	AM phenotype	References
Hit1	*hit1*		∓	−	Murray et al. (2007)
Tco	*tco*[a]		∓		Suzaki et al. (2013)
Nsp1	*nsp1-1*		−	+	Heckmann et al. (2006)
	nsp1-2 to *−5*				Perry et al. (2009)
	nsp1-6[a], *−7*[a]				Takeda et al. (2013)
Nsp2	*nsp2-1*	*sym70*	−	+	Kawaguchi et al. (2002), Murakami et al. (2006), Heckmann et al. (2006)
	nsp2-2	*sym35*	−	+	Murakami et al. (2006), Heckmann et al. (2006)
	nsp2-3		∓	+	Heckmann et al. (2006)
	nsp2-4, −8 to *−11*				Perry et al. (2009)
	nsp2-5 to *−7*				Murray et al. (2006)
	nsp2-12[a]				Takeda et al. (2013)
Nin	*nin-1*	*sym20*	−	+	Schauser et al. (1999)
	nin-2 to *3*		−	+	Schauser et al. (1999)
	nin-4 to 6		+	+	
	nin-7		−	+	Sandal et al. (2006)
	nin-8				Murray et al. (2006)
	nin-9[a]		−		Suzaki et al. (2012)
	nin-9 to *−14*				Perry et al. (2009)
Cerberus	*cerberus-1*	*sym7*	∓		Schauser et al. (1998), Yano et al. (2009)
	cerberus-2	*sym41*	∓		Yano et al. (2009)
	cerberus-3	*sym55*	∓		Yano et al. (2009)
	cerberus-4	*sym57*	∓		Yano et al. (2009)
	cerberus-5	*sym101*	∓		Yano et al. (2009)
	cerberus-6	*LjS28-2B*	∓		Yano et al. (2009)
	cerberus-7 to *11*				Sandal et al. (2006), Yano et al. (2009)
Crinkle	*crinkle-1*	*sym79-1*	∓	+	Kawaguchi et al. (2002), Tansengco et al. (2003)
	crinkle-2, 3	*sym79-2, 3*	∓	+	Yano et al. (2006)
Alb1	*alb1-1*	*sym74-1*	∓	+	Imaizumi-Anraku et al. (1997), Kawaguchi et al. (2002)
	alb1-2, 3	*sym74-2, 3*	∓	+	Kawaguchi et al. (2002), Yano et al. (2006)
Sym8	*sym8*		∓		Schauser et al. (1998), Sandal et al. (2006)
Sym9	*sym9*		∓		Schauser et al. (1998)
Sym10	*sym10*		∓		Schauser et al. (1998), Sandal et al. (2012)
Sst1	*sst1-1*	*sym13*	Fix−	+	Schauser et al. (1998), Krusell et al. 2005
	sst1-2	*sym81*			Senoo et al. 2000; Kawaguchi et al. 2002; Krusell et al. 2005

(continued)

Table 19.1 (continued)

Locus	Allele	Previous name	Nodulation phenotype	AM phenotype	References
Prh1			Fix−		Karas et al. (2005)
Ign1	*ign1*		Fix−	+	Kumagai et al. (2007)
Fen1	*fen1-1*	*sym76*	Fix−	+	Imaizumi-Anraku et al. (1997), Kawaguchi et al. (2002), Hakoyama et al. (2009)
	fen1-2				Hakoyama et al. (2009)
Sen1	*sen1-1*	*sym75*	Fix−	+	Kawaguchi et al. (2002), Suganuma et al. (2003), Hakoyama et al. (2012a)
	sen1-2				Hakoyama et al. (2012a)
	sen1-3, 4	*sym11, sym61*			Schauser et al. (1998), Sandal et al. (2006), Hakoyama et al. (2012a)
	sen1-5				Hakoyama et al. (2012a)
Sym105			Fix−	+	Hossain et al. (2006)
Syp71	*syp71-1*	*EMS*	Fix−		Hakoyama et al. (2012b)
	syp71-2	*ion beam*			Hakoyama et al. (2012b)
Sym12			Fix−		Schauser et al. (1998)
Sym14			Fix−		Schauser et al. (1998)
Sym43	*sym43-1*	*sym43*	Fix∓	+	Sandal et al. (2006)
	sym43-2	*sym103*			Sandal et al. (2006)
Sym102			Fix−		Sandal et al. (2006)
Sym104			Fix−	+	Sandal et al. (2006)
Gln2	*gln2-1*		Fix−		Oreo et al. (2002), Betti et al. (2006), García-Calderón et al. (2012)
	gln2-2		Fix−		Oreo et al. (2002), Betti et al. (2006), García-Calderón et al. (2012)
Enf1	*enf1*		Fix++		Tominaga et al. (2009)
Snf1	*snf1*		Snf+		Tirichine et al. (2006a, b)
Snf2	*snf2-1, 2*		Snf+		Tirichine et al. (2006b), Tirichine et al. (2007)
Snf4	*snf4*		Snf+		Tirichine et al. (2006b)
Har1	*har1-1*	*sym16*	++		Schauser et al. (1998), Krusell et al. (2002)
	har1-2, 3	*sym34-1, 2*	++		Szczyglowski et al. (1998), Krusell et al. (2002)
	har1-4, 5	*sym78-1, 2*	++		Kawaguchi et al. (2002), Nishimura et al. (2002a)
	har1-6		++		Sandal et al. (2006)
	har1-7[a]		++		Miyazawa et al. (2010)
	har1-8[a]		++		Suzaki et al. (2012)
Klavier	*klavier*[a]		++		Miyazawa et al. (2010)

(continued)

Table 19.1 (continued)

Locus	Allele	Previous name	Nodulation phenotype	AM phenotype	References
TML	*tml-1*[a] to *3*[a]		++		Magori et al. (2009), Takahara et al. (2013)
	tml-4[a]	*rdh1*	++		Yokota et al. (2009), Takahara et al. (2013)
PLENTY	*plenty*[a]		++		Yoshida et al. (2010)
Clv2	*clv2-1*		+(+)		Krusell et al. (2011)
SL0154-N			+	∓	Groth et al. (2013)
SL0181-N			+	∓	Groth et al. (2013)
Developmental mutants affecting symbiosis					
Astray	*astray*	*sym77*	+(+)		Nishimura et al. (2002b, c)
Arpc1	*arpc1*		∓	+	Hossain et al. (2012)
Brush	*brush*		∓		Maekawa-Yoshikawa et al. (2009)
Ccw	*ccw*		Low nod		Kawaguchi et al. (2002)
Nap1	*nap1-1* to *3*		∓		Yokota et al. (2009)
Pir1	*pir1-1*	*sym40*	∓		Yokota et al. (2009)
	pir1-2				Yokota et al. (2009)
	pir1-3	*sym80*	∓		Yokota et al. (2009)
	pir1-4, 5				Yokota et al. (2009)
Rhl1	*rhl1-1*		Low nod		Karas et al. (2006, 2009)
	rhl1-2	*Slippery root*	Low nod		Kawaguchi et al. (2002), Karas et al. (2009)
	rhl1-3				Karas et al. (2009)
PhyB	*phyB-1, 2*		∓		Suzuki et al. (2011)

Locus	Allele	Affected in	References
Non-symbiotic mutants			
Kew1	*kew1*	Flower development	Feng et al. (2006)
Pfo	*pfo-1*	Flower development	Zhang et al. (2003)
	pfo-2		Dong et al. (2005), Wang et al. (2013)
Pfm	*pfm*	Flower development	Dong et al. (2005)
Squ1	*squ1*	Flower development	Feng et al. (2006)
Wps1	*wps1*	Flower development	Chen et al. (2006)
Sac	*sac*	Axillary shoot meristems	de Alvarez et al. (2006)
FUL1	*ful1-1*	Leaf shape	Wang et al. (2013)
Rel1	*rel1*	Leaf shape	Yan et al. (2010)
Rel3	*rel3*	Leaf shape	Yan et al. (2010)
UML1	*uml1*	Leaf shape	Wang et al. (2013)
Slp	*sleepless*	Nyctinastic leaf movement	Kawaguchi (2003), Chen et al. (2012)
APL1	*apl1-1, 2*	Starch synthesis	Vriet et al. (2010)
APL2	*apl2-1, 3, 4*	Starch synthesis	Vriet et al. (2010)
APS1	*aps1-1* to *3*	Starch synthesis	Vriet et al. (2010)
Cyd1	*cyd1*	Cyanogenesis	Takos et al. (2010, 2011)
Cyd2	*cyd2*	Cyanogenesis	Takos et al. (2010)

(continued)

Table 19.1 (continued)

Locus	Allele	Affected in	References
GWD1	*gwd1-1, 2*	Starch degradation	Vriet et al. (2010)
GWD3	*gwd3-1, 4*	Starch degradation	Vriet et al. (2010)
PGI	*pgi1-1* to *3*	Starch synthesis	Vriet et al. (2010)
PGM1	*pgm1-3* to *5*	Starch synthesis	Vriet et al. (2010)
Sus1	*sus1-1*	Sucrose synthase	Horst et al. (2007)
Sus3	*sus3-1*	Sucrose synthase	Horst et al. (2007)

t.s. temperature sensitive
Snf represents spontaneous nodule formation
[a] Indicates mutants derived from *L. japonicus* MG-20

insertion and chemical/irradiation mutagenesis, have been generated by research groups around the world. In Table 19.1, we summarize the current status of all mutant lines isolated by forward genetic screens and TILLING analysis. These lines are grouped into 3 categories: symbiotic mutants, developmental mutants affecting symbiosis, and non-symbiotic mutants. This list will be useful not only for research but also for the maintenance of genetic resources of *L. japonicus*.

References

Barykina RP, Kramina TE (2006) A comparative morphological and anatomical study of the model legume *Lotus japonicus* and related species. Wulfenia 13:33–56

Betti M, Arcondeguy T, Marquez AJ (2006) Molecular analysis of two mutants from *Lotus japonicus* deficient in plastidic glutamine synthetase: functional properties of purified GLN2 enzymes. Planta 224:1068–1079

Bonfante P, Genre A, Faccio A et al (2000) The *Lotus japonicus LjSym4* gene is required for the successful symbiotic infection of root epidermal cells. Mol Plant Microbe Interact 13:1109–1120

Borsos O, Somaroo BH, Grant WF (1972) A new diploid species of *Lotus* (Leguminosae) in Pakistan. Can J Bot 50:1865–1870

Chen JH, Pang JL, Wang LL et al (2006) Wrinkled petals and stamens 1, is required for the morphogenesis of petals and stamens in *Lotus japonicus*. Cell Res 16:499–506

Chen T, Zhu H, Ke DX et al (2012) A MAP kinase kinase interacts with SymRK and regulates nodule organogenesis in *Lotus japonicus*. Plant Cell 24:823–838

de Alvarez NDG, Meeking RJ, White DWR et al (2006) The origin, initiation and development of axillary shoot meristems in *Lotus japonicus*. Ann Bot 98:953–963

Dong ZC, Zhao Z, Liu CW et al (2005) Floral patterning in *Lotus japonicus*. Plant Physiol 137:1272–1282

Feng XZ, Zhao Z, Tian ZX et al (2006) Control of petal shape and floral zygomorphy in *Lotus japonicus*. Proc Natl Acad Sci USA 103:4970–4975

García-Calderón M, Chiurazzi M, Espuny MR et al (2012) Photorespiratory metabolism and nodule function: behavior of *Lotus japonicus* mutants deficient in plastid glutamine synthetase. Mol Plant Microbe Interact 25:211–219

Grant WF, Bullen MR, de Nettancourt D (1962) The cytogenetics of Lotus. I Embryo-cultured interspecific diploid hybrids closely related to *L. corniculatus* L. Can J Genet Cytol 4:105–128

Groth M, Takeda N, Perry J et al (2010) NENA, a *Lotus japonicus* homolog of Sec13, is required for Rhizodermal infection by Arbuscular Mycorrhiza Fungi and Rhizobia but dispensable for cortical endosymbiotic development. Plant Cell 22:2509–2526

Groth M, Kosuta S, Gutjahr C et al (2013) Two *Lotus japonicus* symbiosis mutants impaired at distinct steps of arbuscule development. Plant J 75:117–129

Hakoyama T, Niimi K, Watanabe H et al (2009) Host plant genome overcomes the lack of a bacterial gene for symbiotic nitrogen fixation. Nature 462:514–517

Hakoyama T, Niimi K, Yamamoto T et al (2012a) The integral membrane protein SEN1 is required for symbiotic nitrogen fixation in *Lotus japonicus* nodules. Plant Cell Physiol 53:225–236

Hakoyama T, Oi R, Hazuma K et al (2012b) The SNARE Protein SYP71 expressed in vascular tissues is involved in symbiotic nitrogen fixation in *Lotus japonicus* nodules. Plant Physiol 160:897–905

Hayashi M, Miyahara A, Sato S et al (2001) Construction of a genetic linkage map of the model legume *Lotus japonicus* using an intraspecific F2 population. DNA Res 8:301–310

Heckmann AB, Lombardo F, Miwa H et al (2006) Lotus japonicus nodulation requires two GRAS domain regulators, one of which is functionally conserved in a non-legume. Plant Physiol 142:1739–1750

Horst I, Welham T, Kelly S et al (2007) TILLING mutants of *Lotus japonicus* reveal that nitrogen assimilation and fixation can occur in the absence of nodule-enhanced sucrose synthase. Plant Physiol 144:806–820

Hossain MS, Umehara Y, Kouchi H (2006) A novel Fix (−) symbiotic mutant of *Lotus japonicus*, Ljsym105, shows impaired development and premature deterioration of nodule infected cells and symbiosomes. Mol Plant Microbe Interact 19:780–788

Hossain MS, Liao JQ, James EK et al (2012) *Lotus japonicus* ARPC1 is required for rhizobial infection. Plant Physiol 160:917–928

Imaizumi-Anraku H, Kawaguchi M, Koiwa H et al (1997) Two ineffective-nodulating mutants of *Lotus japonicus*—Different phenotypes caused by the blockage of endocytotic bacterial release and nodule maturation. Plant Cell Physiol 38:871–881

Imaizumi-Anraku H, Takeda N, Charpentier M et al (2005) Plastid proteins crucial for symbiotic fungal and bacterial entry into plant roots. Nature 433:527–531

Kanamori N, Madsen LH, Radutoiu S et al (2006) A nucleoporin is required for induction of Ca^{2+} spiking in legume nodule development and essential for rhizobial and fungal symbiosis. Proc Natl Acad Sci USA 103:359–364

Karas B, Murray J, Gorzelak M et al (2005) Invasion of *Lotus japonicus* root hairless 1 by Mesorhizobium loti involves the nodulation factor-dependent induction of root hairs. Plant Physiol 137:1331–1344

Karas B, Amyot L, Johansen C et al (2009) Conservation of Lotus and Arabidopsis basic helix-loop-helix proteins reveals new players in root hair development. Plant Physiol 151:1175–1185

Kawaguchi M (2000) *Lotus japonicus* "Miyakojima" MG-20: an early flowering accession suitable for indoor handling. J Plant Res 113:507–509

Kawaguchi M (2003) *SLEEPLESS*, a gene conferring nyctinastic movement in legume. J Plant Res 116:151–154

Kawaguchi M, Imaizumi-Anraku H, Koiwa H et al (2002) Root, root hair, and symbiotic mutants of the model legume *Lotus japonicus*. Mol Plant Microbe Interact 15:17–26

Kawaguchi M, Pedorosa-Harand A, Yano K et al (2005) *Lotus burttii t*ake a position of the third corner in the *Lotus* molecular genetics triangle. DNA Res 12:69–77

Kistner C, Winzer T, Pitzschke A et al (2005a) Seven *Lotus japonicus* genes required for transcriptional reprogramming of the root during fungal and bacterial symbiosis. Plant Cell 17:2217–2229

Kistner C, Winzer T, Pitzschke A et al (2005b) Seven *Lotus japonicus* genes required for transcriptional reprogramming of the root during fungal and bacterial symbiosis. Plant Cell 17:2217–2229

Kosuta S, Held M, Hossain MS et al (2011) *Lotus japonicus symRK-14* uncouples the cortical and epidermal symbiotic program. Plant J 67:929–940

Krusell L, Madsen LH, Sato S et al (2002) Shoot control of root development and nodulation is mediated by a receptor-like kinase. Nature 420:422–426

Krusell L, Krause K, Ott T et al (2005) The sulfate transporter SST1 is crucial for symbiotic nitrogen fixation in *Lotus japonicus* root nodules. Plant Cell 17:1625–1636

Krusell L, Sato N, Fukuhara I et al (2011) The *Clavata2* genes of pea and *Lotus japonicus* affect autoregulation of nodulation. Plant J 65:861–871

Kumagai H, Hakoyama T, Umehara Y et al (2007) A novel ankyrin-repeat membrane protein, IGN1, is required for persistence of nitrogen-fixing symbiosis in root nodules of *Lotus japonicus*. Plant Physiol 143:1293–1305

Madsen EB, Madsen LH, Radutoiu S et al (2003) A receptor kinase gene of the LysM type is involved in legume perception of rhizobial signals. Nature 425:637–640

Maekawa-Yoshikawa M, Muller J, Takeda N (2009) The temperature-sensitive brush mutant of the legume *Lotus japonicus* reveals a link between root development and nodule infection by rhizobia. Plant Physiol 149:1785–1796

Magori S, Oka-Kira E, Shibata S et al (2009) TOO MUCH LOVE, a root regulator associated with the long-distance control of nodulation in *Lotus japonicus*. Mol Plant Microbe Interact 22:259–268

Miyazawa H, Oka-Kira E, Sato N et al (2010) The receptor-like kinase KLAVIER mediates systemic regulation of nodulation and non-symbiotic shoot development in *Lotus japonicus*. Development 137:4317–4325

Murakami Y, Miwa H, Imaizumi-Anraku H et al (2006) Positional cloning identifies *Lotus japonicus* NSP2, a putative transcription factor of the GRAS Family, required for NIN and ENOD40 Gene expression in nodule initiation. DNA Res 13:255–265

Murray J, Karas B, Ross L et al (2006) Genetic suppressors of the *Lotus japonicus* har1-1 hypernodulation phenotype. Mol Plant Microbe Interact 19:1082–1091

Murray JD, Karas BJ, Sato S et al (2007) A cytokinin perception mutant colonized by rhizobium in the absence of nodule organogenesis. Science 315:101–104

Nishimura R, Hayashi M, Wu GJ et al (2002a) HAR1 mediates systemic regulation of symbiotic organ development. Nature 420:426–429

Nishimura R, Ohmori M, Kawaguchi M (2002b) The novel symbiotic phenotype of enhanced-nodulating mutant of *Lotus japonicus*: astray mutant is an early nodulating mutant with wider nodulation zone. Plant Cell Physiol 43:853–859

Nishimura R, Ohmori M, Fujita H et al (2002c) A Lotus basic leucine zipper protein with a RING-finger motif negatively regulates the developmental program of nodulation. Proc Natl Acad Sci USA 99:15206–15210

Oreo A, Pajuelo P, Pajuelo E et al (2002) Isolation of photorespiratory mutants from *Lotus japonicus* deficient in glutamine synthetase. Physiol Plant 115:352–361

Perry JA, Wang TL, Welham TJ et al (2003) A TILLING reverse genetics tool and a web-accessible collection of mutants of the legume *Lotus japonicus*. Plant Physiol 131:866–871

Perry J, Brachmann A, Welham T et al (2009) TILLING in *Lotus japonicus* identified large allelic series for symbiosis genes and revealed a bias in functionally defective ethyl methanesulfonate alleles toward glycine replacements. Plant Physiol 151:1281–1291

Radutoiu S, Madsen LH, Madsen EB et al (2003) Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases. Nature 425:585–592

Saito K, Yoshikawa M, Yano K et al (2007) NUCLEOPORIN85 is required for calcium spiking, fungal and bacterial symbioses, and seed production in *Lotus japonicus*. Plant Cell 19:610–624

Sandal N, Krusell L, Radutoiu S et al (2002) A genetic linkage map of the model legume *Lotus japonicus* and strategies for fast mapping of new loci. Genetics 161:1673–1683

Sandal N, Petersen TR, Murray J et al (2006) Genetics of symbiosis in *Lotus japonicus*: recombinant inbred lines, comparative genetic maps, and map position of 35 symbiotic loci. Mol Plant Microbe Interact 19:80–91

Sandal N, Jin H, Rodriguez-Navarro DN et al (2012) A set of *Lotus japonicus* gifu x *Lotus burttii* recombinant inbred lines facilitates map-based cloning and QTL mapping. DNA Res 19:317–323

Schauser K, Handberg N, Sandal J et al (1998) Symbiotic mutants deficient in nodule establishment identified after T-DNA transformation of *Lotus japonicus*. Mol Gen Genet 259:414–423

Schauser L, Roussis A, Stiller J et al (1999) A plant regulator controlling development of symbiotic root nodules. Nature 402:191–195

Senoo K, Solaiman MZ, Kawaguchi M et al (2000) Isolation of two different phenotypes of mycorrhizal mutants in the model legume plant *Lotus japonicus* after EMS-treatment. Plant Cell Physiol 41:726–732

Stracke S, Kistner C, Yoshida S et al (2002) A plant receptor-like kinase required for both bacterial and fungal symbiosis. Nature 417:959–962

Suganuma N, Nakamura Y, Yamamoto M et al (2003) The *Lotus japonicus* Sen1 gene controls rhizobial differentiation into nitrogen-fixing bacteroids in nodules. Mol Genet Genomics 269:312–320

Suzaki T, Yano K, Ito M et al (2012) Positive and negative regulation of cortical cell division during root nodule development in *Lotus japonicus* is accompanied by auxin response. Development 139:3997–4006

Suzaki T, Kim CS, Takeda N et al (2013) TRICOT encodes an AMP1-related carboxypeptidase that regulates root nodule development and shoot apical meristem maintenance in *Lotus japonicus*. Development 140:353–361

Suzuki A, Suriyagoda L, Shigeyama T et al (2011) *Lotus japonicus* nodulation is photomorphogenetically controlled by sensing the red/far red (R/FR) ratio through jasmonic acid (JA) signaling. Proc Natl Acad Sci USA 108:16837–16842

Szczyglowski K, Shaw RS, Wopereis J et al (1998) Nodule organogenesis and symbiotic mutants of the model legume *Lotus japonicus*. Mol Plant Microbe Interact 11:684–697

Takahara M, Magori S, Soyano T et al (2013) TOO MUCH LOVE, a Novel Kelch repeat-containing F-box Protein, functions in the long-distance regulation of the legume-rhizobium symbiosis. Plant Cell Physiol 54:433–447

Takeda N, Tsuzuki S, Suzaki T et al (2013) CERBERUS and NSP1 of *Lotus japonicus* are common symbiosis genes that modulate arbuscular mycorrhiza development. Plant Cell Physiol 54:1711–1723

Takos A, Lai D, Mikkelsen L et al (2010) Genetic screening identifies cyanogenesis-deficient mutants of *Lotus japonicus* and reveals enzymatic specificity in hydroxynitrile glucoside metabolism. Plant Cell 22:1605–1619

Takos AM, Knudsen C, Lai D et al (2011) Genomic clustering of cyanogenic glucoside biosynthetic genes aids their identification in *Lotus japonicus* and suggests the repeated evolution of this chemical defence pathway. Plant J 68:273–286

Tansengco ML, Hayashi M, Kawaguchi M et al (2003) Crinkle, a novel symbiotic mutant that affects the infection thread growth and alters the root hair, trichome, and seed development in *Lotus japonicus*. Plant Physiol 131:1054–1063

Tirichine L, Imaizumi-Anraku H, Yoshida S et al (2006a) Deregulation of a Ca^{2+}/calmodulin-dependent kinase leads to spontaneous nodule development. Nature 441:1153–1156

Tirichine L, James EK, Sandal N et al (2006b) Spontaneous root-nodule formation in the model legume *Lotus japonicus*: a novel class of mutants nodulates in the absence of rhizobia. Mol Plant Microbe Interact 19:373–382

Tirichine L, Sandal N, Madsen LH (2007) A gain-of-function mutation in a cytokinin receptor triggers spontaneous root nodule organogenesis. Science 315:104–107

Tominaga A, Nagata M, Futsuki K et al (2009) Enhanced nodulation and nitrogen fixation in the abscisic acid low-sensitive mutant enhanced nitrogen fixation1 of *Lotus japonicus*. Plant Physiol 151:1965–1976

Vriet C, Welham T, Brachmann A et al (2010) A suite of *Lotus japonicus* starch mutants reveals both conserved and novel features of starch metabolism. Plant Physiol 154:643–655

Wang Z, Chen J, Weng L et al (2013) Multiple components are integrated to determine leaf complexity in *Lotus japonicus*. J Integr Plant Biol 55:419–433

Yan J, Cai X, Luo J et al (2010) The REDUCED LEAFLET genes encode key components of the trans-acting small interfering RNA pathway and regulate compound leaf and flower development in *Lotus japonicus*. Plant Physiol 152:797–807

Yano K, Tansengco ML, Hio T et al (2006) New nodulation mutants responsible for infection thread development in *Lotus japonicus*. Mol Plant Microbe Interact 19:801–810

Yano K, Yoshida S, Muller J et al (2008) CYCLOPS, a mediator of symbiotic intracellular accommodation. Proc Natl Acad Sci USA 105:20540–20545

Yano K, Shibata S, Chen WL et al (2009) CERBERUS, a novel U-box protein containing WD-40 repeats, is required for formation of the infection thread and nodule development in the legume–Rhizobium symbiosis. Plant J 60:168–180

Yokota K, Fukai E, Madsen LH et al (2009) Rearrangement of actin cytoskeleton mediates invasion of *Lotus japonicus* roots by Mesorhizobium loti. Plant Cell 21:267–284

Yoshida C, Funayama-Noguchi S, Kawaguchi M (2010) Plenty, a novel hypernodulation mutant in *Lotus japonicus*. Plant Cell Physiol 51:1425–1435

Zhang SL, Sandal N, Polowick PL (2003) Proliferating Floral Organs (Pfo), a *Lotus japonicus* gene required for specifying floral meristem determinacy and organ identity, encodes an F-box protein. Plant J 33:607–619

Forward and Reverse Genetics: The *LORE1* Retrotransposon Insertion Mutants

20

Eigo Fukai, Anna Małolepszy, Niels Sandal, Makoto Hayashi, and Stig U. Andersen

Abstract

The endogenous *Lotus retrotransposon 1* (*LORE1*) transposes in the germ line of *Lotus japonicus* plants that carry an active element. This feature of *LORE1* has been exploited for generation of a large non-transgenic insertion mutant population, where insertions have been annotated using next-generation sequencing approaches. The *LORE1* mutant lines are freely available and can be ordered online. Endogenous retrotransposons are also active in many other plant species. Based on the methods developed for *LORE1* mutagenesis, it should be simple to establish similar systems in other species, once an appropriate element has been identified.

20.1 Introduction

The interaction between legumes and rhizobia that allows symbiotic nitrogen fixation has been extensively studied. Since the 1990s, the molecular characterization of plant symbiotic genes has been accelerated by the use of two model legumes, *Lotus japonicus* and *Medicago truncatula*. Developing genetic tools, such as mutant collections, will help to expand the application of model legumes to a wider range of research areas and will facilitate the identification of legume genes of agronomical importance. Next-generation sequencing has greatly increased the availability of genomic information for many plant species. To facilitate exploitation of this information, ready access to loss-of-function alleles from mutant collections will be essential in a wide range of plant species, including legumes. In this chapter, the establishment of a *Lotus* mutant resource using the endogenous retrotransposon *Lotus retrotransposon 1* (*LORE1*) is described along with the features of this mutant collection. In addition, future possibilities for establishing mutant collections in other legume species are discussed.

A. Małolepszy · N. Sandal · S.U. Andersen (✉)
Centre for Carbohydrate Recognition and Signalling, Department of Molecular Biology and Genetics, Aarhus University, Gustav Wieds Vej 10, 8000 Aarhus C, Denmark
e-mail: sua@mb.au.dk

E. Fukai · M. Hayashi
National Institute of Agrobiological Sciences, Tsukuba 305-8602, Japan

S. Tabata and J. Stougaard (eds.), *The Lotus japonicus Genome*, Compendium of Plant Genomes,
DOI 10.1007/978-3-662-44270-8_20,

20.2 Mutagens and Identification of the Mutations

Forward genetic screening for mutant phenotypes followed by gene cloning using genetic mapping has frequently and successfully been used for identifying genetic components controlling a wide variety of plant traits. Usually, the mutagenized population is single use, meaning that mutations that do not affect the traits of interest will be lost. In the post-genomic era, a large demand for loss-of-function alleles of genes of interest for hypothesis testing quickly became apparent. Mutant collections comprising large mutagenized plant populations complete with information about all induced mutations in the genome of each individual have proved the most efficient way to meet this need, as evidenced by the Arabidopsis T-DNA mutant collection (Alonso et al. 2003).

There are three main combinations of mutagens and procedures for mutation characterization in mutant collections (Table 20.1). The most widely used mutagens are chemicals, such as EMS (ethyl methanesulfonate), that induce nucleotide substitutions and small indels. These mutations can be detected using a technique called targeting induced local lesions in genomes (TILLING), which relies on cleavage by single strand-specific nucleases at sites where mismatches between wild-type and mutant sequences occur (McCallum et al. 2000). At the moment, however, comprehensive characterization of all mutations in an entire mutagenized population has not been undertaken, probably because it would require a very extensive sequencing effort.

High-energy radiation is another common way to mutagenize plants. The first fast neutron mutant population for reverse-genetic use was established by Li et al. (2001). They employed a simple PCR technique to detect deletions in a genomic region of interest through identification of amplicons smaller than those amplified from wild-type alleles. It has been shown that genome-wide array analysis allows detection of deletions induced by fast neutron radiation in soybean (Bolon et al. 2011). However, so far it has not been feasible to identify the majority of the deletions in the population because of the excessive cost and the difficulty in identifying smaller deletions.

Another type of mutagenesis relies on insertion of DNA sequences. The advantage of using DNA insertion for mutagenesis is the simplicity of identifying the induced mutations, provided that the sequence information for the inserted DNA fragments is available. DNA fragment insertion sites can then easily be determined by sequencing flanking region amplicons. The simplicity also makes it possible to comprehensively characterize entire mutagenized populations more efficiently than when using the alternative methods described above. In most cases, T-DNAs or transposable elements (TEs) are used for generating the insertions. Both have been used in Arabidopsis, and in all cases, exogenous DNA fragments were used as mutagens. This makes the resulting mutants transgenic, requiring biocontainment precautions to be taken when handling the plant material. Considering the relatively small size of Arabidopsis plants and its common use in controlled laboratory environments, this is not a major issue. However, in many other plant species, including most crops, the use of transgenic-based insertional mutagenesis is not attractive because of larger plant sizes and recalcitrance to transformation. For these reasons, comprehensive mutant collections in other plant species are scarce, and there remains a great need for such resources in most plant research communities.

Table 20.1 Mutagens and mutation detection methods

Mutagen	Expected types of mutations	Detection method
Chemical (e.g., EMS)	Nucleotide substitutions, indels	TILLING, sequencing
High-energy radiation (e.g., Fast neutron)	Deletions, rearrangements	PCR, DNA array, de-TILLING
Insertion tagging (e.g., TEs, T-DNA)	Insertions	Linker-assisted PCR

20.3 Legume Mutant Collections

From the very beginning of the history of genetics, marked by Mendel's work on garden peas, legumes have been extensively exploited as research tools (Reid and Ross 2011). They include many agronomically important food crops such as soybean, pea, and peanuts as well as pasture crops including alfalfa (*Medicago sativa*), clover (*Trifolium repens* and *Trifolium pratense*), and bird's foot trefoil (*Lotus corniculatus*). Symbiotic nitrogen fixation is a trait of major agronomical importance for legumes, and much legume research has been focused on identifying the genetic components regulating the symbiotic interaction. Reflecting the importance of legumes in research and breeding, several mutant collections are being established in different legume species (Imaizumi et al. 2005; Triques et al. 2007; Tadege et al. 2008; Rogers et al. 2009; Bolon et al. 2011; Hancock et al. 2011; Mathieu et al. 2009; Perry et al. 2003;Pislariu et al. 2012; Cui et al. 2013). One of them is the *LORE1* insertion mutant resource, which was generated in *L. japonicus* using the endogenous retrotransposon *LORE1*. At the time of writing, it was the only non-transgenic legume insertion mutant collection.

20.4 Identification of *LORE1* and Its Germline Transposition

LORE1 elements constitute a family of *Gypsy* retrotransposons endogenous to *Lotus*. They contain a chromodomain at the C-terminal of the integrase, which categorizes them as chromoviruses (Gorinsek et al. 2004). *LORE1* was first identified because of its insertion into genes required for symbiotic nitrogen fixation (Madsen et al. 2005; Schauser et al.1999). The symbiotic mutants were isolated from a gene tagging population established by introducing the exogenous maize *Ac/Ds* DNA transposon into *Lotus* (Thykjaer et al. 1995). In addition to *LORE1*, transpositions of another *Gypsy* retrotransposon *Lotus Retrotransposon 2* (*LORE2*) was also identified in the Ac/Ds population. This suggested that at least two different endogenous retrotransposon families were concurrently active in the *Ac/Ds* population. Their transpositions were also identified in other plant populations regenerated from non-transformed cultured cells (Umehara et al., personal communication; Fukai et al. 2010). The only condition shared by the two activation events was the tissue culture step, suggesting that the activation of *LORE1 a*nd *LORE2* was associated with tissue culture.

Fukai et al. (2010) found that *LORE1a*, one of the *LORE1* family members, can be epigenetically activated in the regenerated intact plants from dedifferentiated cells of the *Lotus* B-129 (Gifu) accession. They also found that activated *LORE1a* transposes in the germ line, mainly in pollen, but not during tissue culture. In agreement with this finding, the promoter of *LORE1a* showed high activity in pollen (Fukai et al. 2010). Variation in DNA methylation patterns in the *LORE1a* promoter region among regenerated plants suggested instability of epigenetic regulation of *LORE1a* during tissue culture (Fukai et al. 2010), indicating that tissue culture processes could induce epigenetic activation of *LORE1a*. *LORE2* transpositions were also observed in plants with active *LORE1* copies (Fukai et al. 2012), suggesting that activation of *LORE2* could be induced in a similar way to *LORE1*. Although the precise characteristics of the transpositional pattern of *LORE2* remains unknown, its transposition frequency was an order of magnitude lower than that seen for *LORE1a* in the mutagenized populations (Urbanski, personal communication).

20.5 Establishment of a *Lotus* Mutant Collection Using the Germline-Specific Retrotransposon *LORE1*

Establishment of mutant collections requires a large number of independent insertions, and the germ line-specific transposition of *LORE1a* greatly facilitates large-scale mutagenesis. Since the number of independent insertions increases in proportion to the number of seeds harvested from a founder plant carrying an active *LORE1a*

element, a single clonally amplified founder plant is sufficient to generate a mutant population of any desired size. Another advantage of *LORE1a* is the absence of untagged phenotypic mutants derived from somaclonal mutations induced during tissue culture. This is often problematic when TE activation in cell culture is used for construction of mutant collections. Since *LORE1* is endogenous to *Lotus*, the *LORE1* mutants are not transgenic, and no biological containment precautions are required, enabling applications such as large-scale mutant screening in open fields.

As a pilot experiment, two groups established medium-scale mutant populations using *LORE1* (Fukai et al. 2012; Urbanski et al. 2012). From the two mutant populations composed of 2,450 and 3,744 plant lines, 4,532 and 8,935 insertions sites were identified, respectively (Fukai et al. 2012; Urbanski et al. 2012). The large-scale identification of insertion sites revealed that *LORE1* has a preference for insertion into genes, limiting the number of insertions required for saturation mutagenesis (Urbanski et al. 2012).

20.6 Identification of *LORE1* Insertions in the Mutant Population by Deep Sequencing

Mutated genes can be identified by the sequencing of flanking regions of new *LORE1* insertions. A sequence-specific amplified polymorphism (SSAP) technique has been used for small-scale experiments of insertion site identification (Madsen et al. 2005; Yokota et al. 2009), and a simplified SSAP protocol has been set up specifically for identifying *LORE1* insertions (Urbanski et al. 2013). The SSAP method, however, is not well suited for high-throughput insertion site identification. Instead, population-wide characterization of insertion sites was carried out by taking advantage of molecular barcoding combined with next-generation sequencing technology using the *FSTpoolit* protocol and software. The laboratory protocol relies on specific amplification of *LORE1* flanking regions from plants pooled in rows and columns of a two-dimensional array. Nested Splinkerette PCR, where molecular barcodes are introduced as part of the PCR primers, is used to ensure highly specific amplification of the *LORE1* flanking sequences (Urbanski et al. 2012, 2013). The two-dimensional setup and barcoding allow pooling of more than 9,000 plants per sequencing library and subsequent automated identification of insertions in all pooled individuals using the *FSTpoolit* software package (Urbanski et al. 2012).

20.7 Line Availability and Current Status of the *LORE1* Mutant Panel

Following the successful pilot experiments, the groups in both Japan and Denmark have continued to expand the *LORE1* populations. At the time of writing, the *LORE1* mutant collection holds more than 80,000 lines with in excess of 340,000 annotated insertions. The number of lines is planned to increase to ~120,000 during 2014/2015, which will result in more than 450,000 annotated insertions, bringing the mutant panel to near-saturation with mutants readily available for the majority of *Lotus* genes. The number of insertions per line ranges from 1 to 15 with an average of four, and all insertions are annotated with information about gene model overlaps and details on the number of reads supporting the insertion call.

All *LORE1* lines, including supporting information, can be browsed and queried using the BLAST tools and genome browser at the Kazusa DNA Research Institute Web site (http://www.kazusa.or.jp/lotus/index.html) or at the Centre for Carbohydrate Signalling and Recognition (CARB) (http://carb.au.dk/lore1/). The Japanese

Table 20.2 Endogenous TEs in legumes with evidence of recent activity

Host	Name of TE	class I/II	Evidence of activity	Activity identified in/as	References
L. japonicus	*LORE1*	I	Transpositions	Regenerated plants from cultured cells	Fukai et al. (2010), Madsen et al. (2005)
	LORE2	I	Transpositions	Regenerated plants from cultured cells	Fukai et al. (2008)
			M. truncatula	*MERE1*	I
			Transpositions	Cultured cells	Rakocevic et al. (2009)
Soybean	*Tgm*	II	Excisions, New mutations	Somatic and germ cells	Xu et al. (2010), Zabala and Vodkin (2005, 2008)
	SORE-1	I	Recent transpositions	Spontaneous mutations during breeding	Kanazawa et al. (2009), Liu et al. (2008)
Peanut	*AhMITE1*	I	Excision	Spontaneous mutation, excision induced by chemical mutagen, Gamma ray, Tissue culture	Gowda et al. (2010, 2011), Patel et al. (2004), Shirasawa et al. (2012)

lines are ordered through Legume Base (http://www.legumebase.brc.miyazaki-u.ac.jp/lore1BrowseAction.do), whereas the Danish lines are ordered from the CARB Web site (http://carb.au.dk/lore1/). As of January 2014, more than 1,800 *LORE1* lines had been dispatched to 19 different countries.

20.8 Future Perspectives

From studying the *LORE1* retrotransposon, we now know that epigenetic activation of TEs can be induced by tissue culture and that their active states can be inherited by the regenerated plants. However, when the TEs lack activity in cultured cells, the activation cannot be detected as transpositions in the first generation of regenerated plants (R0), since germinal transpositions will only be detectable from R1 and later generations. These activation events may occur in tissue culture experiments for any plant species and would usually be overlooked. Looking at the segregation of mutant phenotypes derived from a regenerated plant population, while keeping such a scenario in mind, could lead to the identification of endogenous TEs, which can be activated by tissue culture. The generation of the *LORE1* mutant resource has demonstrated the cost-effectiveness of using endogenous TEs for mutagenesis. A literature search for evidence of recent transpositions or empirical confirmation of transposition activity revealed a number of endogenous legume TEs that could possibly be used for large-scale mutagenesis (Table 20.2). Combined use of model and crop legume mutant collections will strongly facilitate legume molecular biological studies and will contribute to solving agronomical and environmental issues worldwide.

References

Alonso JM et al (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. Science 301(5633):653–657

Bolon YT, Haun WJ, Xu WW, Grant D, Stacey MG, Nelson RT, Gerhardt DJ, Jeddeloh JA, Stacey G, Muehlbauer GJ, Orf JH, Naeve SL, Stupar RM, Vance CP (2011) Phenotypic and genomic analyses of a fast neutron mutant population resource in soybean. Plant Physiol 156:240–253

Cui Y, Barampuram S, Stacey MG, Hancock CN, Findley S, Mathieu M, Zhang Z, Parrott WA, Stacey G (2013) Tnt1 retrotransposon mutagenesis: a tool for soybean functional genomics. Plant Physiol 161:36–47

Fukai E, Dobrowolska AD, Madsen LH, Madsen EB, Umehara Y, Kouchi H, Hirochika H, Stougaard J (2008) Transposition of a 600 thousand-year-old LTR retrotransposon in the model legume *Lotus japonicus*. Plant Mol Biol 68:653–663

Fukai E, Umehara Y, Sato S, Endo M, Kouchi H, Hayashi M, Stougaard J, Hirochika H (2010) Derepression of the plant Chromovirus *LORE1* induces germline transposition in regenerated plants. PLoS Genet 6:e1000868

Fukai E, Soyano T, Umehara Y, Nakayama S, Hirakawa H, Tabata S, Sato S, Hayashi M (2012) Establishment of a *Lotus japonicus* gene tagging population using the exon-targeting endogenous retrotransposon *LORE1*. Plant J 69:720–730

Gorinsek B, Gubensek F, Kordis D (2004) Evolutionary genomics of chromoviruses in eukaryotes. Mol Biol Evol 21:781–798

Gowda MVC, Bhat RS, Motagi BN, Sujay V, Kumari V, Sujatha B (2010) Association of high-frequency origin of late leaf spot resistant mutants with *AhMITE1* transposition in peanut. Plant Breed 129:567–569

Gowda MVC, Bhat RS, Sujay V, Kusuma P, Varshakumari BS, Varshney RK (2011) Characterization of *AhMITE1* transposition and its association with the mutational and evolutionary origin of botanical types in peanut (Arachis spp.). Plant Syst Evol 291:153–158

Hancock CN, Zhang F, Floyd K, Richardson AO, Lafayette P, Tucker D, Wessler SR, Parrott WA (2011) The rice miniature inverted repeat transposable element mPing is an effective insertional mutagen in soybean. Plant Physiol 157:552–562

Imaizumi R, Sato S, Kameya N, Nakamura I, Nakamura Y, Tabata S, Ayabe S, Aoki T (2005) Activation tagging approach in a model legume, *Lotus japonicus*. J Plant Res 118:391–399

Kanazawa A, Liu B, Kong F, Arase S, Abe J (2009) Adaptive evolution involving gene duplication and insertion of a novel *Ty1/copia*-like retrotransposon in soybean. J Mol Evol 69:164–175

Li X, Song Y, Century K, Straight S, Ronald P, Dong X, Lassner M, Zhang Y (2001) A fast neutron deletion mutagenesis-based reverse genetics system for plants. Plant J 27:235–242

Liu B, Kanazawa A, Matsumura H, Takahashi R, Harada K, Abe J (2008) Genetic redundancy in soybean photoresponses associated with duplication of the phytochrome A gene. Genetics 180:995–1007

Madsen LH, Fukai E, Radutoiu S, Yost CK, Sandal N, Schauser L, Stougaard J (2005) *LORE1*, an active low-copy-number TY3-gypsy retrotransposon family in the model legume *Lotus japonicus*. Plant J 44:372–381

Mathieu M, Winters EK, Kong F, Wan J, Wang S, Eckert H, Luth D, Paz M, Donovan C, Zhang Z, Somers D, Wang K, Nguyen H, Shoemaker RC, Stacey G, Clemente T (2009) Establishment of a soybean (*Glycine max* Merr. L) transposon-based mutagenesis repository. Planta 229:279–289

McCallum CM, Comai L, Greene EA, Henikoff S (2000) Targeted screening for induced mutations. Nat Biotechnol 18:455–457

Patel M, Jung S, Moore K, Powell G, Ainsworth C, Abbott A (2004) High-oleate peanut mutants result from a *MITE* insertion into the *FAD2* gene. Theor Appl Genet 108:1492–1502

Perry JA, Wang TL, Welham TJ, Gardner S, Pike JM, Yoshida S, Parniske M (2003) A TILLING reverse genetics tool and a web-accessible collection of mutants of the legume *Lotus japonicus*. Plant Physiol 131:866–871

Pislariu CI, Murray JD, Wen J, Cosson V, Muni RR, Wang M, Benedito VA, Andriankaja A, Cheng X, Jerez IT, Mondy S, Zhang S, Taylor ME, Tadege M, Ratet P, Mysore KS, Chen R, Udvardi MK (2012) A *Medicago truncatula* Tobacco Retrotransposon insertion mutant collection with defects in nodule development and symbiotic nitrogen fixation. Plant Physiol 159:1686–1699

Rakocevic A, Mondy S, Tirichine L, Cosson V, Brocard L, Iantcheva A, Cayrel A, Devier B, Abu El-Heba GA, Ratet P (2009) *MERE1*, a low-copy-number copia-type retroelement in *Medicago truncatula* active during tissue culture. Plant Physiol 151:1250–1263

Reid JB, Ross JJ (2011) Mendel's genes: toward a full molecular characterization. Genetics 189:3–10

Rogers C, Wen J, Chen R, Oldroyd G (2009) Deletion-based reverse genetics in *Medicago truncatula*. Plant Physiol 151:1077–1086

Schauser L, Roussis A, Stiller J, Stougaard J (1999) A plant regulator controlling development of symbiotic root nodules. Nature 402:191–195

Shirasawa K, Hirakawa H, Tabata S, Hasegawa M, Kiyoshima H, Suzuki S, Sasamoto S, Watanabe A, Fujishiro T, Isobe S (2012) Characterization of active miniature inverted-repeat transposable elements in the peanut genome. Theor Appl Genet 124:1429–1438

Tadege M, Wen J, He J, Tu H, Kwak Y, Eschstruth A, Cayrel A, Endre G, Zhao PX, Chabaud M, Ratet P, Mysore KS (2008) Large-scale insertional mutagenesis using the Tnt1 retrotransposon in the model legume *Medicago truncatula*. Plant J 54:335–347

Thykjaer T, Stiller J, Handberg K, Jones J, Stougaard J (1995) The maize transposable element Ac is mobile in the legume *Lotus japonicus*. Plant Mol Biol 27:981–993

Triques K, Sturbois B, Gallais S, Dalmais M, Chauvin S, Clepet C, Aubourg S, Rameau C, Caboche M, Bendahmane A (2007) Characterization of *Arabidopsis thaliana* mismatch specific endonucleases: application to mutation discovery by TILLING in pea. Plant J 51:1116–1125

Urbanski DF, Malolepszy A, Stougaard J, Andersen SU (2012) Genome-wide *LORE1* retrotransposon mutagenesis and high-throughput insertion detection in *Lotus japonicus*. Plant J 69:731–741

Urbanski DF, Malolepszy A, Stougaard J, Andersen SU (2013) High-throughput and targeted genotyping of *Lotus japonicus* LORE1 insertion mutants. In: Rose RJ (ed) Legume genomics: methods and protocols, methods in molecular biology, vol 1069, pp 119–146. doi:10.1007/978-1-62703-613-9_10

Xu M, Brar HK, Grosic S, Palmer RG, Bhattacharyya MK (2010) Excision of an active CACTA-like transposable element from *DFR2* causes variegated flowers in soybean [*Glycine max* (L.) Merr.]. Genetics 184:53–63

Yokota K et al (2009) Rearrangement of actin cytoskeleton mediates invasion of *Lotus japonicus* roots by *Mesorhizobium loti*. Plant Cell 21(1):267–284

Zabala G, Vodkin LO (2005) The wp mutation of *Glycine max* carries a gene-fragment-rich transposon of the CACTA superfamily. Plant Cell 17:2619–2632

Zabala G, Vodkin L (2008) A putative autonomous 20.5 kb-CACTA transposon insertion in an F3'H allele identifies a new CACTA transposon subfamily in *Glycine max*. BMC Plant Biol 8:124

TILLING in *Lotus japonicus*

21

Trevor L. Wang and Fran Robson

Abstract

Following the seminal work on TILLING in *Arabidopsis thaliana*, a population of EMS-mutagenized plants was established for *Lotus japonicus* 'Gifu' to be used for both forward and reverse screening. This was developed into the Lotus TILLING platform which subsequently became *RevGen*UK covering not only TILLING in *L. japonicus*, but also *Medicago truncatula* and other species. Over the last 10 years, nearly two thousand mutations for more than 160 genes have been identified using the platform. In this article, we cover the history, development and current technology employed in the process and explore the impact TILLING has had on Lotus research.

21.1 Introduction

Targeting Induced Local Lesions IN Genomes or TILLING was devised by Claire McCallum and developed for use with Arabidopsis in Seattle, USA, under a collaboration between the groups of Steve Henikoff at the Fred Hutchinson Cancer Research Centre and Luca Comai at the University of Washington (McCallum et al. 2000). It is a reverse genetics method for identifying mismatches in DNA heteroduplexes. Initially, the method relied on denaturing HPLC to identify the mismatches, but it was later modified to permit a higher throughput (Colbert et al. 2001) and establish the Arabidopsis TILLING platform (ATP) in Seattle. This latter method uses an endonuclease, CEL1, from celery and has been adopted almost universally. It remains the method of choice, although others have been used such as high-resolution melting (e.g. Parry et al. 2009). The platform expanded to cover other species and eventually became the Seattle TILLING Project (STP), but ceased operation in 2011, with most of the operation being moved to the Comai lab at UC Davis (http://tilling.ucdavis.edu/index.php/Main_Page) where it continues to date.

TILLING is particularly suited to species with limited genomic information, those that lack other robust reverse genetic methods and especially crops, since it is based on established mutation breeding methods without the use of genetic manipulation. Furthermore, it generates

T.L. Wang (✉) · F. Robson
Metabolic Biology, John Innes Centre, Norwich Research Park, Colney, Norwich NR4 7UH, UK
e-mail: trevor.wang@jic.ac.uk

S. Tabata and J. Stougaard (eds.), *The Lotus japonicus Genome*, Compendium of Plant Genomes,
DOI 10.1007/978-3-662-44270-8_21,

an allelic series of mutants that permits a refined analysis of gene function, in contrast to knockout mutants. The mutagens most often used, because they are effective in most species, are chemical and those that generate single or few nucleotide changes, such as ethylmethane sulphonate (EMS), methyl nitrosourea or sodium azide. The last has often been used with cereals since it has been found to be more effective than the alkylating agents. The effect of the mutagens is random which is important for covering the genome, but is restricted to certain transitions (G/C to A/T for alkylating agents and A/T to G/C for azide) (Olsen et al. 1993; van Arten 1998).

21.2 History of Lotus TILLING

For all the reasons given above, TILLING was deployed to advance *Lotus japonicus* as a model for legume biology and especially the rhizobium–legume symbiosis. At the time the ATP was being established, Martin Parniske at the Sainsbury Laboratory, Norwich, UK, decided to raise a population of *L. japonicus* Gifu EMS-treated plants to isolate nodulation mutants and develop the TILLING method for this species. He was joined by groups working on legumes at the adjacent John Innes Centre, and this collaboration established the Lotus TILLING Platform (LTP), the first TILLING platform outside the USA. The platform started operation in 2003 and now operates under the *RevGen*UK banner (http://revgenuk.jic.ac.uk/). It was originally established for Lotus alone, but now includes *Medicago truncatula*, brassicas and cereals. The unique feature of LTP was that the mutagenized populations included forward screens. This not only helped to identify mutants of interest, but it also provided a measure of how well the population was mutagenized, in terms of mutation density and saturation. The initial platform was based on a Gifu population (Perry et al. 2003, 2009), but has since been supplemented with one in the MG20 background.

21.3 The Mutagenized Populations

The detailed structures of the populations (totalling nearly 9,000 individuals) are presented in Fig. 21.1. Extensive data have already been provided on the Gifu population (Fig. 21.1a) in Perry et al. (2009). Essentially, three forward (thematic) screens were conducted: one for nodulation (nodule morphology, nodulation ability and root morphology; NODPOP), one for general plant morphology (by gross observation; DEVPOP) and one on leaf starch accumulation (by iodine staining of alcohol-cleared leaves at the end of the day and end of the night) (Vriet et al. 2010). DNA was collected from each plant in the screens to use for TILLING since they represent specific mutant-enriched subpopulations. Single plants from each family were also used to collect leaves, extract DNAs, harvest seeds and form a general TILLING population (GENPOP). For the MG20 population (Fig. 21.1a), further forward screens were carried out. Firstly, to add to the nodulation population by screening for nodule production and nodule function (nod^-/fix^-), and to add to GENPOP. In addition, two further screens were completed: one for the ability of leaves to generate cyanide which was highly successful in identifying cyanogenic glucoside mutants (Takos et al. 2010) and an unsuccessful one based on the well-established method of inducing isoflavonoids by applying glutathione (Shelton et al. 2012). In the latter screen, leaves were placed in glutathione-containing induction medium and the fluorescence of the medium measured after 24 h. It was hypothesised that a lack of fluorescence would indicate an inability to produce isoflavonoids and increased fluorescence, an overproduction. Although 61 mutant families showing fluorescence above or below the background level were isolated (Fig. 21.1b), none proved to be mutant on further screening (Takos et al. 2011).

Two further populations were also generated. Seeds from all the plants in a selection of ca. 2000 Gifu families were bulked and retained as a backup population. Several groups have used this

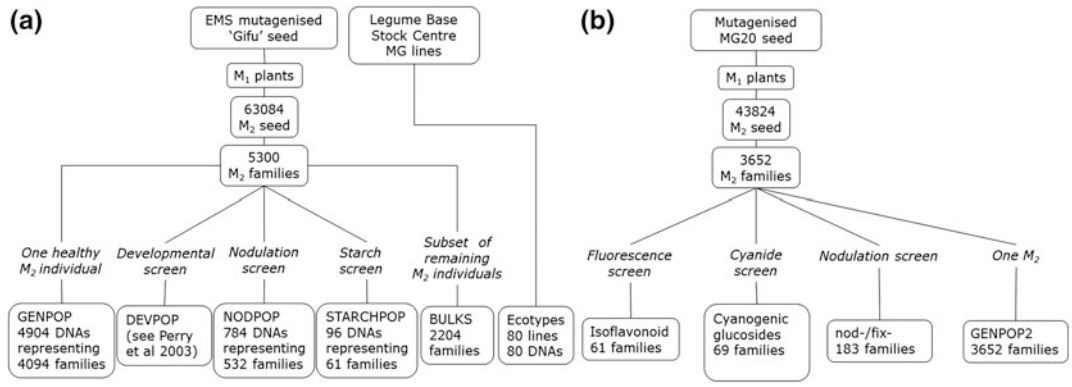


Fig. 21.1 The *Lotus japonicus* plant populations used for forward and reverse genetics. The initial populations were developed in ecotype Gifu (**a**) (Perry et al. 2003, 2009) and new populations generated in MG20 in 2007 (**b**). The type of forward screen is included in italics. All bar one (for isoflavonoids) of these screens was used successfully to isolate mutants (see text for further details)

population to screen for their own particular phenotypes, such as root architecture and, in particular, spontaneous nodulation (see Sect. 21.5). A collection of 80 Lotus ecotypes was also assembled for use as an ecoTILLING population (Fig. 21.1a).

A summary of the types of mutation found by TILLING across the Lotus populations is shown in Table 21.1. Not all mutations can generate a phenotype. In our populations, 43 % of mutations were in non-coding regions or were silent, with an average of seven mutations identified for each sequence targeted in the Gifu population and 11 in the MG20 population, although the latter sample is much smaller. Of note are the 4.9 % of mutations causing gene knockouts by creating premature stop codons or missplicing at splice–site junctions. Interestingly, there is a difference between the two populations in the ratio of mutations existing in heterozygous form to those in homozygous ones. In initial observations of the Gifu population, the heterozygote-to-homozygotes ratio was found to be 10:1, indicating three germline cells contributed to male and female gametes (Perry et al. 2009 Supplemental Fig. S1). If a single mutagenized cell was carried through into the germline, one would anticipate a 2:1 ratio as observed in the original Arabidopsis population (Greene et al. 2003) and seen here in the MG20 population (although in a much smaller sample than for the Gifu population). Current data accumulated for the Gifu population indicate a ratio of 5.5:1. This is more in line with two germline cells in each flower contributing to gametes. The reason for the difference between the two ecotypes is not obvious as one would expect germline genetics of the two to be the same. It may indicate some kind of selection in the generation of the original Gifu population favouring heterozygotes.

Table 21.1 Summary of mutations found by TILLING in *Lotus japonicus*

	No. of genes TILLed	No. of fragments TILLed	Mutation type						Average no. per TILLING fragment	Het: Hom ratio
			Missense	Premature Stop	Splice Junction	Non-coding	Silent (no AA change)	Total		
Gifu	150	235	895	64	18	412	337	1726	7.4	5.5:1
MG20	20	20	140	13	1	23	46	223	11.2	2.1:1
Total	170[a]	255	1,035	77	19	435	383	1,949		

Data were taken from Table 21.3, but includes embargoed genes. The ratio of the number of M_2 plants bearing mutations in heterozygous form to homozygous ones is given in the final column. *Note* the difference between the two populations in this ratio (see text for further comments). Note also the mutations causing premature stop codons and splicing errors amount to 4.9 %, equivalent to the theoretical value (McCallum et al. 2000)

[a] Some of the total of 163 genes have been TILLed in both populations

21.4 The Process

TILLING can be carried out using a range of different technologies, but the central principle is the same: the detection of rare mutations in pooled DNA samples from large mutant populations in a first round of screening, followed by deconvolution of the pool to identify the individual carrying the mutation. The majority of current TILLING protocols utilise two different and sequential techniques—one to screen the pools and another to identify the mutant individual. The definitive method for initial screening uses single-strand-specific endonucleases that cleave DNA at the site of mismatched nucleotides in otherwise double-stranded DNA (Colbert et al. 2001). The mismatches form a tiny single-stranded 'bulge' that is recognised and nicked by these enzymes, and the resulting cleaved and fluorescently labelled PCR fragments are separated and visualised. The process used for *L. japonicus* has been described previously in detail for the initial instrument employed to separate fragments, an ABI 377 DNA analyzer (Perry et al. 2003, 2005), and also for the LI-COR® 4300 DNA analyzer (Perry et al. 2009). Both instruments use denaturing polyacrylamide gel electrophoresis for fragment separation, and presently, most laboratories use this technique for TILLING.

A method that replaces the gel-based system with fragment separation by capillary sequencing (Le Signor et al. 2009) was developed by the John Innes Genome Laboratory under the EU GLIP project (Grain Legumes Integrated Project, http://www.pcgin.org/GLIP/pubrep.pdf). This technique was adopted by *RevGen*UK and is currently in use for all its populations. In this method, heteroduplexed PCR products from pooled DNA samples are cleaved using purified celery juice extract (CJE) and the resulting labelled fragments separated on an Applied Biosystems™ 3730xl DNA sequencer. The *RevGen*UK TILLING pipeline utilises a number of resources to assist with the design and implementation of the best possible TILL for each gene. To be effective, the genomic and coding or protein sequence of the gene of interest needs to be known so that it can be annotated correctly with all intron/exon boundaries marked. To identify regions of the gene that are likely to give the highest concentration of deleterious mutations generated by EMS-induced G/C to A/T transitions, the Web software CODDLE (Codons Optimised to Discover Deleterious LEsions; http://www.proweb.org/coddle) simplifies this selection procedure (Colbert et al. 2001). The program automatically searches homology databases using either suggested or user-defined homology blocks (the latter by means of the program Sorts Intolerant From Tolerant (SIFT) to predict whether a substitution affects protein function) (Kumar et al. 2009) and identifies conserved regions important for protein function. Taking this and the amino acid composition of the protein into account, CODDLE identifies a region suitable for TILLING. The CODDLE program can also design primer pairs following standard quality criteria based on Primer3 (Untergrasser et al. 2012) and in addition considers homology between similar genes. However, where genes are interrupted by numerous and/or large introns, CODDLE can recommend a fragment unsuitable for our TILLING protocol because it attempts to encompass more than one exon and hence may contain an unacceptably large proportion of non-coding sequence. In addition, long tracts of homopolymers are not excluded by CODDLE. These are best avoided because *Taq* polymerase slippage can introduce mismatches in such regions. Consequently, in some circumstances, CODDLE is used for guidance only.

As part of our quality controls before TILLING a gene, we generally test two primer pairs for each fragment for their amplification efficiency and fidelity. We select one pair and then retest with fluorescently labelled versions using unlabelled/labelled at a ratio of 3:2 to replicate our TILL PCR conditions. Our current DNA populations are mostly pooled eightfold; hence, each TILL starts by using four to six eightfold pooled 96-well plates. Products are tested to determine how much of the PCR to use in the subsequent CJE digest. After the digest step and subsequent clean-ups, the samples are run on the

sequencer (Le Signor et al. 2009). Potential mutations are visualised as anomalous 'empty' peaks on overlaid chromatograms using the GeneMapper® software. This identifies a pool containing DNA from a number of plants. The plant carrying the mutation is pinpointed after conventional PCR and sequencing of the individual DNA samples that made up the relevant pool followed by analysis using the software MutationSurveyor® which can analyse multiple sequence trace files simultaneously. Seed of the plant is then sown, and progeny genotyped and phenotyped to check or select as necessary homozygous mutants.

21.5 Global Reach and Impact

While it was under development, Lotus TILLING was available to relatively few researchers within a collaborative programme of research on the species. When this programme ended however, and the service was open to all, it rapidly became globalised. To date, 32 Lotus groups from 11 countries have used the service (Table 21.2). The depth and spread of TILLING across the globe demonstrates that it was universally adopted as the method of choice for reverse genetics in *L. japonicus*. To date, 163 genes have been targeted, some in both ecotypes. These are listed in Table 21.3 together with the types of mutation discovered. Due to the random nature of the chemical mutagenesis, TILLING can be used to target any sequence with the expectation that mutations will be identified and the table shows that this has been the case to date. TILLING has thus had a significant impact on numerous aspects of legume research, but especially the legume–rhizobium symbiosis (Perry et al. 2009); it was used very early on to establish such genes as *CASTOR* and *POLLUX* (Imaizumi-Anraku et al. 2005) as part of the symbiosis signalling pathway and to obtain numerous alleles at symbiotic loci (Perry et al. 2009). The random nature of the mutagen ensures that mutations can be found in most genes whether in nodule signal transduction, as above, nodule metabolism (e.g. Horst et al. 2007), plant development (Welham et al. 2009) or indeed many processes for which it would be very difficult to develop forward screens. Although knockouts, through the generation of premature stop codons or incorrect splicing, are rare—about 5 % as calculated by McCallum et al. (2000) and as found in practice—potentially useful mutations are found at a much higher frequency in most cases. TILLING has been used, therefore, not only to identify gene function, but also to confirm a particular phenotype (e.g. Yano et al. 2008; Krusell et al. 2011) where a knockout may not be required.

The impact of the mutations discovered from our populations has not been confined to TILLING, however. Researchers have used the mutagenized seed for a number of forward screens, from simple visual screens (e.g. flower development; Dong et al. 2005), through microscopic screens (e.g. Groth et al. 2013), to chemical screens for secondary metabolism (e.g. Takos et al. 2010). Most notable of these, however, was the heroic effort made to isolate spontaneously nodulating plants by Tirichine and co-workers (Tirichine et al. 2007) using material from the bulked populations of Gifu (Fig. 21.1a).

Table 21.2 Research groups by country that have accessed TILLING in *Lotus japonicus*

	No. of Research groups	No. of Genes TILLed
Canada	1	5
Denmark	2	40
France	2	3
Germany	4	43
Greece	1	2
Italy	1	4
Japan	9	27
Spain	1	1
Sweden	1	3
UK	8	34
USA	2	2
Total	*32*	*164*[a]

[a] A total of 163 genes has been TILLed, one twice

Table 21.3 Mutations identified in *Lotus japonicus* genes targeted by TILLING

Description	Gene symbol	Predicted gene location	Release 3.0 predicted gene	Eco-type	pre- screen or Reverse	Mis-sense	Premature stop	Splice junction	non-coding	silent	Total mutations per gene	Genbank Protein ID.	Reference for gene	Reference for TILLING mutants
Adenylate Isopentenyltransferase	IPT3	chr5.CM1125.340.r2.m	Lj5g0962690.1	G	R	11	0	0	0	5	16	ABD93934.1		
adenylate Isopentenyltransferase	IPT4	chr6.CM0139.790.r2.d	Lj0g0154359.1	G	R	7	0	0	0	6	13	ABD93935.1		
ADP glucose pyrophosphorylase LSU	APL2	chr1.CM0952.140.r2.m	Lj1g2126130.1	G	F/R	3	0	1	8	4	16	n/e		Vriet et al. 2010
ADP glucose Pyrophosphorylase SSU	APS1	chr2.CM0191.60.r2.m	Lj2g3337310.1	G	R	9	0	0	5	4	18	n/e		Vriet et al. 2010
Allene oxide cyclase1		chr2.CM0903.700.r2.m	Lj2g1168900.1	G	R	11	0	0	4	1	16	n/e		
Allene oxide cyclase2		chr1.CM0012.1230.r2.m	Lj1g4590880.1	G	R	3	0	0	5	5	13	n/e		
Allene oxide synthase		chr5.CM0089.580.r2.d	Lj5g0080890.1	G	R	2	0	0	0	2	4	n/e		
AM-induced MYB transcription factor	MAMI	chr1.LjB18K24.70.r2.a	n/a (pseudo)	G	R	9	0	0	3	0	12	n/e	Volpe et al. 2013	
Ankyrin repeat domain-containing protein		LjSGA_008026.1	Lj0g0049599.1	G	F	4	0	0	0	0	4	n/e		
AP2/ERF transcription factor		chr1.CM0104.2670.r2.m	Lj1g3975310.1	G	R	12	0	0	4	4	20	n/e		
AP2-EREBP transcription factor	ERF1	chr5.CM0341.90.r2.m	Lj5g1167370.1	G	F/R	14	0	0	1	2	17	BAG50053.1	Asamizu et al. 2008	
AP2-EREBP transcription factor	ERF17	chr2.LjT11L13.40.r2.m	Lj2g0227060.1	G	F	1	0	0	0	0	1	BAG50064.1	Asamizu et al. 2008	
AP2-EREBP transcription factor		chr6.CM0013.730.r2.d	Lj6g1175110.1/ L|6g1175120.1+	G	R	11	0	0	0	1	12	n/e		
Arabinogalactan protein		chr1.CM0104.2620.r2.m	n/a	G	F	1	0	0	0	0	1	n/e		
Ascorbate peroxidase		chr3.CM0616.30.r2.d	n/a	G	F	0	0	0	2	1	3	n/e		
Asparaginase	NSE1	chr5.CM0096.20.r2.m	Lj5g0296030.1	G	F/R	4	0	0	5	1	10	n/e	Credali et al. 2013	
Aspartyl protease family		chr1.CM0982.580.r2.d	Lj1g1584680.1	G	R	12	0	0	0	8	20	n/e		
ATP Binding protein	ABCB1	chr1.CM0122.1360.r2.m	Lj1g4830300.1	G	R	14	0	0	2	0	16	n/e	Takanashi et al. 2012	
ATP binding protein		chr1.LjT14I13.150.r2.a	Lj1g3137890.1	G	R	6	0	0	0	4	10	n/e		
Beta-Glucosidase2	CYD2/ BGD2	chr3.LjT33P07.150.r2.m	Lj0g0362549.1	G	R	0	0	0	7	1	8	ACD65510.1		Takos et al. 2010
Beta-Glucosidase4	BGD4	chr5.LjT34K16.50.r2.m	Lj5g0391410.1	G/M	R	8	0	0	9	2	19	ACD65509.2		Takos et al. 2010
bHLH DNA-binding protein		chr1.CM0147.640.r2.d	Lj0g0157999.1	G	R	6	0	0	1	2	9	n/e		
Blade-on-petiole BTB/POZ-Ankyrin domain protein		chr1.CM0009.300.r2.m	Lj1g4155980.1	M	R	12	0	0	0	3	15	n/e		
BOI-RELATED GENE 1, RING finger domain protein		chr6.CM0066.230.r2.a	Lj6g1692670.1	G	R	8	0	0	1	3	12	n/e		

(continued)

Table 21.3 (continued)

Description	Gene symbol	Predicted gene location	Release 3.0 predicted gene	Eco-type	pre- screen or Reverse	Mis-sense	Premature stop	Splice junction	non-coding	silent	Total mutations per gene	Genbank Protein ID.	Reference for gene	Reference for TILLING mutants
bZIP transcription factor	bZIP-R91	chr1.LjT35C22.60.r2.m	Lj1g2035100.2	G	F	0	0	0	1	0	1	BAG50056.1	Asamizu et al. 2008	
bZIP transcription factor	bZIP-M43	chr5.CM0909.500.r2.m	Lj5g1697630.1	G	F	2	0	0	0	0	2	BAG50070.1	Asamizu et al. 2008	
bZIP transcriptional repressor		chr4.CM0165.530.r2.d	Lj4g0934560.4	G	R	17	1	0	8	4	30	n/e		
Ca2+ and calmodulin-dependent protein kinase	CCaMK	chr3.LjT02O17.60.r2.m	Lj3g1739280.1	G	F	0	1	0	4	0	5	CAJ76700.1		Perry et al. 2009
Carbonic anhydrase	CAA1	chr1.CM0398.590.r2.a	Lj1g4226880.1	G	R	3	0	1	6	0	10	CAM59682.1	Tsikou et al. 2011	
Carbonic anhydrase	CAA2	chr5.CM2155.40.r2.a	Lj5g0780660.1	G	R	1	0	0	2	0	3	CAM59683.1	Tsikou et al. 2011	
CCAAT-binding transcription factor	CBF-A22	chr1.CM0029.540.r2.m	Lj1g4752710.1	G	F	1	0	0	0	0	1	BAG50055.1	Asamizu et al. 2008	
Ceramide glucosyltransferase		chr4.CM0617.430.r2.d	Lj4g3044960.1	G	R	6	0	0	10	0	16	n/e		
Clavata3/ESR-related CLE		CM0446.165.r2.a	Lj0g0000559.1	G	R	6	0	0	4	2	12	n/e		
Clavata3/ESR-related CLE		chr3.CM2103.25.r2.a	n/a	G	R	6	0	0	11	4	21	n/e		
Coiled-coiled domain protein	CYCLOPS	chr2.CM0803.150.r2.m	Lj2g 1549600.1	G	F	0	2	0	1	0	3	ABU63668.1	Yano et al. 2008	
Cryptochrome 2		chr6.CM1613.50.r2.m	Lj6g0029980.1	G	R	15	5	0	6	5	31	n/e		
Cyclic nucleotide-gated ion channel protein		chr2.CM0435.1090.r2.d	Lj2g0632220.2	G	R	1	1	0	1	1	4	n/e	Maekawa-Yoshikawa et al. 2009	
Cystein-rich polycomb-like protein	CPP1	chr5.CM0345.1440.r2.m	Lj5g0539820.1	G	F	1	0	0	0	1	2	CAF02298.1	Andersen et al. 2003	
Cytochrome P450		chr1.CM0017.80.r2.m	Lj1g3317470.1	G	R	5	1	0	2	3	11	n/e		
Cytochrome P450	CYP71D353	chr3.CM0292.110.r2.m	Lj3g1983600.1	G	R	1	0	0	0	0	1	n/e	Krokida et al. 2013	
Cytochrome P450	CYP88D5	chr3.CM0292.120.r2.m	Lj3g1983620.1	G	R	3	0	0	0	1	4	BAG68928.1	Krokida et al. 2013	
Cytochrome P450	CYP79D3	chr3.CM0241.700.r2.m	Lj3g0755150.1	G	R	2	0	0	0	1	3	AAT11920.1	Takos et al. 2011	
Cytochrome P450	CYP79D4	chr3.CM0241.310.r2.m	Lj3g0744720.1	G	R	6	1	0	0	0	7	AAT11921.1	Morant et al. 2008	
Cytochrome P450		chr6.CM1091.90.r2.m	Lj6g0898690.2	G	R	1	0	0	0	1	2	n/e		
Dehydration-responsive element binding protein		chr5.CM0359.290.r2.m	Lj5g0712260.1	G	R	4	0	0	0	2	6	n/e		
Dicer-like 1		chr1.CM0105.1760.r2.m	Lj1g5061000.1	G	R	2	0	0	1	2	5	n/e	Bustos-Sanmamed et al. 2013	

(continued)

Table 21.3 (continued)

Description	Gene symbol	Predicted gene location	Release 3.0 predicted gene	Eco-type	pre- screen or Reverse	Mis-sense	Premature stop	Splice junction	non-coding	silent	Total mutations per gene	Genbank Protein ID.	Reference for gene	Reference for TILLING mutants
Dicer-like 2		chr6.CM0437.480.r2.m	Lj6g1629810.1	G	R	3	0	1	0	1	5	n/e	Bustos-Sanmamed et al. 2013	
Digalactosyldiacylglycerol synthase 1	DGD1	chr1.CM0029.370.r2.d	Lj1g4741430.1	G	R	7	1	0	0	1	9	AAT67422.1	Gaude et al. 2004	
Digalactosyldiacylglycerol synthase 2	DGD2	chr1.CM0284.340.r2.d	Lj1g2627380.1	G	R	7	0	0	3	2	12	AAT67423.1	Gaude et al. 2004	
DNA polymerase epsilon, catalytic subunit		chr3.CM1144.290.r2.m	Lj3g2995970.1	G	R	3	0	0	3	4	10	n/e		
E3 Ubiquitin ligase	SINA1	chr1.CM0105.860.r2.m	Lj1g5034850.1	G	R	3	0	0	7	1	11	CCG06551.1	Den Herder etal. 2012	
E3 Ubiquitin ligase-like protein	SINA2	chr1.LjT05B18.80.r2.d	Lj1g4372170.1	G	R	3	0	0	3	1	7	CCG06552.1	Den Herder et al. 2012	
ENOD promoter DNA binding protein, jumonji domain		chr1.LjT26E16.120.r2.m	Lj1g0725870.1	G	F	1	0	0	1	2	4	n/e		
Ethylene receptor		chr3.CM0634.490.r2.m	Lj3g0461750.1	G	R	25	1	1	5	4	36	n/e		
Ethylene response sensor		chr1.CM0122.2470.r2.m	Lj1g4865380.1	G	R	26	0	0	5	8	39	n/e		
FAD linked oxidase		chr6.CM0778.220.r2.d	Lj5g0692300.1	G	R	7	1	0	5	1	14	n/e		
Glucan water dikinase 3		chr5.LjT42F22.160.r2.m	Lj5g2302970.1	G	R	19	0	0	0	7	26	n/e		Vriet et al. 2010
GRAS family protein	NSP1	chr3.CM0416.1260.r2.d	Lj3g2579340.1	G	F	2	0	0	0	1	3	ABK35066.1		Perry etal. 2009
GRAS family protein	NSP2	chr1.CM 1976.90.r2.m	Lj1g0785930.1	G	F	3	1	0	0	1	5	ABG49438.1		Perry etal. 2009
GRAS family transcription factor		LjT15C06.70.r2.m	Lj0g0017249.1	G/M	R	27	3	0	0	10	40	n/e		
High affinity ammonium transporter	AMT2	LjSGA_014720.1	Lj0g0115479.	G	R	3	4	0	0	1	8	ACQ91094.1	Guether etal. 2009	
His-Asp phosphotransmitter		chr6.CM0139.1730.r2.m	Lj6gl915980.2	G	F/R	5	0	0	3	0	8	n/e		
Histidine kinase	HK2	chr2.LjBHM03.80.r2.m	Lj2g0136400.1	G	R	5	1	0	2	2	10	ABI48270.1	Murray et al. 2007	
Histidine kinase		chr2.CM0191.270.r2.a	Lj2g3338860.2	G	R	10	0	0	2	5	17	n/e		
Histone deacetylase		chr4.CM0680.20.r2.m	Lj4g2742710.1	M	R	6	0	0	1	2	9	n/e		
Ion channel protein	CASTOR	chr1.CM0105.1940.r2.m	Lj1g5061360.1	G	F	10	1	1	3	0	15	BAD89021.1		Perry etal. 2009
Ion channel protein	POLLUX	chr6.CM0508.260.r2.m	Lj6g2275010.2	G	F	11	3	2	5	2	23	BAD89022.1		Perry etal. 2009
Lipopolysaccharide Binding protein		chr2.CM0667.100.r2.a	Lj2g3224210.1	G	R	2	2	0	1	1	6	n/e		

(continued)

Table 21.3 (continued)

Description	Gene symbol	Predicted gene location	Release 3.0 predicted gene	Eco-type	pre- screen or Reverse	Mis-sense	Premature stop	Splice junction	non-coding	silent	Total mutations per gene	Genbank Protein ID.	Reference for gene	Reference for TILLING mutants
Lipopolysaccharide Binding protein			chr6.CM1514.340.r2.m		Lj6g0819000.1	G	R	12	0	0	3	1	16	n/e
Lipopolysaccharide Binding protein		chr2.CM1285.90.r2.a	Lj2g2856000.2	G	R	4	2	0	6	1	13	n/e		
LysM type receptor kinase	Lys11	chr4.CM0165.270.r2.d	Lj4g0912430.1	G/M	R	13	4	0	0	8	25	BAI79285.1	Lohmann et al. 2010	
LysM type receptor kinase	LYS16	chr1.CM0064.880.r2.m	Lj1g3834250.1	G	R	6	0	0	0	0	6	BAI79280.1	Lohmann et al. 2010	
LysM type receptor kinase	LYS5	chr3.CM0213.610.r2.m	Lj3g2318170.1	G	R	4	0	0	3	2	9	BAI79272.1	Lohmann et al. 2010	
LysM type receptor kinase	LYS15	chr3.CM0649.210.r2.d	Lj3g3082380.1	G	R	12	2	0	0	1	15	BAI79290.1	Lohmann et al. 2010	
LysM type receptor kinase	LYS6	chr6.CM0041.460.r2.a	Lj6g1055580.1	G	R	10	0	0	0	2	12	BAI79273.1	Lohmann et al. 2010	
LysM type receptor kinase	LYS7	chr6.CM0367.820.r2.m	Lj6g1812110.1	G	R	5	0	1	5	2	13	BAI79274.1	Lohmann et al. 2010	
LysM type receptor kinase	LYS3	chr2.CM0008.630.r2.m	Lj2g1415410.1	G	R	10	1	1	9	5	26	BAI79284.1	Lohmann et al. 2010	
LysM type receptor kinase	LYS20	chr1.CM1413.200.r2.a	Lj1g2808030.1	G	R	8	0	0	0	6	14	BAI79288.1	Lohmann et al. 2010	
LysM type receptor kinase	LYS13	chr2.CM0826.330.r2.m	Lj2g2899910.1	G	R	8	0	0	0	2	10	BAI79287.1	Lohmann et al. 2010	
LysM type receptor kinase	LYS14	chr2.CM0826.310.r2.m	Lj2g2899900.1	G	R	10	1	0	0	1	12	BAI79278.1	Lohmann et al. 2010	
LysM type receptor kinase	LYS4	chr3.CM1488.630.r2.m	Lj3g0290100.1	G	R	5	0	0	4	1	10	BAI79270.1	Lohmann et al. 2010	
LysM type receptor kinase	LYS1	chr2.CM0545.230.r2.m	Lj2g2904610.1	G	R	4	0	0	14	2	20	BAI79267.1	Lohmann et al. 2010	
LysM type receptor kinase	LYS2	chr2.CM0545.240.r2.m	Lj2g2904710.1	G	R	2	0	1	7	3	13	BAI79283.1	Lohmann et al. 2010	
LysM type receptor kinase	NFR1a	chr2.CM0545.250.r2.m	Lj2g2904690.1	G	F	3	1	0	2	1	7	CAE02589.1		
LysM type receptor kinase	NFR5	chr2.CM0323.400.r2.d	Lj2g1828350.1	G	F/R	5	3	0	0	2	10	CAE02597.1		Perry et al. 2009
Malate dehydrogenase		chr4.CM0087.120.r2.m	Lj0g0169339.1	G/M	F/R	28	0	0	0	10	38	n/e		
Multi-functional beta-amyrin synthase	AMY2	chr3.CM0292.280.r2.d	Lj3g1983430.1	G	R	2	0	0	1	0	3	AA033580.1	Krokida et al. 2013	
MYB-like DNA binding protein		chr1.CM0295.1040.r2.m	Lj1g4515910.1	G/M	R	12	3	1	1	5	22	n/e		
NDX homeobox gene	NDX2	chr3.LjT09L18.40.r2.m/ chr3.LjT09L18.60.r2.m	Lj3g1652340.1	G	F	2	0	1	0	1	4	CAA09792.1	Gronlund et al. 2003	

(continued)

Table 21.3 (continued)

Description	Gene symbol	Predicted gene location	Release 3.0 predicted gene	Eco-type	pre- screen or Reverse	Mis-sense	Premature stop	Splice junction	non-coding	silent	Total mutations per gene	Genbank Protein ID.	Reference for gene	Reference for TILLING mutants
Neutral/alkaline invertase	INV1	chr5.CM1667.130.r2.a	Lj5g1853130.1	G	F/R	7	1	0	2	0	10	CAG30577.1		Welham et al. 2009
Neutral/alkaline invertase	INV2	chr5.CM1598.160.r2.m	Lj5g1530080.1	G	F/R	3	1	0	3	5	12	n/e		Welham et al. 2009
Nicotianamine synthase	NAS1	chr1.CM0122.180.r2.m	Lj1g4790630.1	G	R	6	3	0	0	4	13	BAH22562.1	Hakoyama et al. 2009	
Nicotianamine synthase	NAS2	chr6.CM0139.1430.r2.d	Lj6g1914550.1	G	R	4	0	0	1	7	12	BAH22563.1	Hakoyama et al. 2009	
NIN-like protein		chr3.CM0091.230.r2.m	Lj3g3336070.1	G	F	0	0	0	0	1	1	n/e		
Nod factor binding lectin-nucleotide phosphohydrolase	LNP	chr1.CM0104.1920.r2.m	Lj1g3948070.1	G	F/R	20	0	0	6	2	28	AAF00609.1	Roberts et al. 1999	
Nod factor binding lectin-nucleotide phosphohydrolase	LNP2	chr1.CM0104.1830.r2.d	Lj1g3945990.1	G	R	6	0	1	5	3	15	n/e	Roberts et al. 2013	
Nonsymbiotic hemoglobin	HB1	chr3.CM0091.620.r2.m	Lj3g3338170.1	G	R	6	0	0	9	2	17	BAE46739.1	Nagata et al. 2008	
Nucleoporin	NUP133	chr2.CM0191.150.r2.m	Lj2g3337540.1	G	F	1	2	0	1	4	8	CAI64811.1	Kanamori et al. 2006	
Nucleoporin	NUP85	chr1.CM0171.120.r2.m	Lj1g0318200.1	G	F	2	0	0	1	0	3	BAF45348.1		Perry et al. 2009
Oligopeptide transporter		chr1.CM0125.390.r2.m	Lj1g1183390.1	G	F/R	4	0	0	2	1	7	AAB69642.1		
Oligopeptide transporter		chr1.CM0295.980.r2.m	Lj1g4515810.2	G	R	2	0	0	0	0	2	n/e		
Pathogenesis-related protein		LjSGA_029184.1	n/a	G	R	4	0	0	1	3	8	n/e		
Pathogenesis-related protein		LjSGA_134572.1	Lj6g2170750.1	G	R	9	0	0	0	5	14	n/e		
Pentatricopeptide repeat-containing protein		chr1.CM0012.540.r2.a	Lj1g4578920.1	G	R	5	0	0	4	1	10	n/e		
Phosphatidylinositol transferlike protein IV	PLP-IV	chr2.CM0249.900.r2.m	n/a	G	F/R	5	0	0	3	5	13	AAK63248.1	Kapranov et al. 2001	
Phosphoglucomutase	PGM1	chr5.CM0200.200.r2.m	Lj5g2029660.1	G	F/R	3	2	1	4	0	10	n/e		Vriet et al. 2010
Phospholipase D	PLD	chr3.CM0142.570.r2.d	Lj0g0116389.1	G	R	8	0	0	0	0	8	n/e	Serna-Sanz et al. 2011	
PITH domain-containing protein		chr4.CM0307.240.r2.d	Lj4g2821280.1	M	R	5	0	0	5	0	10	n/e		
R2R3-MYB transcription factor	MYB15	chr6.CM1613.30.r2.m	Lj6g0029920.1	G	R	0	0	0	1	0	1	n/e	Shelton et al. 2012	
R2R3-MYB transcription factor	MYB3	chr3.LjT15F17.70.r2.d	Lj3g3465520.1	G	R	3	0	0	0	0	3	n/e	Shelton et al. 2012	
R2R3-MYB transcription factor	MYB111	chr6.CM0385.170.r2.d	Lj0g0128909.1	G	R	3	0	0	0	0	3	n/e	Shelton et al. 2012	
R2R3-MYB transcription factor	MYB17	chr5.CM0148.590.r2.m	Lj5g2013860.1	G	R	0	0	0	9	0	9	n/e	Shelton et al. 2012	

(continued)

Table 21.3 (continued)

Description	Gene symbol	Predicted gene location	Release 3.0 predicted gene	Eco-type	pre- screen or Reverse	Mis-sense	Premature stop	Splice junction	non-coding	silent	Total mutations per gene	Genbank Protein ID.	Reference for gene	Reference for TILLING mutants
R2R3-MYB transcription factor	MYB16	chr1.CM0141.320. r2.m	Lj1g2979210.1	G	R	1	0	1	7	1	10	n/e	Shelton et al. 2012	
R2R3-MYB transcription factor	MYB13	chr1.CM0295.180. r2.m	Lj1g4483810.1	G	R	5	0	0	4	2	11	n/e	Shelton et al. 2012	
R2R3-MYB transcription factor	TT2a/ MYB136	chr6.CM0013.1540.r2.a	Lj6g1201340.3	G	R	8	0	0	3	3	14	BAG12893.1	Yoshida et al. 2008	
R2R3-MYB transcription factor	TT2a/ MYB137	chr6.CM0013.1500.r2.m	Lj6g1201220.1	G	R	2	0	0	1	6	9	BAG12894.2	Yoshida et al. 2008	
R2R3-MYB transcription factor	MYB4	n/e	Lj1g3342760.1	G	R	2	0	0	0	1	3	n/e	Shelton et al. 2012	
R2R3-MYB transcription factor		chr1.CM0104.3630.r2.m	Lj1g4012910.1	G	R	2	0	0	0	1	3	n/e		
Receptor-like kinase	CLV3	chr3.LjT37K03.60.r2.m	Lj3g1239970.1	G	R	7	0	0	0	7	14	n/e		
Receptor-like kinase	SYM RK	chr2.CM0177.340.r2.m	Lj2g1467920.1	G	F/R	26	2	1	17	9	55	AAM 67418.1		Perry etal. 2009
Remorin protein		chr2.CM0021.3290.r2.m	Lj2g2017500.1	G	R	0	0	0	1	0	1	n/e		
Remorin protein		chr4.CM0004.60.r2.d	Lj4g2928720.1	G	R	3	1	0	3	2	9	AFK48724.1		
Remorin protein		chr2.CM0021.3290.r2.m	Lj2g2017500.1	G	R	5	1	1	4	2	13	n/e		
RHO protein GDP dissociation inhibitor	Rho GDI	chr4.CM0288.670.r2.m	Lj4g0120120.2	G	R	3	0	0	6	0	9	AFK42510.1		
RWP-RK family protein	NIN	chr2.CM0102.250.r2.m	Lj2g3373110.1	G	F	5	0	0	0	1	6	CAB61243.1		Perry etal. 2009
Sterol 3-beta-glycosyl transferase		LjSGA_049701.1	Lj0g0206149.2	G	R	0	1	0	7	2	10	n/e		
Sterol 3-beta-glycosyl transferase		chr5.CM0052.810.r2.m	Lj5g1949520.1	G	R	2	2	0	2	1	7	n/e		
Subtilase	SbtM1	chr2.CM0021.2780.r2.m	Lj2g2002910.1	G	R	8	0	0	0	1	9	BAF95755.1	Takeda et al. 2011	
Subtilase	SbtS	chr3.CM1144.130.r2.m	Lj3g2995720.1	G	R	4	0	0	2	3	9	BAF95887.1	Takeda et al. 2011	
Subtilase	SbtM4	chr4.CM0126.510.r2.m	Lj4g1327480.1	G	R	18	0	0	0	7	25	BAF95753.1	Takeda et al. 2011	
Sucrose synthase	SUS1	chr6.CM0013.460.r2.m	Lj6g1162830.2	G	R	16	1	0	13	9	39	n/e		Horst et al. 2007
Sucrose synthase	SUS2	chr1.CM0122.2540.r2.m	Lj1g4875640.1	G	R	6	0	0	3	4	13	n/e		Horst et al. 2007
Sucrose synthase	SUS3	chr4.CM0006.540.r2.m	Lj4g2215210.1	G	R	3	1	0	2	1	7	n/e		Horst et al. 2007
Sucrose synthase	SUS4	chr5.CM0239.860.r2.d	Lj0g0242009.1	G	R	4	0	0	2	0	6	n/e		Horst et al. 2007

(continued)

Table 21.3 (continued)

Description	Gene symbol	Predicted gene location	Release 3.0 predicted gene	Eco-type	pre- screen or Reverse	Mis-sense	Premature stop	Splice junction	non-coding	silent	Total mutations per gene	Genbank Protein ID.	Reference for gene	Reference for TILLING mutants
Symbiotic sulfate transporter	SST1	chr2.CM0610.70.r2.m	Lj2g0776860.1	G	R	0	0	0	3	2	5	CAL36108.1	Krusell etal. 2005	
Thiaminethiazole synthase		chr3.CM1543.40.r2.m	Lj3g1010900.1	G/M	R	4	0	0	1	4	9	n/e		
Thiaminethiazole synthase		chr5.CM0200.1720.r2.m	Lj5g2060670.1	G	R	7	0	0	0	10	17	n/e		
Transcription factor CPP	CPP-L56	chr1.CM0591.160.r2.m	Lj1g0114040.1	G	R	1	0	0	0	0	1	BAG50072.1	Asamizu et al. 2008	
Tubulin A		hr4.CM0046.780.r2.a	Lj4g2604350.1	G/M	R	12	0	0	9	3	24	n/e		
U-box domain-containing protein		chr6.LjB02K20.60.r2.m	n/a	G	R	15	1	0	0	4	20	n/e		
UDP-Glucosyl transferase		LjSGA_003166.2	Lj0g0140509.1	M	R	9	1	0	0	4	14	n/e	Takos etal. 2011	
UDP-Glucosyl transferase		chr4.CM0414.390.r2.d	Lj0g0193249.1	M	R	8	3	0	1	5	17	n/e	Takos etal. 2011	
UDP-Glucosyl transferase		chr3.CM0241.610.r2.m	Lj3g0754960.1	M	R	9	1	0	0	2	12	n/e	Takos etal. 2011	
Vesicle-associated membrane protein		chr6.CM0066.240.r2.a	Lj6g1692680.1	G	R	8	0	2	9	2	21	n/e		
Vesicle-associated membrane protein		chr6.CM0553.210.r2.d	Lj6g0132830.1	G	R	7	0	0	11	1	19	n/e		
Vesicle-associated membrane protein		chr3.CM0164.260.r2.d	Lj3g2990300.1	G	R	2	0	0	8	2	12	n/e		
Vesicle-associated membrane protein		chr2.CM0308.520.r2.m	Lj2g3022900.1	G	R	6	1	0	5	5	17	n/e		
Vesicle-associated membrane protein		chr4.CM0003.1110.r2.m	Lj4g1881960.1	G	R	11	0	0	10	3	24	n/e		

*RevGen*UK, by agreement with requestors, releases selected information on each target TILLed one year after completion of the work. This table represents information available for Lotus to date (Nov 2013). An additional 15 gene targets are under embargo at the time of going to press
N/a not annotated (some genes were not identified in release 3.0)

[a] Two entries are present most likely because the two separate annotations, one from routine annotation and the other from gene-based deposition to public databases, were assigned same gene model

21.6 The Future

For more than 10 years, TILLING in Lotus has been the predominant reverse genetic platform for this species, primarily because relatively few laboratories had efficient conventional (rather than hairy root) transformation systems and no insertional mutagenesis population had been developed. With the development of the LORE1 transposon population (see Chap. 20), the playing field has changed and Lotus TILLING will no doubt take a backstage since most researchers prefer knockout mutations as an indicator of gene function, especially in the initial stages of their research. Since it is the only way of producing a range of stable allelic variation, however, TILLING will remain the most effective and refined way of allowing detailed investigations of gene function. Furthermore, one cannot uncover the function of a gene which when fully disrupted is lethal, other than by producing a range of weaker, non-lethal alleles.

The advent of efficient resequencing technologies will also have an impact on Lotus TILLING. As the cost of resequencing falls, it will become cost-effective to identify all the mutations in our populations, thus making conventional TILLING redundant in favour of in silico TILLING (Wang et al. 2012); this is already in hand for species with small genomes or species such as hexaploid wheat. Each wheat plant can carry large numbers of mutations, and hence, such species require the resequencing of relatively small plant populations (thus reducing the cost) to identify numerous mutations in each gene. The future for conventional TILLING in *Lotus japonicus,* therefore, will be briefer than its past.

References

Andersen SU, Cvitanich C, Hougaard BK, Roussis A, Grønlund M, Jensen DB, Frøkjær LA, Jensen EØ (2003) The glucocorticoid-inducible GVG system causes severe growth defects in both root and shoot of the model legume *Lotus japonicus*. Mol Plant-Microbe Interact 16:1069–1076

Asamizu E, Shimoda Y, Kouchi H, Tabata S, Sato S (2008) A positive regulatory role for *LjERF1* in the nodulation process is revealed by systematic analysis of nodule-associated transcription factors of *Lotus japonicus*. Plant Physiol 147:2030–2040

Bustos-Sanmamed P, Bazin J, Hartmann C, Crespi M, Lelandais-Brière C (2013) Small RNA pathways and diversity in model legumes: lessons from genomics. Front Plant Sci 4, art 236. doi:10.3389/fpls.2013.00236

Credali A, García-Calderón M, Dam S, Perry J, Díaz-Quintana A, Parniske M, Wang TL, Stougaard J, Vega JM, Márquez AJ (2013) The K+-dependent asparaginase, NSE1, is crucial for plant growth and seed production in *Lotus japonicus*. Plant Cell Physiol 54:107–118

Colbert T, Till BJ, Tompa R, Reynolds S, Steine MN, Yeung AT, McCallum CM, Comai L, Henikoff S (2001) High-throughput screening for induced point mutations. Plant Physiol 126:480–484

Den Herder G, Yoshida S, Antolín-Llovera M, Ried MK, Parniske M (2012) *Lotus japonicus* E3 Ligase SEVEN IN ABSENTIA4 destabilizes the symbiosis receptor-like kinase SYMRK and negatively regulates rhizobial Infection. Plant Cell 24:1691–1707

Dong Z-C, Zhao Z, Liu C-W, Luo J-H, Yang J, Huang W-H, Hu X-H, Wang Tl, Luo D (2005) Floral patterning. In *Lotus japonicus*. Plant Physiol 137:1272–1282

Gaude N, Tippmann H, Flemetakis E, Katinakis P, Udvardi M, Dörmann P (2004) The galactolipid digalactosyldiacylglycerol accumulates in the peribacteroid membrane of nitrogen-fixing nodules of soybean and Lotus. J Biol Chem 279:34624–34630

Groth M, Kosuta S, Haage K, Hardel S, Gutjahr C, Brachmann A, Sato S, Tabata S, Findlay K, Wang T, Parniske M (2013) Two *Lotus japonicus* symbiosis mutants impaired at distinct steps of arbuscule development. Plant J 75:117–129

Greene EA, Codomo CA, Taylor NE, Henikoff JG, Till BJ, Reynolds SH, Enns LC, Burtner C, Johnson JE, Odden AR et al (2003) Spectrum of chemically induced mutations from a large-scale reverse-genetic screen in Arabidopsis. Genetics 164:731–740

Grønlund M, Gustafsen C, Roussis A, Jensen D, Nielsen L, Marcker K, Jensen E (2003) The *Lotus japonicus NDX* gene family is involved in nodule function and maintenance. Plant Mol Biol 52:303–316

Guether M, Neuhäuser B, Balestrini R, Dynowski M, Ludewig U, Bonfante P (2009) A mycorrhizal-specific ammonium transporter from *Lotus japonicus* acquires nitrogen released by arbuscular mycorrhizal fungi. Plant Physiol 150:73–83

Hakoyama T, Watanabe H, Tomita J, Yamamoto A, Sato S, Mori Y, Kouchi H, Suganuma N (2009) Nicotianamine synthase specifically expressed in root nodules of *Lotus japonicus*. Planta 230:309–317

Horst I, Welham T, Kelly S, Kaneko T, Sato S, Tabata S, Parniske M, Wang TL (2007) TILLING mutants of *Lotus japonicus* reveal that nitrogen assimilation and fixation can occur in the absence of nodule-enhanced sucrose synthase. Plant Physiol 144:806–820

Imaizumi-Anraku H, Takeda N, Charpentier M, Perry J, Miwa H, Umehara Y, Kouchi H, Murakami Y, Mulder L,

Vickers K et al (2005) Plastid proteins crucial for symbiotic fungal and bacterial entry into plant roots. Nature 433:527–531

Kanamori N, Madsen LH, Radutoiu S, Frantescu M, Quistgaard EMH, Miwa H, Downie JA, James EK, Felle HH, Haaning LL, Jensen TH, Sato S, Nakamura Y, Tabata S, Sandal N, Stougaard J (2006) A nucleoporin is required for induction of Ca2+ spiking in legume nodule development and essential for rhizobial and fungal symbiosis. Proc Nat Acad Sci USA 103:359–364

Kapranov P, M. Routt S, Bankaitis VA, de Bruijn FJ, Szczyglowski K (2001) Nodule-specific regulation of phosphatidylinositol transfer protein expression in *Lotus japonicus*. Plant Cell 13:1369-1382

Koressaar T, Remm M (2007) Enhancements and modifications of primer design program Primer3. Bioinformatics 23:1289–1291

Krokida A, Delis C, Geisler K, Garagounis C, Tsikou D, Peña-Rodríguez LM, Katsarou D, Field B, Osbourn AE, Papadopoulou KK (2013) A metabolic gene cluster in *Lotus japonicus* discloses novel enzyme functions and products in triterpene biosynthesis. New Phytol 200:675–690

Krusell L, Krause K, Ott T, Desbrosses G, Krämer U, Sato S, Nakamura Y, Tabata S, James EK, Sandal N, Stougaard S, Kawaguchi M, Miyamoto A, Suganuma N, Udvardi MK (2005) The sulphate transporter SST1 is crucial for symbiotic nitrogen fixation in *Lotus japonicus* root nodules. Plant Cell 17:1625–1636

Krusell L, Sato N, Fukuhara I, Koch BEV, Grossmann C, Okamoto S, Oka-Kira E, Otsubo Y, Aubert G, Nakagawa T, Sato S, Tabata S, Duc G, Perry J, Wang TL, Kawaguchi M, Stougaard J (2011) The Clavata2 genes of pea and *Lotus japonicus* affect autoregulation of nodulation and shoot architecture. Plant J 65:861–871

Kumar P, Henikoff S, Ng PC (2009) Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc 4:1073–1081

Le Signor C, Savois V, Aubert G, Verdier J, Nicolas M, Pagny G, Moussy F, Sanchez M, Baker D, Clarke J, Thompson R (2009) Optimizing TILLING populations for reverse genetics in *Medicago truncatula*. Plant Biotechnol 7:430–441

Lohmann GV, Shimoda Y, Nielsen MW, Jørgensen FG, Grossmann C, Sandal N, Sørensen K, Thirup S, Madsen LH, Tabata S, Sato S, Stougaard J, Radutoiu S (2010) Evolution and regulation of the *Lotus japonicus* LysM receptor gene family. Mol Plant-Microbe Interact 23:510–521

Maekawa-Yoshikawa M, Müller J, Takeda N, Maekawa T, Sato S, Tabata S, Perry J, Wang TL, Groth M, Brachmann A, Parniske M (2009) The temperature-sensitive brush mutant of the legume *Lotus japonicus* reveals a link between root development and nodule infection by rhizobia. Plant Physiol 149:1785–1796

McCallum CM, Comai L, Greene EA, Henikoff S (2000) Targeting Induced Local Lesions IN Genomes (TILLING) for plant functional genomics. Plant Physiol 123:439–442

Morant AV, Bjarnholt N, Kragh ME, Kjærgaard CH, Jørgensen K, Paquette SM, Piotrowski M, Imberty A, Olsen CE, Møller BL, Bak S (2008) The β-Glucosidases responsible for bioactivation of hydroxynitrile glucosides in *Lotus japonicus*. Plant Physiol 147:1072–1091

Murray JD, Karas BJ, Sato S, Tabata S, Amyot L, Szczyglowski K (2007) A cytokinin perception mutant colonized by rhizobium in the absence of nodule organogenesis. Science 315:101–104

Nagata M, Murakami EI, Shimoda Y, Shimoda-Sasakura F, Kucho KI, Suzuki A, Abe M, Higashi S, Uchiumi T (2008) Expression of a class 1 hemoglobin gene and production of nitric oxide in response to symbiotic and pathogenic bacteria in *Lotus japonicus*. Mol Plant-Microbe Interact 21:1175–1183

Olsen O, Wang X, Von Wettstein D (1993) Sodium azide mutagenesis: preferential generation of AT → GC transitions in the barley *Antl8* gene. Proc Nat Acad Sci USA 90:8043–8047

Parry MAJ, Madgwick PJ, Bayon C, Tearall K, Hernandez-Lopez A, Baudo M, Rakszegi M, Hamada W, Al-Yassin A, Ouabbou H, Labhilili M, Phillips AL (2009) Mutation discovery for crop improvement. J Exp Bot 60:2817–2825

Perry JA, Wang TL, Welham TJ, Gardner S, Pike JM, Yoshida S, Parniske M (2003) A TILLING reverse genetics tool and a web accessible collection of mutants of the legume *Lotus japonicus*. Plant Physiol 131:866–871

Perry J, Welham T, Cheminant S, Parniske M, Wang T (2005) TILLING. In: Márquez AJ (ed) *Lotus japonicus* Handbook Springer, Dordrecht, Chapter 5.3, pp 197–210

Perry J, Welham T, Brachmann A, Binder A, Charpentier M, Groth M, Haage K, Markmann K, Wang TL, Parniske M (2009) TILLING in *Lotus japonicus* identified large allelic series for symbiosis genes and revealed a bias in non-functional alleles towards hits in in glycine codons. Plant Physiol 151:1281–1291

Roberts NJ, Brigham J, Wu B, Murphy JB, Volpin H, Phillips DA, Etzler ME (1999) A nod factor-binding lectin is a member of a distinct class of apyrases that may be unique to the legumes. Mol Gen Genet 262:261–267

Roberts NJ, Morieri G, Kalsi G, Rose A, Stiller J, Edwards A, Xie F, Gresshoff PM, Oldroyd GED, Downie JA, Etzler ME (2013) Rhizobial and mycorrhizal symbioses in *Lotus japonicus* require lectin nucleotide phosphohydrolase, which acts upstream of calcium signalling. Plant Physiol 161:556–567

Serna-Sanz A, Parniske M, Peck SC (2011) Phosphoproteome analysis of *Lotus japonicus* roots reveals shared and distinct components of symbiosis and defense. Mol Plant-Microbe Interact 24:932–937

Shelton D, Stranne M, Mikklesen L, Pakseresht N, Welham T, Hiraka H, Tabata S, Sato S, Paquette S, Wang TL, Martin C, Bailey P (2012) Transcription factors of *Lotus japonicus*: regulation of isoflavonoid biosynthesis requires co-ordinated changes in transcription factor activity. Plant Physiol 159:531–547

Takanashi K, Sugiyama A, Sato S, Tabata S, Yazaki K (2012) LjABCB1, an ATP-binding cassette protein

specifically induced in uninfected cells of *Lotus japonicus* nodules. J Plant Physiol 169:322–326

Takeda N, Tsuzuki S, Suzaki T, Parniske M, Kawaguchi M (2013) CERBERUS and NSP1 of *Lotus japonicus* are common symbiosis genes that modulate arbuscular mycorrhiza development. Plant Cell Physiol 54:1711–1723

Takos A, Lai D, Mikkelsen L, Hachem MA, Shelton D, Bak S, Motawia MS, Olsen CE, Møller BL, Wang TL, Martin C, Rook F (2010) Genetic screening identifies cyanogenesis deficient mutants of *Lotus japonicus* and reveals enzymatic specificity in hydroxynitrile glucoside metabolism. Plant Cell 22:1605–1619

Takos AM, Knudsen C, Lai D, Kannangara R, Mikkelsen L, Motawia MS, Olsen CE, Sato S, Tabata S, Jørgensen K, Møller BL, Rook F (2011) Genomic clustering of cyanogenic glucoside biosynthetic genes aids their identification in *Lotus japonicus* and suggests the repeated evolution of this chemical defence pathway. Plant J 68:273–286

Tirichine L, Sandal N, Madsen LH, Radutoiu, Albrektsen AS, Sato S, Asamizu E, Tabata S, Stougaard J (2007) A gain-of-function mutation in a cytokinin receptor triggers spontaneous root nodule organogenesis. Science 315:104-107

Tsikou D, Stedel C, Kouri ED, Udvardi MK, Wang TL, Katinakis P, Labrou NE, Flemetakis E (2011) Characterization of two novel nodule-enhanced α-type carbonic anhydrases from *Lotus japonicus*. Biochim Biophys Acta 1814:496–504

Untergrasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG (2012) Primer3—new capabilities and interfaces. Nucl Acids Res 40:e115. doi:10.1093/nar/gks596

Van Arten AM (1998) Mutation breeding. Cambridge University Press, Cambridge, Theoretical and practical applications

Volpe V, Dell'Aglio E, Bonfante P (2013) The *Lotus japonicus MAMI* gene links root development, arbuscular mycorrhizal symbiosis and phosphate availability. Plant Signal Behav 8:e23414. doi:10.4161/psb.23414

Vriet C, Welham T, Brachmann A, Pike M, Pike J, Perry J, Parniske M, Sato S, Tabata S, Smith AM, Wang TL (2010) A suite of *Lotus japonicus* starch mutants reveals both conserved and novel features of starch metabolism. Plant Physiol 154:643–655

Wang TL, Uauy C, Robson F, Till B (2012) TILLING in extremis. Plant Biotech J 10:761–772

Welham T, Pike J, Horst I, Flemetakis E, Katinakis P, Kaneko T, Sato S, Tabata S, Perry J, Parniske M, Wang TL (2009) A cytosolic invertase is required for normal growth and cell development in the model legume, *Lotus japonicus*. J Exp Bot 60:3353–3365

Yano K, Yoshida S, Müller J, Singh S, Imaizumi-Anraku H, Vickers K, Markmann K, White C, Schuller B, Sato S, Asamizu E, Tabata S, Perry J, Wang TL, Kawaguchi M, Hayashi M, Parniske M (2008) CYCLOPS, a mediator of symbiotic intracellular accommodation. Proc Nat Acad Sci 105:20540–20545

The National BioResource Project in Japan

22

Masatsugu Hashiguchi and Ryo Akashi

Abstract

The objective of the National BioResource Project (NBRP) in Japan is to collect, conserve, and distribute biological materials for life sciences research. The project consortium of 28 core facilities of animal, plant, microorganisms, and DNA resources, and an information center. NBRP *Lotus* and *Glycine* aims to support the development of legume research through the collection, conservation, and distribution of these bioresources. *Lotus japonicus* is a perennial legume that grows naturally throughout Japan and is widely used as a model plant for legumes because of such advantages as its small genome size and short life cycle. Soybean (*Glycine max*) has been cultivated as an important crop since ancient times, and numerous research programs have generated a large amount of basic research information and valuable genetic resources for this crop. We have also developed "LegumeBase" a specialized database for the genera *Lotus* and *Glycine* and are maintaining this database as a part of the NBRP.

22.1 Introduction

The family Fabaceae (formerly Leguminaceae) is one of the most morphologically diverse taxa consisting of over 20,000 species divided into 730 genera (Doyle and Luckow 2003). This agronomically and ecologically important group of plants is responsible for much of the plant-associated biological nitrogen fixation and contains major food-producing crops including soybean (*Glycine max*), pea (*Pisum sativum)*, azuki bean (*Vigna angularis*), and sources of traditional medicines, such as *Astragalus* species. The genus *Lotus* consists of more than 200 species with the greatest diversity occurring in the Mediterranean. *Lotus japonicus* (Japanese trefoil) is distributed across East and Central Asia, including Japan, Korea, and China, extending west into Afghanistan (Pajuelo and Stougaard 2005). *L. japonicus* was initially proposed as a legume research model due to its small genome (472 Mb) (Ito et al. 2000), generation time of 3–4 months, small plant size, large and abundant flowers, easy hand pollination, high levels of seed production, easy cultivation, and amenability to *Agrobacterium*-mediated transformation (Handberg and Stougaard 1992). Furthermore,

M. Hashiguchi · R. Akashi (✉)
Faculty of Agriculture, University of Miyazaki, Miyazaki 889-2192, Japan
e-mail: rakashi@cc.miyazaki-u.ac.jp

S. Tabata and J. Stougaard (eds.), *The Lotus japonicus Genome*, Compendium of Plant Genomes,
DOI 10.1007/978-3-662-44270-8_22,

Sato et al. (2008) reported the first whole genome sequence of Legume plants using *L. japonicus* "Miyakojima MG-20." Currently, large amounts of important bioresources such as experimental strains, mutants, and DNA libraries have been developed through numerous independent research programs and scientific research projects. These bioresources will continue to serve as valuable materials for basic and applied studies. The National BioResource Project (NBRP) was launched by the Japanese government in 2002 with the objective of collecting, conserving, and distributing such valuable, independent resources and making them easily available for the larger research community. At present, the NBRP is a consortium of 28 core facilities of animal, plant, microorganisms, and DNA resources, and an information center (Yamazaki et al. 2010). NBRP plant consists of nine resources: *Arabidopsis* (*Arabidopsis thaliana*), rice (*Oryza sativa*), *Lotus*/*Glycine*, wheat (*Triticum aestivum*), barley (*Hordeum vulgare*), tomato (*Solanum lycopersicum*), *Chrysanthemum*, morning glory (*Ipomoea nil*), and algae (Kurata et al. 2010). As a part of this project, the *L. japonicus* and *G. max* program was developed beginning at the end of 2003. Here, we will provide an overview of the *Lotus* resources available from the NBRP *Lotus* and *Glycine* database site, called "LegumeBase."

22.2 Experimental Strains

Two important experimental strains of *L. japonicus*, "Gifu B-129" and "Miyakojima MG-20," have become global standards for legume research along with *Medicago truncatula*. "Gifu B-129" was the first established experimental *L. japonicus* strain, collected by Hirayoshi in Gifu Prefecture Japan, named by Grant and self-pollinated nine times by Stougaard (Handberg and Stougaard 1992; Stougaard and Beuselink 1996). More recently, Kawaguchi (2000) established the "Miyakojima MG-20" strain by self-pollinating *L. japonicus* strains from Miyakojima Island, Okinawa Prefecture, Japan. This strain is characterized by a short generation time and flowers easily under fluorescent lights. Although clearly of the same species, "Gifu B-129" and "Miyakojima MG-20" have very different morphological traits. "Gifu B-129" has an erect growth pattern and anthocyanin accumulation in the stem. "Miyakojima MG-20" is an early-flowering plant with a creeping growth habit that lacks stem anthocyanin. In addition, the leaflets and petals of "Miyakojima MG-20" are wider than those of "Gifu B-129," stems and petioles are thicker, and seeds are darker and larger. The third *Lotus* experimental strain, *Lotus burttii* B-303, was collected in Pakistan, named by Burtt (Sz-Borso et al. 1972) and self-pollinated nine times by Kawaguchi et al. (2005). There is a great demand for these experimental strains that have played central roles in studying legume-specific characteristics such as nodulation, and large numbers of mutants have been isolated in the past two decades (Charpentier and Oldroyd 2010; Kawaguchi et al. 2002; Novák 2010; Popp and Ott 2011; Szczyglowski et al. 1998). All of these experimental strains are available from "LegumeBase" (Table 22.1).

22.3 Wild Accessions

Genetic variation of *Lotus* is expected to have a broad range since *Lotus* adapts readily to diverse environmental conditions, such as temperature or soil type. The strains that we currently maintain and distribute at "LegumeBase" were collected across several climatic zones from as far north as Rebun Island, Hokkaido (45°17′46″N) to Miyakojima Island, Okinawa (24°43′57″N) to the south (Fig. 22.1). These strains were collected mainly for three purposes: (1) to evaluate the potential of *L. japonicus* as a pasture plant by Shimada in 1979 and for the Gene Bank Project of the Ministry of Agriculture, Forestry and Fisheries of Japan in 1981 (Suginobu et al. 1988); (2) to assess the suitability of this plant species to serve as a model organism for leguminous plants by Kawaguchi and Aoki since 1996 (Kawaguchi et al. 2001); and (3) to collect *L. japonicus* bioresources for the NBRP. At present, 180 accessions are stocked and 108

Table 22.1 *Lotus* resources preserved in LegumeBase

Name of resource	No. of stocked resources	Depositor	Reference
Wild accessions	180	*Lotus* research community, NBRP	Suginobu et al. (1988) Kawaguchi et al. (2001)
Core collection[a]	20	NBRP	Kai et al. (2010)
Experimental strain Gifu B-129 Miyakojima MG-20 *L. burttii* B-303	3	J. Stougaard M. Kawaguchi, W. F. Grant	Stougaard and Beuselinck (1996) Kawaguchi (2000) Kawaguchi et al. (2005)
RILs			
Gifu B-129 × Miyakojima MG-20	205	Kazusa DNA Institute	Hayashi et al. (2001)
Gifu B-129 × *L. burttii* B-303[a]	160	Aarhus University	Sandal et al. (2012)
Gifu B-129 × *L. filicaulis* B-37[a]	100	Aarhus University	Sandal et al. (2002)
EMS mutants	171	RIKEN	
M_2 bulked seeds	162[b]	NBRP	
Superroot	1	Univ. of Miyazaki	Akashi et al. (1998)
Activation-tagged lines[a]	960	Nihon University	Imaizumi et al. (2005)
LORE1 tag line	975	NIAS	Fukai et al. (2012)
Mesorhizobium loti STM mutants	6,671	Kazusa DNA Institute	Shimoda et al. (2008)
TAC clones	72,192	Kazusa DNA Institute	Sato et al. (2001)
BAC clones	14,976	Kazusa DNA Institute	Sato et al. (2007), (2008)
cDNA clones	140,544	Kazusa DNA Institute	Asamizu et al. (2004)
Binary vectors	6	M. Hayashi	Maekawa et al. (2008)
Full-length cDNA clones	104,064	Kazusa DNA Institute	Sakurai et al. unpublished
M. loti plasmid clone	4,196	Kazusa DNA Institute	Kaneko et al. (2000)

[a] This resource is in preparation
[b] Number of batches. We provide seed sets containing 10–20 batches. Each set consists of 5,000–9,000 M_2 seeds derived from 1,000 to 2,000 M_1 plants

accessions are available via "LegumeBase" (Table 22.1).

After the launch of the NBRP, we studied the variation in nine morphological characteristics of *L. japonicus* wild accessions and openly provided the information on "LegumeBase." Recently, Kai et al. (2010) selected 20 accessions to serve as a representative core collection based on simple sequence repeats (SSR) polymorphisms and morphological traits (Fig. 22.2). The range of morphological traits in the core collection was representative of that found in the entire collection. This core collection will be useful for genome wide studies and data obtained for this model species should lead to numerous practical applications for crop legumes.

22.4 Recombinant Inbred Lines (RIL)

A total of 205 "LjMG RI Lines" (Table 22.1) were derived from an F2 seed cross between "Miyakojima MG-20" and "Gifu B-129" and were self-pollinated eight times at the Kazusa DNA Research Institute. A total of 96 SSR markers were mapped on the chromosomes of *L. japonicus* using the F2 generation (Hayashi et al. 2001), and AFLP and SSR marker-based high-density linkage maps of *L. japonicus* were constructed (Wang et al. 2008). Gondo et al. (2007) reported the first quantitative trait locus (QTL) analysis of 13 phenotypic traits in two consecutive years in *L. japonicus*. A total of 40 QTLs

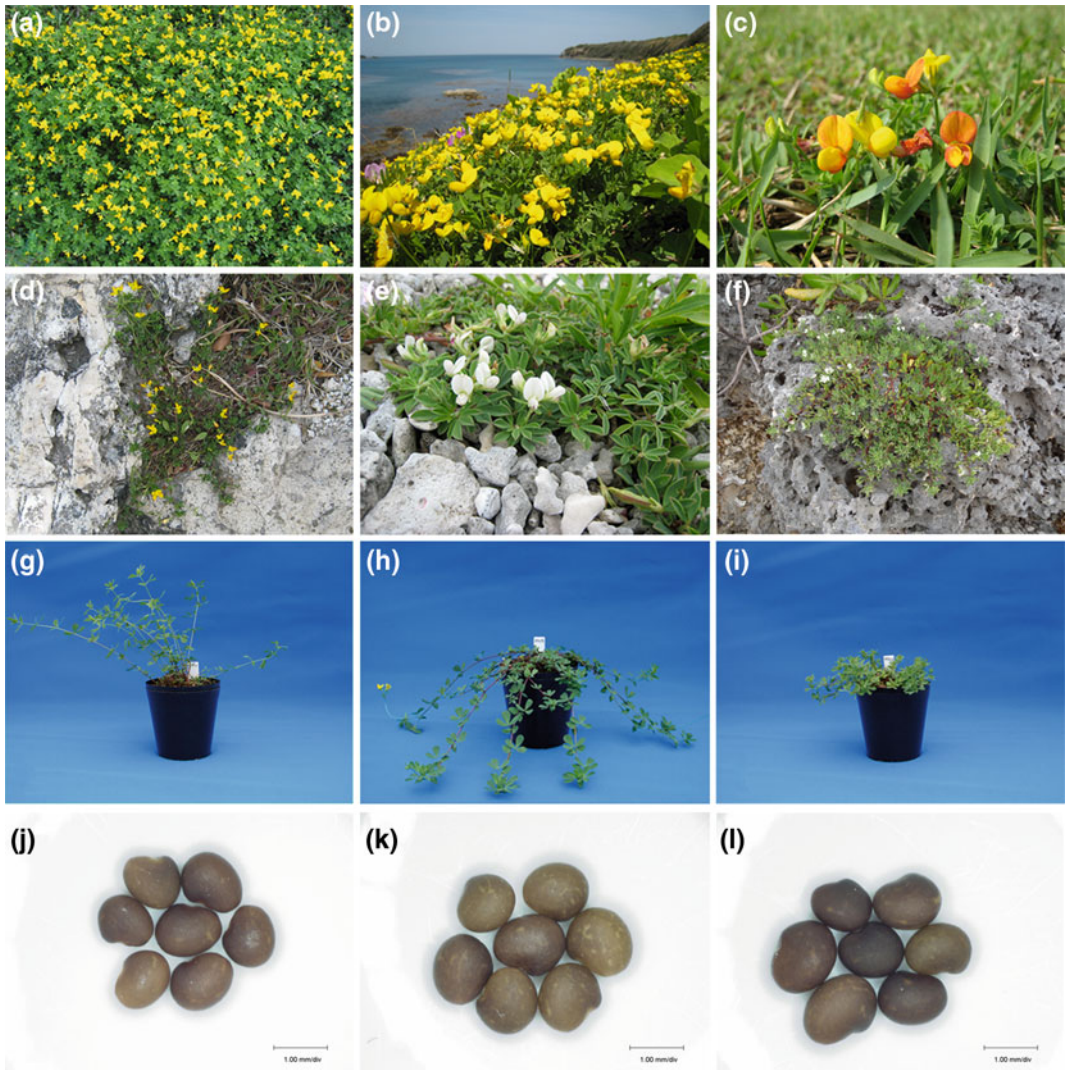


Fig. 22.1 Habitat of the two *Lotus* genera (*Lotus japonicus* and *Lotus pacificus*) and a comparison of plant types and seed size in the *Lotus japonicus* wild accessions. **a** *L. japonicus* native to Nichinan coast, Miyazaki; **b**, **d** *L. japonicus* grows naturally on cliffs of the cape (**b**) and on limestone (**d**) in the Tsushima Islands, Nagasaki; **c** native *L. japonicus* on a cape on Ojika Island, Nagasaki; **e** *L. pacificus* (previous name; *Lotus australis*) grows naturally on coral sand on Hateruma Island, Okinawa; **f** native *L. pacificus* grows on limestone on Kurima Island, Okinawa; **g–l** Pot (φ9.0 cm) cultuer and seeds of *L. japonicus* wild accessions collected in Japan, MG-74, Ehime, erect type (**g**, **j**), MG-34, Hokkaido, creeping type (**h**, **k**), MG-23, Aomori, dwarf type (**i, l**)

were detected that explained 5–6 % of the total variation. The QTL that explained the most variation was that for stem color, which was detected in the same region of chromosome 2 for both years (Fig. 22.3). "LegumeBase" contains an interactive database for selecting LjMG RILs based on the 13 phenotypic traits evaluated by Gondo et al. (2007). Moreover, macrosynteny between soybean and *L. japonicus* was analyzed with the objective of applying the genomic information of the model legume *L. japonicus* to soybean (Tsubokura et al. 2008). In addition, two other kinds of RILs were deposited from Aarhus University, Denmark. The first RIL is from crosses between *L. japonicus* "Gifu B-129" and *L. burttii* "B-303" (G × LB population) (Sandal

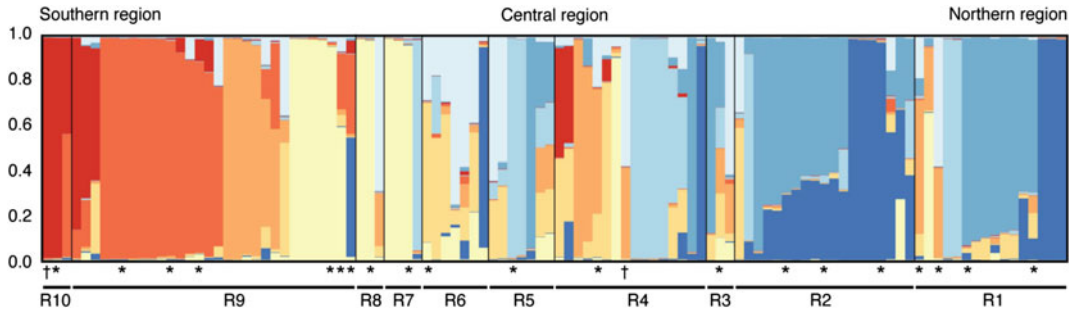


Fig. 22.2 Population structures estimated by STRUCTURE analysis at K = 9 using data for 25 nuclear genetic loci (Kai et al. 2010). The analysis was designed to produce an admixture ancestry model, assuming no correlations, with a burn-in length of 10^5 and a run length of 10^6 iterations. Each accession is represented by a single vertical bar partitioned into nine colored segments that represent the accession's estimated ancestry proportion. *R1* Hokkaido, *R2* Tohoku, *R3* Kanto, *R4* Hokuriku, *R5* Tokai, *R6* Kansai, *R7* Chugoku, *R8* Shikoku, *R9* Kyusyu, *R10* Okinawa. † "Miyakojima MG-20"; ‡ "Gifu B-129"

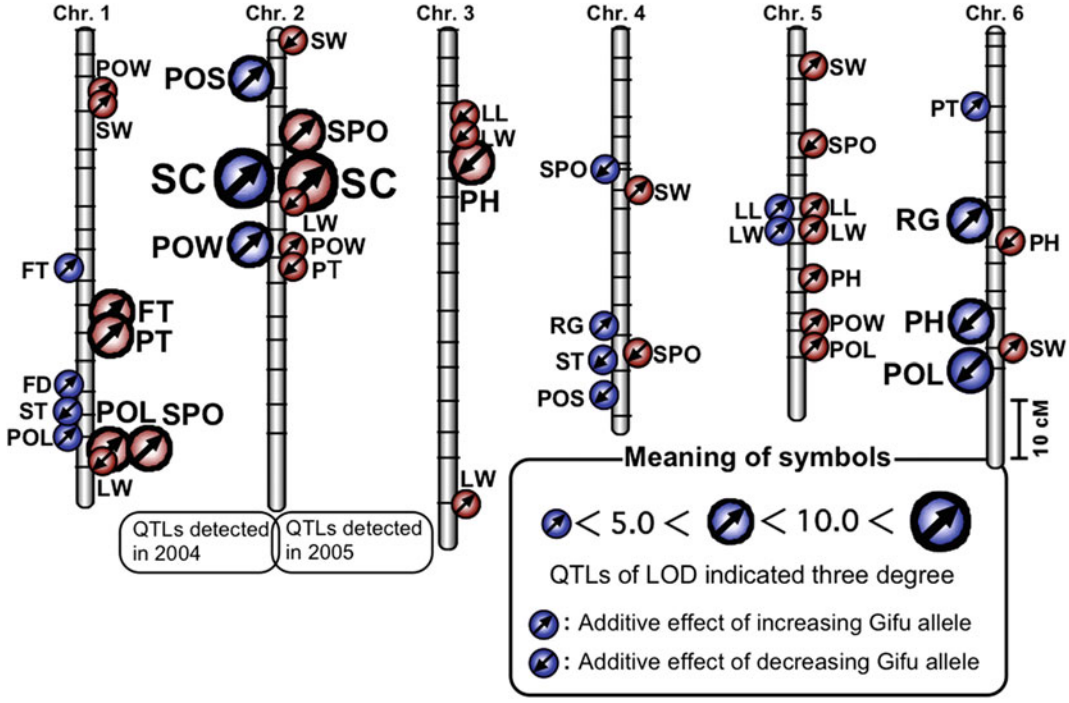


Fig. 22.3 Molecular linkage map of *L. japonicus* and the location of QTLs identified for 13 agronomic traits in 2004 and 2005 (Gondo et al. 2007). Arrowheads indicate the positions of QTLs and the key to symbols is summarized in the box. The arrowheads to the left (*blue*) and right (*red*) of each chromosome show the QTLs detected in 2004 and 2005, respectively. *FT* flowering time; *FD* flowering degree; *PH* plant height; *PS* plant shape; *ST* stem thickness; *SC* stem color; *LL* leaf length; *LW* leaf width; *PR* plant regrowth; *POL* pod length; *POW* pod width; *SPO* seeds per pod; *SM* mass of 1,000 seeds

et al. 2012), and the second RIL is from crosses between *L. japonicus* "Gifu B-129" and *L. filicaulis* "B-37" (G × F population) (Sandal et al. 2002, 2009). Both of these RILs will be available from "LegumeBase" in the near future (Table 22.1). Please refer to Chap. 4 of this book for more detailed information about the RILs described above.

22.5 EMS Mutants and M2 Bulked Seeds

Ethyl methanesulfonate (EMS)-treated mutants of *L. japonicus* were isolated from "Miyakojima MG-20" at the RIKEN Plant Science Research Center. There are two categories of mutants: above-ground mutants (plantlet, leaf, stem, flower, etc.) and root morphological mutants (root elongation, root thickness, root hair length, and the number of root hairs, etc.). At present, 98 homozygous mutants are available (Fig. 22.4). Recently, Suzuki et al. (2011) reported that root nodule formation is photomorphogenetically controlled by sensing the red/far-red ratio through jasmonic acid signaling by using the 01-0017 (Fig. 22.4f) and 01-1428 mutants. In addition, we have prepared ten sets of EMS-treated bulked M2 seeds of *L. japonicus* "Miyakojima MG-20" (Table 22.1). Each set consists of 5,000–9,000 M2 seeds derived from 1,000 to 2,000 M1 plants treated with a 0.4 % EMS solution for 8–10 h. Users may screen the mutants themselves and use the screened mutants for their research. Once their study is published, users are required to deposit the isolated mutant lines derived from this resource with our resource center.

22.6 Activation-Tagged Lines

Activation tagging is a method to produce gain-of-function mutants by random insertion of tandemly repeated CaMV 35S enhancer sequences into the plant genome. This method allows the analysis of functionally redundant gene families and essential genes, whose knockout mutants cannot be obtained. Although this powerful approach has been used in *Arabidopsis thaliana* (Weigel et al. 2000), its application to leguminous plants was not popular because of the difficulty in genetically transforming legumes. Imaizumi et al. (2005) improved the transformation technique for *L. japonicus* and produced more than 3,500 T-DNA insertional lines, demonstrating the feasibility of activation tagging *L. japonicus*. Activation-tagged populations of this model legume should provide a powerful tool for identifying novel genes involved in morphology, accumulation of seed storage proteins, biosynthesis of legume-specific natural products, symbiotic nitrogen fixation, and mycorrhizal formation. These activation-tagged lines will also serve as suitable materials for post-genomic analyses, such as transcriptomics, proteomics, and metabolomics, and will be available via "LegumeBase" in the future (Table 22.1).

22.7 Root Culture (Superroots)

We discovered super-growing roots (superroots) from *Lotus corniculatus* that grow efficiently after removal of the above-ground organs under growth regulator-free culture conditions (Akashi et al. 1998) (Fig. 22.5). Superroots are highly competent for plant regeneration. Moreover, protoplasts can be easily obtained from superroots and proliferate well in vitro (Fig. 22.5f–i). These characteristics are still maintained 15 years after the discovery of superroots (Akashi et al. 1998, 2003). Superroots can be used in physiological research as well as in functional analysis of genes using *A. tumefaciens* (Tanaka et al. 2008) (Fig. 22.5j–l) or *A. rhizogenes*-mediated transformation (Jian et al. 2009). Recently, Himuro et al. (2011) developed 130 *Arabidopsis* full-length cDNA overexpressing (FOX)-superroot lines using the FOX hunting system. FOX-superroot lines provide a new tool for genetic analysis and control of root growth in leguminous plants (Fig. 22.5o).

22.8 LORE1 Tag Lines

Fukai et al. (2012) established a gene-tagging population of the model legume *L. japonicus* using an endogenous long-terminal repeat (LTR) retrotransposon, *Lotus Retrotransposon 1* (*LORE1*). Part of the LORE1 population, 975 lines, has been made available at our Web site (Table 22.1). Details of this resource are provided in detail in Chap. 21 of this book.

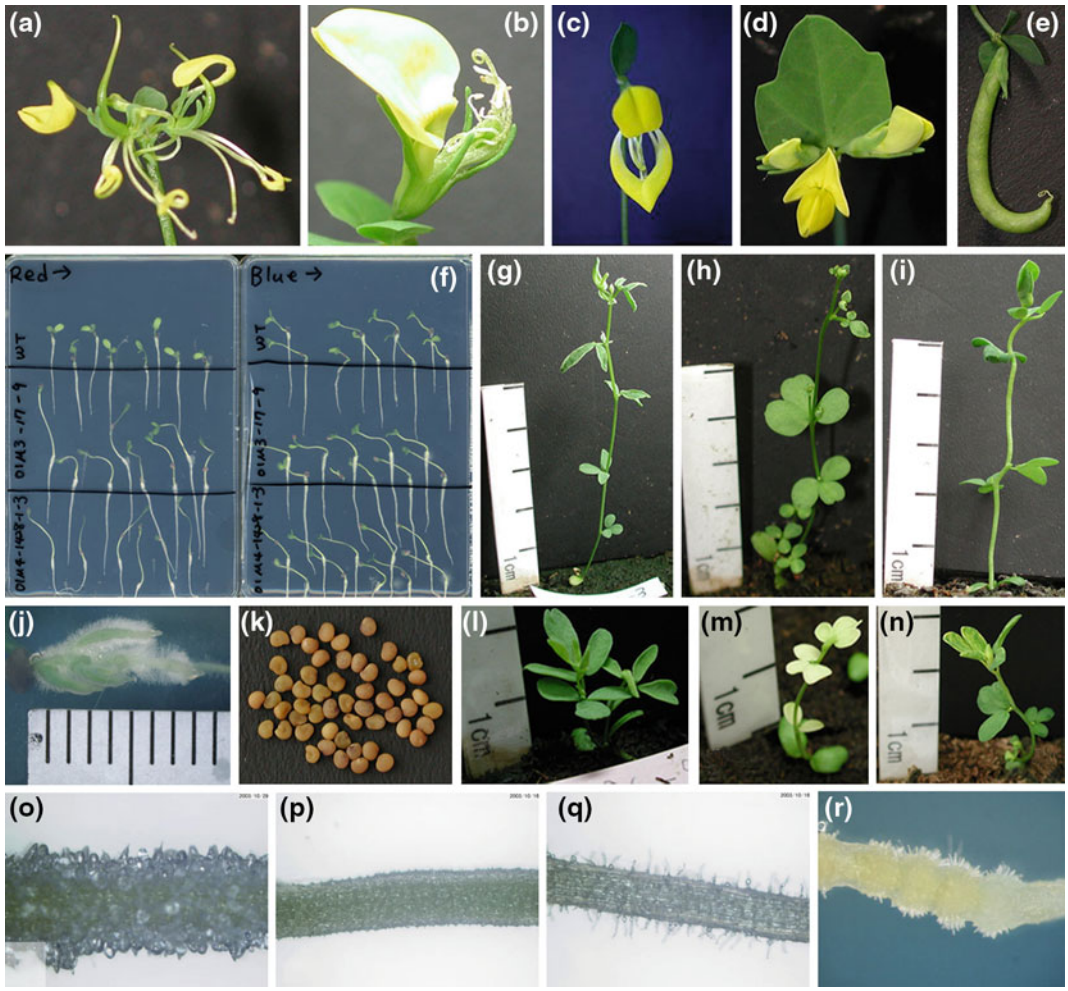


Fig. 22.4 EMS-treated mutants derived from *L. japonicus* "Miyakojima MG-20." The numbers in parentheses indicate the line number of the mutant. **a**, **j** (*212-003*); abnormal flower and many short roots; **b**, **e**, **i** (*212-103*) abnormal flower, crooked pod and twisted stem; **c** (*01-0210*) abnormal flower; **d** (*212-596*) indiscrete leaf bract; **f** (*01-0017*) long hypocotyl and defective hypocotyl elongation with red-light irradiation; **g** (*207-008*) narrow leaf; **h** (*210-088*) small leaf tip; **k** (*211-039*) pale brown seed coat; **l** (*01-0196*) dwarf plant; **m** (*206-105*) white leaf; **n** (*205-074*) variegated leaf; **o** (*01-0348*) swollen root hair; **p** (*01-1066*) short root hair; **q** (*01-0640*) reduced number of root hair; **r** (*207-034a*) abnormal root shape and shorter root hair. These mutant lines were developed by T. Sakai, RIKEN (present address, Niigata University)

22.9 cDNA, TAC, and BAC Clones

Sato et al. (2008) sequenced the entire *L. japonicus* genome using the "Miyakojima MG-20" strain. Various material resources such as transformation-competent artificial chromosomes (TAC) (Asamizu et al. 2003; Kaneko et al. 2003; Kato et al. 2003; Nakamura et al. 2002; Sato et al. 2001), bacterial artificial chromosomes (BAC) (Sato et al. 2007, 2008), and cDNA libraries (Asamizu et al. 2000, 2004) were developed during the genome sequencing projects, and the information was made available at the "miyakogusa.jp" Web site or the "*Lotus japonicus* EST Index" hosted by the Kazusa DNA Institute (Table 22.2). These important products of *L. japonicus* genome sequence

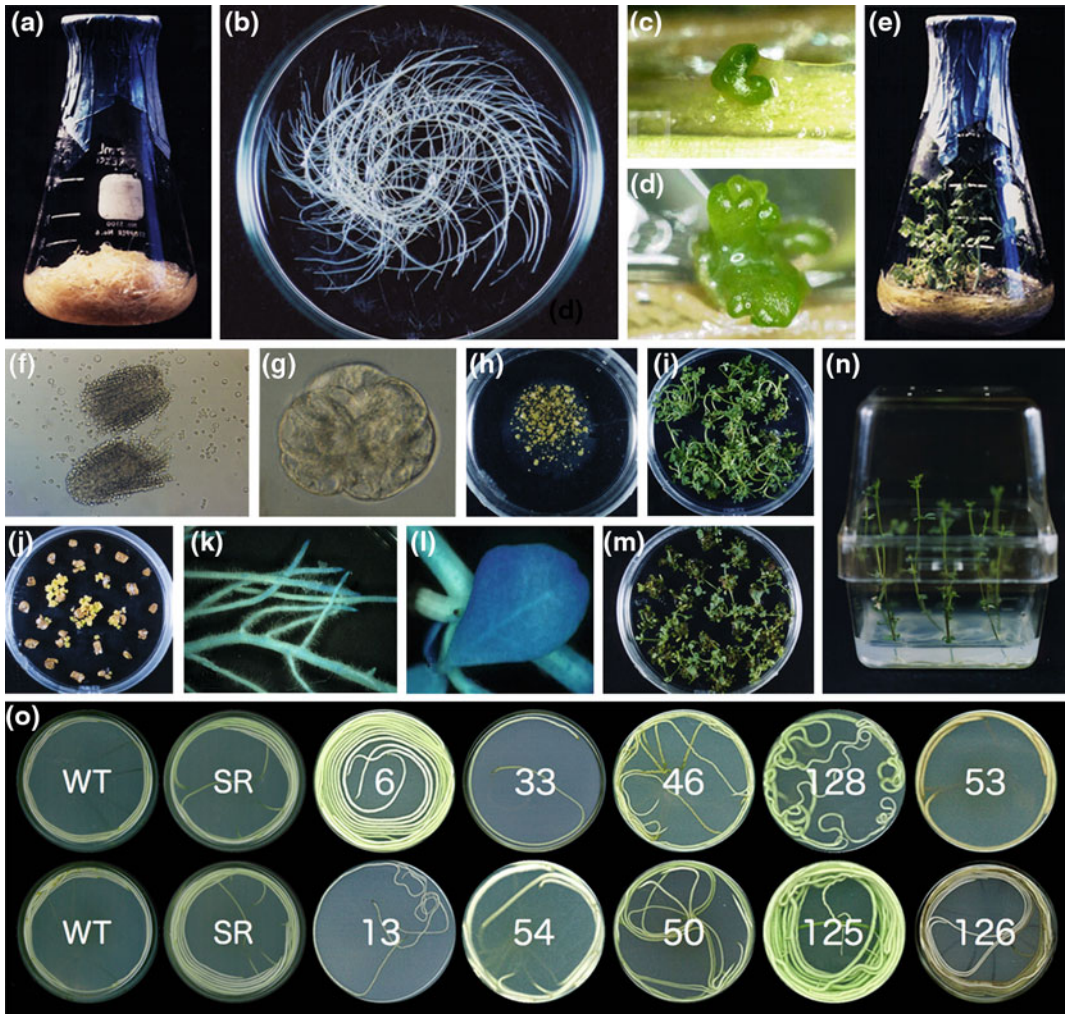


Fig. 22.5 Root cloning, somatic embryogenesis, plant regeneration, protoplast culture, transformation, and FOX lines in superroots (super-growing root cultures) of *Lotus corniculatus* (bird's-foot trefoil) (Akashi et al. 1998, 2003; Tanaka et al. 2008; Himuro et al. 2011). **a** a 125-mL Erlenmeyer flask containing a superroot culture displaying typical root density at the end of a subculture period; **b** root cultures derived from 10 lateral roots 28 days after subculture (transferred to a 9-cm Petri plate for photography); **c**, **d** somatic embryogenesis from cultured roots; **e** shoots, formed at wound sites of superroots, grow vigorously when placed on a lighted shelf; **f** enzymatically separated root tips after two hours of enzyme treatment; **g** microcallus formation from isolated root protoplasts; **h** microcolonies formed after 4 weeks of culture in agarose disks in a 6-cm Petri plate; **i** shoot formation and shoot elongation on protoplast-derived calli in a 9-cm Petri plate; **j** resistant callus on MS medium containing 100 mg/L kanamycin; **k**, **l** GUS activity in root (**k**) and leaf (**l**) tissues of regenerated plants; **m** Petri plate with shoot-producing callus derived from a leaf explant after 8 weeks; **n** a regenerated plant from superroots, **o** *Arabidopsis* FOX-superroots (numbers), wild type (WT) and superroots (SR) of *Lotus corniculatus*. Shoots of FOX lines were rooted in flat-bottomed test tubes. The roots had to penetrate 3 cm of agar before coiling up at the bottom. Images of the test-tube bottom were taken with a photocopier after 4 weeks of culture

projects are exceedingly valuable tools for genetic and physiological studies and/or synteny analysis of leguminous plants. These resources have been deposited with our resource center and are available from "LegumeBase" for researchers (Table 22.1).

Table 22.2 NBRP *Lotus* and *Glycine* related Web sites and databases

Name of database	Contents	URL
LegumeBase	NBRP *Lotus* and *Glycine* main page	http://www.legumebase.brc.miyazaki-u.ac.jp/
Lotus japonicus database	NBRP Database for *L. japonicus*	http://www.legumebase.brc.miyazaki-u.ac.jp/lotus/
NBRP information site	NBRP Web site	http://www.nbrp.jp/
BioResource World: BRW	An integrated NBRP database retrieval system	http://resourcedb.nbrp.jp/top.jsp
Worldwide Legume science information desk	Social bookmark site for legume-related Web pages	http://www.shigen.nig.ac.jp/infodesk/topSpeciesAction.do?speciesId=4
Research resource circulation *lotus* /*glycine*	Database of publications related to the NBRP resources	http://www.shigen.nig.ac.jp/rrc/gatewayAction.do?speciesId=17
miyakogusa.jp	*Lotus japonicus* genome database	http://www.kazusa.or.jp/lotus/
Lotus japonicus EST Index	*Lotus japonicus* EST database	http://est.kazusa.or.jp/en/plant/lotus/EST/
RhizoBase	Genetic information of ORFs for 39 strains of rhizobacteria	http://genome.microbedb.jp/rhizobase/

22.10 Full-Length cDNAs

Full-length cDNAs are useful resources for the functional analysis of genes or proteins and are available for several plants, such as *Arabidopsis* (Seki et al. 1998), rice (Kikuchi et al. 2003), wheat (Ogihara et al. 2004), soybean (Umezawa et al. 2008), maize (*Zea mays*) (Soderlund et al. 2009), tomato (Aoki et al. 2010), and barley (Matsumoto et al. 2011). *L. japonicus* full-length cDNAs were developed at the Kazusa DNA Research Institute and have been deposited with "LegumeBase" (Table 22.1). There are approximately 100,000 *L. japonicus* cDNA clones from a full-length-enriched cDNA library, including 3,874 full-read sequences that were derived from plants and roots, as well as from in vitro-cultured cells of *L. japonicus* that were cultured under diverse chemical treatment conditions (Sakurai et al. unpublished).

22.11 Binary Vectors

Promoter analysis studies have demonstrated that the polyubiquitin promoter from *L. japonicus* plants (Ljubq1) possesses higher activity than the CaMV35S promoter in *L. japonicus* leaves, stems, roots, nodules, and pollen (Maekawa et al. 2008). The GATEWAY conversion technology-compatible binary vectors that were constructed for overexpression and RNAi under the control of the Ljubq1 promoter provide alternative choices for studies in *L. japonicus* (Fig. 22.6). The six kinds of vectors deposited by M. Hayashi are now available in LegumeBase. For one of these vectors, Nakagawa et al. (2011) investigated the expression profiles of the Nod factor (NFs) receptor gene in roots of *L. japonicus* through a complementation test using *Agrobacterium rhizogenes*-mediated transgenic *L. japonicus* with pUB-GW-GFP as shown in Fig. 22.6.

22.12 *Mesorhizobium loti* STM Mutants

The mutant library of *Mesorhizobium loti* was developed by the Kazusa DNA Research Institute through transposon mutagenesis. These transposon insertion mutants were generated using the signature-tagged mutagenesis (STM) technique (Shimoda et al. 2008). At present, 6,671 STM *M. loti* mutants are available from "LegumeBase" (Table 22.1). Detailed information about *M. loti* ORFs, such as the operon structure, predicted protein domains and orthologous protein groups, is available at RhizoBase (http://bacteria.kazusa.or.jp/rhizobase/

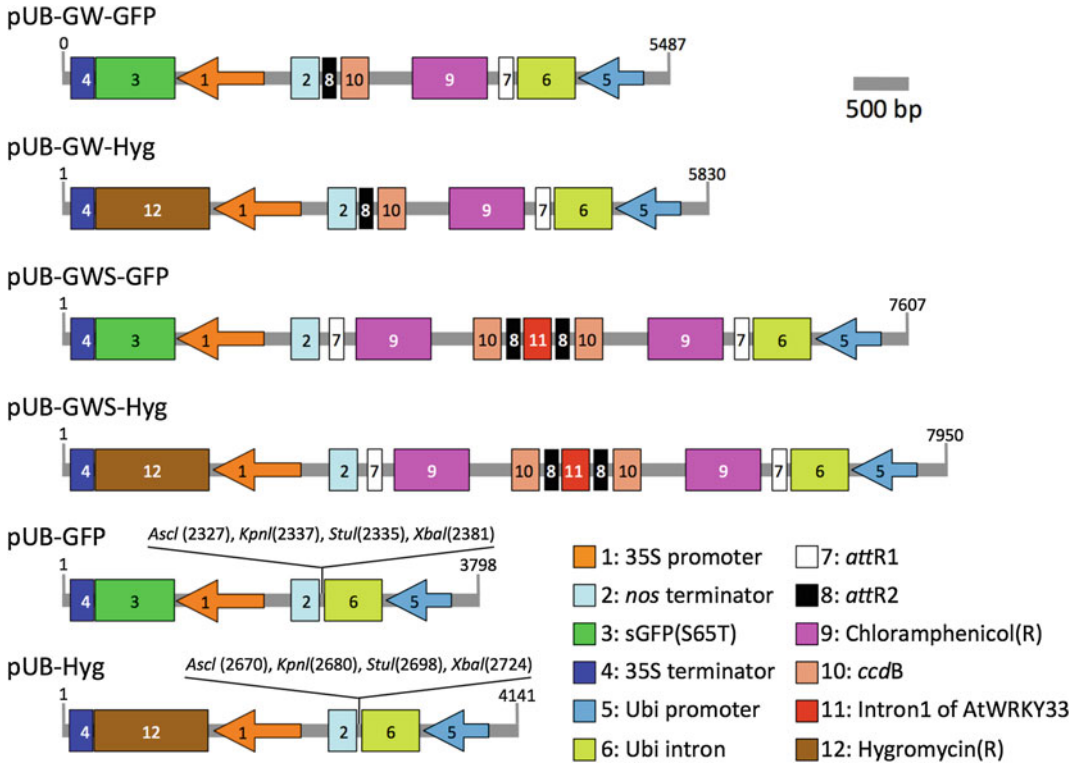


Fig. 22.6 Maps of *Ljubq1* promoter-based binary vectors (Maekawa et al. 2008). The binary vectors are represented as linear segments from the left to right borders. These constructs were based on a vector in which the 535-bp fragment of *Ljubq1* containing the 5′ untranslated region intron was cloned into pCAMBIA1300. The hygromycin resistance gene originated from the backbone vector, pCAMBIA1300. Reprinted and modified by permission of the author

Mesorhizobium/index.html), a database constructed by the Kazusa DNA Research Institute. The *M. loti* mutant STM 5 that contains an inserted transposon in the 3-phosphoglycerate dehydrogenase gene played an important role in identifying the molecular basis for effective symbiosis between *M. loti* and *L. japonicus* (Thapanapongworakul et al. 2010).

22.13 *Mesorhizobium loti* Plasmid Clones

Recently, 4,196 *Mesorhizobium loti* plasmid clones were made available at "LegumeBase" (Table 22.1). This resource, deposited by the Kazusa DNA Research Institute, comprises of clones that were used in the genome sequencing project of *M. loti* (Kaneko et al. 2000). Manageable sizes of genomic DNA of *M. loti* were cloned into a pUC plasmid. These DNA clones were linked mutually with the Genome browser in RhizoBase (http://genome.microbedb.jp/rhizobase/), a database for *Rhizobium* DNA hosted by the Kazusa DNA Institute (Table 22.2). In the Genome browser, users are able to select the DNA clones of interest and compare the DNA sequence of the clone with other loci or domains.

22.14 Databases

We have constructed a Web page for NBRP *Lotus* and *Glycine* "LegumeBase" (http://www.legumebase.brc.miyazaki-u.ac.jp/) at our resource center that is composed of two databases, the "*Lotus japonicus* database" (http://

www.legumebase.brc.miyazaki-u.ac.jp/lotus/) and the "*Glycine max/soja* database" (http://www.legumebase.brc.miyazaki-u.ac.jp/glycine/) (Table 22.2). In "LegumeBase," users may select the strains of interest by accession number, name, collector, morphological data, genotype, or detailed information about the collection site including meteorological data. Sequence data for DNA resources are also available for each database (Table 22.2). In addition, there are several related Web sites, such as the social bookmark site "Worldwide Legume Science Information Desk" or sites providing lists of relevant papers in the research area, such as Research Resource Circulation *Lotus*/*Glycine* that was established by the NBRP Information Center at the National Institute of Genetics (Yamazaki et al. 2010) (Table 22.2). The latter site provides useful information about legume research using the resources of NBRP *Lotus* and *Glycine*.

22.15 Conclusions

We have developed extensive resources for two important leguminous plants, *Lotus japonicus* and *Glycine max*, and have constructed a database called "LegumeBase" hosted at our resource center. NBRP *Lotus* and *Glycine* aims to make research materials and resources readily available to the legume research community. We make an effort to collect valuable resources for legume research, maintain the resources in good condition, and provide superior quality information and biological materials. When using our resources, the user is required to sign a material transfer agreement (MTA) and to explicitly acknowledge our resource center as the source in any publication that ensues from the study. We started collecting handling fees for providing resources in April, 2010. The fees can be paid online by credit card payment. We are accepted deposits of resources from researchers worldwide. Care will be taken to adhere to the protective conditions that are stipulated by depositors when distributing the bioresources. Previously, researchers wasted a lot of time with labor costs for procuring and maintaining their resources. NBRP *Lotus* and *Glycine* "LegumeBase" will alleviate these problems by accepting valuable research materials, maintaining the resources and distributing them as needed.

Acknowledgments This work was supported by the National BioResource Project of the Ministry of Education, Culture, Sports, Science and Technology (MEXT) Japan. We thank Dr. Yamazaki and the staff of the Genetic Informatics Laboratory at the National Institute of Genetics (NIG) for constructing and maintaining "LegumeBase." In addition, we profoundly thank the *Lotus* and *Glycine* research communities for their cooperation and support of our project, NBRP in *Lotus* and *Glycine*.

References

Akashi R, Hoffmann-Tsay S-S, Hoffmann F (1998) Selection of a super-growing legume root culture that permits controlled switching between root cloning and direct embryogenesis. Theor Appl Genet 96:758–764

Akashi R, Kawano T, Hashiguchi M et al (2003) Super roots in *Lotus corniculatus*: A unique tissue culture and regeneration system in a legume species. Plant Soil 255:27–33

Aoki K, Yano K, Suzuki A et al (2010) Large-scale analysis of full-length cDNAs from the tomato (*Solanum lycopersicum*) cultivar Micro-Tom, a reference system for the Solanaceae genomics. BMC Genom 11:210

Asamizu E, Nakamura Y, Sato S et al (2000) Generation of 7137 non-redundant expressed sequence tags from a Legume, *Lotus japonicus*. DNA Res 7:127–130

Asamizu E, Kato T, Sato S et al (2003) Structural analysis of a *Lotus japonicus* genome. IV. Sequence features and mapping of seventy-three TAC clones which cover the 7.5 mb regions of the genome. DNA Res 10:115–122

Asamizu E, Nakamura Y, Sato S et al (2004) Characteristics of the *Lotus japonicus* gene repertoire deduced from large-scale expressed sequence tag (EST) analysis. Plant Mol Biol 54:405–514

Charpentier M, Oldroyd G (2010) How close are we to nitrogen-fixing cereals? Curr Opin Plant Biol 13:556–564

Doyle JJ, Luckow MA (2003) The rest of the iceberg. Legume diversity and evolution in a phylogenetic context. Plant Physiol 131(3):900–910

Fukai E, Soyano T, Umehara Y et al (2012) Establishment of a *Lotus japonicus* gene tagging population using the exon-targeting endogenous retrotransposon LORE1. Plant J 69:720–730

Gondo T, Sato S, Okumura K et al (2007) Quantitative trait locus analysis of multiple agronomic traits in the model legume *Lotus japonicus*. Genome 50:627–637

Handberg K, Stougaard J (1992) *Lotus japonicus*, an autogamous, diploid legume species for classical and molecular genetics. Plant J 2:487–496

Hayashi M, Miyahara A, Sato S et al (2001) Construction of a genetic linkage map of the model legume *Lotus japonicus* using an intraspecific F2 population. DNA Res 8:301–310

Himuro Y, Tanaka H, Hashiguchi M et al (2011) FOX-superroots of *Lotus corniculatus*, overexpressing *Arabidopsis* full-length cDNA, show stable variations in morphological traits. J Plant Physiol 168:181–187

Ito M, Miyamoto J, Mori Y et al (2000) Genome and chromosome dimensions of *Lotus japonicus*. J Plant Res 113:435–442

Imaizumi R, Sato S, Kameya N et al (2005) Activation tagging approach in a model legume, *Lotus japonicus*. J Plant Res 118:391–399

Jian B, Hou W, Wu C et al (2009) *Agrobacterium rhizogenes*-mediated transformation of superroot-derived *Lotus corniculatus* plants: a valuable tool for functional genomics. BMC Plant Biol 9:78

Kai S, Tanaka H, Hashiguchi M et al (2010) Analysis of genetic diversity and morphological traits of Japanese *Lotus japonicus* for establishment of a core collection. Breed Sci 60(4):36–446

Kaneko T, Nakamura Y, Sato S et al (2000) Complete genome structure of the nitrogen-fixing symbiotic bacterium *Mesorhizobium loti*. DNA Res 7 (6):331–338

Kaneko T, Asamizu E, Kato T et al (2003) Structural analysis of a *Lotus japonicus* genome. III. Sequence features and mapping of sixty-two TAC clones which cover the 6.7 Mb regions of the genome. DNA Res 10:27–33

Kato T, Sato S, Nakamura Y et al (2003) Structural analysis of a *Lotus japonicus* genome. V. Sequence features and mapping of sixty-four TAC clones which cover the 6.4 mb regions of the genome. DNA Res 10:277–285

Kawaguchi M (2000) *Lotus japonicus* 'Miyakojima' MG-20': An early-flowering accession suitable for indoor handling. J Plant Res 133:507–509

Kawaguchi M, Motomura T, Imaizumi-Anraku H et al (2001) Providing the basis for genomics in *Lotus japonicus*: the accessions Miyakojima and Gifu are appropriate crossing partners for genetic analyses. Mol Genet Genomics 266(2):157–166

Kawaguchi M, Imaizumi-Anraku H, Koiwa H et al (2002) Root, root hair, and symbiotic mutants of the model legume *Lotus japonicus*. Mol Plant Microbe Interact 15:17–26

Kawaguchi M, Andrea PH, Koji Y et al (2005) *Lotus burttii* takes a position of the third corner in the *lotus* molecular genetics triangle. DNA Res 12:69–77

Kikuchi S, Satoh S, Nagata S et al (2003) Collection, mapping, and annotation of over 28,000 cDNA clones from *japonica* rice. Science 301:376–379

Kurata N, Satoh H, Kitano H et al (2010) NBRP, National Bioresource Project of Japan and plant bioresource management. Breed Sci 60:461–468

Maekawa T, Kusakabe M, Shimoda Y et al (2008) Polyubiquitin promoter-based binary vectors for over-expression and gene silencing in *Lotus japonicus*. Mol Plant Microb Interact 21:375–382

Matsumoto T, Tanaka T, Sakai H et al (2011) Comprehensive sequence analysis of 24,783 barley full-length cDNAs derived from 12 clone libraries. Plant Physiol 156:20–28

Nakagawa T, Kaku H, Shimoda Y et al (2011) From defense to symbiosis: limited alterations in the kinase domain of LysM receptor-like kinases are crucial for evolution of legume-*Rhizobium* symbiosis. Plant J 65:169–180

Nakamura Y, Kaneko T, Asamizu E et al (2002) Structural analysis of a *Lotus japonicus* genome. II. Sequence features and mapping of sixty-five TAC clones which cover the 6.5-mb regions of the genome. DNA Res 9:63–70

Novák K (2010) On the efficiency of legume supernodulating mutants. Ann Appl Biol 157:321–342

Ogihara Y, Mochida K, Kawaura K et al (2004) Construction of a full-length cDNA library from young spikelets of hexaploid wheat and its characterization by large-scale sequencing of expressed sequence tags. Genes Genet Syst 79:227–232

Pajuelo E, Stougaard J (2005) *Lotus japonicus* as a model system. In: Márquez AJ (ed) *Lotus japonicus* Handbook. Springer, Netherlands, pp 3–24

Popp C, Ott T (2011) Regulation of signal transduction and bacterial infection during root nodule symbiosis. Curr Opin Plant Biol 14:1–10

Sandal N, Krusell L, Radutoiu S et al (2002) A genetic linkage map of the model legume *Lotus japonicus* and strategies for fast mapping of new loci. Genetics 161:1673–1683

Sandal N, Petersen TR, Murray J et al (2009) Genetics of Symbiosis in *Lotus japonicus*: recombinant inbred lines, comparative genetic maps, and map position of 35 symbiotic loci. Mol Plant Microb Interact 19:80–91

Sandal N, Jin H, Rodriguez-Navarro DN et al (2012) A set of *Lotus japonicus* Gifu × *Lotus burttii* recombinant inbred lines facilitates map-based cloning and QTL mapping. DNA Res 19:317–323

Sato S, Kaneko T, Nakamura Y et al (2001) Structural analysis of a *Lotus japonicus* genome. I. Sequence features and mapping of fifty-six TAC clones which cover the 5.4 Mb regions of the genome. DNA Res 8:311–318

Sato S, Nakamura Y, Asamizu E et al (2007) Genome sequencing and genome resources in model legumes. Plant Physiol 144:588–593

Sato S, Nakamura Y, Kaneko T et al (2008) Genome structure of the legume, *Lotus japonicus*. DNA Res 15:227–239

Seki M, Carninci P, Nishiyama Y et al (1998) High-efficiency cloning of *Arabidopsis* full-length cDNA by biotinylated CAP trapper. Plant J. 15:707–720

Shimoda Y, Mitsui H, Kamimatsuse H et al (2008) Construction of signature-tagged mutant library in

Mesorhizobium loti as a powerful tool for functional genomics. DNA Res 15:297–308

Soderlund C, Descour A, Kudrna D et al (2009) Sequencing, mapping, and analysis of 27,455 maize full-length cDNAs. PLoS Genet 5:e1000740

Stougaard J, Beuselink PR (1996) Registration of GIFU B-129-S9 *Lotus japonicus* germplasm. Crop Sci 36:476

Suginobu K, Suzuki S, Komatsu T (1988) Evaluation of the characteristics in Miyakogusa (*Lotus corniculatus* L. var. *japonicus* Regel) 3. Characteristics of local strains collected from different regions in Japan. J Jpn Grassl Sci 34:13–19 (in Japanese with English abstract)

Suzuki A, Suriyagoda L, Shigeyama T et al (2011) *Lotus japonicus* nodulation is photomorphogenetically controlled by sensing the red/far red (R/FR) ratio through jasmonic acid (JA) signaling. Proc Natl Acad Sci USA 108:16837–16842

Sz-Borso O, Somaroo BH, Grant WF (1972) A new diploid species of *Lotus japonicus* (Leguminosae) in Pakistan. Can J Bot 50:1865–1870

Szczyglowski K, Shaw RS, Wopereis J et al (1998) Nodule organogenesis and symbiotic mutants of the model legume *Lotus japonicus*. Mol Plant Microbe Interact 11:684–697

Tanaka H, Toyama J, Hashiguchi M et al (2008) Transgenic superroots of *Lotus corniculatus* can be regenerated from superroot-derived leaves following *Agrobacterium*-mediated transformation. J Plant Physiol 165:1313–1316

Thapanapongworakul N, Nomura M, Dao TV et al (2010) 3-Phosphoglycerate dehydrogenase in *Mesorhizobioum loti* is essential for maintaining symbiotic nitrogen fixation of *Lotus japonicus* root nodules. Plant Soil 336:233–240

Tsubokura Y, Onda R, Sato S et al (2008) Characterization of soybean genome based on synteny analysis with *Lotus japonicus*. Breed Sci 58:157–167

Umezawa T, Sakurai T, Totoki Y et al (2008) Sequencing and analysis of approximately 40,000 Soybean cDNA clones from a full-length-enriched cDNA library. DNA Res 15:333–346

Wang X, Sato S, Tabata S et al (2008) A high-density linkage map of *Lotus japonicus* based on AFLP and SSR markers. DNA Res 15:323–332

Weigel D, Ahn JH, Blázquez MA et al (2000) Activation tagging in *Arabidopsis*. Plant Physiol 122:1003–1013

Yamazaki Y, Akashi R, Banno Y et al (2010) NBRP databases: databases of biological resources in Japan. Nucleic Acids Res 38:D26–D32

Legume and *Lotus japonicus* Databases 23

Hideki Hirakawa, Terry Mun, Shusei Sato, and Stig U. Andersen

Abstract

Since the genome sequence of *Lotus japonicus*, a model plant of family Fabaceae, was determined in 2008 (Sato et al. 2008), the genomes of other members of the Fabaceae family, soybean (*Glycine max*) (Schmutz et al. 2010) and *Medicago truncatula* (Young et al. 2011), have been sequenced. In this section, we introduce representative, publicly accessible online resources related to plant materials, integrated databases containing legume genome information, and databases for genome sequence and derived marker information of legume species including *L. japonicus*.

23.1 Database for Plant Materials

Legume Base (http://www.legumebase.brc.miyazaki-u.ac.jp/top.jsp), a *L. japonicus* and *G. max* database, is developed by the National Bio-Resource Project (NBRP) (http://www.nbrp.jp) for distribution of plant materials. Users can order seeds of *L. japonicus* and *G. max*, bulked M2 seeds of *L. japonicus*, DNA, vector and rhizobium, root culture systems (superroot), and EMS M2 bulked seeds of *G. max* as plant materials. The biological materials pertaining to *L. japonicus* hitherto available through Legume Base (http://www.legumebase.brc.miyazaki-u.ac.jp/lotus/top/top.jsp) are: three experimental strains, 108 wild accessions, 205 LjMG recombinant inbred lines (RILs, those of crosses between Miyakojima MG-20 and Gifu B-129), 171 mutant lines (EMS-treated mutants), 6,671 *Mesorhizobium loti* STM mutant clones, 10,839 BAC clones of Miyakojima MG-20, 16,656 TAC clones of Miyakojima MG-20, 92,389 cDNA clones (Miyakojima MG-20 and Gifu B-129), endogenous long-terminal repeat (LTR) retrotransposon, and Lotus Retrotransposon 1 (LORE1) insertion tag lines. Meanwhile, biological materials related to *G. max* (http://www.legumebase.brc.miyazaki-u.ac.jp/glycine/top/top.jsp) are: 265 cultivars, 715 wild accessions,

H. Hirakawa (✉) · S. Sato
Kazusa DNA Research Institute, Kazusa-Kamatari 2-6-7, Kisarazu, Chiba 292-0818, Japan
e-mail: hh@kazusa.or.jp

T. Mun · S.U. Andersen
Centre for Carbohydrate Recognition and Signalling, Department of Molecular Biology and Genetics, Aarhus University, Gustav Wieds Vej 10, 8000 Aarhus C, Denmark
e-mail: sua@mb.au.dk

S. Sato
Graduate School of Life Sciences, Tohoku University, Katahira 2-1-1, Aoba-Ku, Sendai 980-8577, Japan
e-mail: shuseis@ige.tohoku.ac.jp

S. Tabata and J. Stougaard (eds.), *The Lotus japonicus Genome*, Compendium of Plant Genomes,
DOI 10.1007/978-3-662-44270-8_23,

164 RILs (those of crosses between Misuzudaizu and Mashidou Gong 503), 96 RILs (MxS lines, those of crosses between TK780 and B01167), 21 mutants, 39 strains of Edamame, and 37,890 clones of full-length cDNA.

23.2 Integrated Databases

23.2.1 PlantGDB

PlantGDB (http://www.plantgdb.org) is an integrated database for plant genomics, in which 16 dicots and 7 monocots have been registered. For Fabaceae, the genome and gene models of *L. japonicus, M. truncatula,* and *G. max*, spliced alignment of EST, cDNA, PUT (assembled unique transcript) against genome sequences are available in LjGDB, MtGDB, and GmGDB.

23.2.2 The Gene Index Project

The Gene Index Project (http://compbio.dfci.harvard.edu/tgi/) encompasses tentative consensus (TC) contigs assembled from expressed sequence tag (EST) data for plants, animals, protest, and fungi. For plants, EST and TC sequences have been provided for 60 species. As for Fabaceae, bean (*Phaseolus vulgaris*), cowpea (*Vigna unguiculata*), *L. japonicus*, *M. truncatula*, scarlet bean (*Phaseolus coccineus*), and soybean (*G. max*) have been registered in the database.

23.2.3 Phytozome

Phytozome (http://www.phytozome.net) contains releases of the nodes, clusters and consensus sequences, and annotation of transcripts for 41 plant species. The source data for the database were constructed using the BioMart framework (http://www.biomart.org). The source data were subsequently integrated with physical map positions and displayed using GBrowse (http://gmod.org/wiki/GBrowse). For Fabaceae, bean (*P. vulgaris*), *M. truncatula* and soybean (*G. max*) have been registered in the database.

23.2.4 PGDBj (Plant Genome DataBase Japan)

PGDBj (http://pgdbj.jp/?ln=en) is a portal site integrating the databases related to plant omics studies (Asamizu et al. 2014). The information related to DNA markers and QTL for 55 plant species has been collected from the literature by manual curation. Information for the following leguminous species has been collected—*L. japonicus* (1,073 SSR markers and 82 dCAPS markers), *G. max* (7,020 SSR markers), *M. truncatula* (65 SNP markers, 544 SSR markers, 246 QTL, 827 other markers (CAPS, RFLP, etc.)), red clover (*Trifolium pratense*) (7,468 SSR markers, 314 other markers (RFLP etc.)), and white clover (*T. repens*) (1,993 SSR markers). PGDBj contains cross-links to other databases, such as SABRE2 (http://sabre.epd.brc.riken.jp/SABRE2.html) for plant resources of *L. japonicus* and *G. max*, and KNApSAcK (http://kanaya.naist.jp/KNApSAcK/) for species-metabolite relationships in *L. japonicus*, *G. max*, *M. truncatula*, *T. pratense*, and *T. repens*.

23.3 Database for Legume Genome Information

23.3.1 Soybean (*G. max*)

SoyBase (http://www.soybase.org) was established to publish the integrated information on the genetics, genome information, and molecular biology of soybean. The database contains ontologies, metabolic pathways (SoyCyc), and microarray data (SoyChip developed by Affymetrix®). Genetic maps and physical maps can be browsed by CMap and GBrowse from the menus, "Maps" and "Genome," respectively. The database also contains genomic sequence, gene predictions, potential SSRs, and ESTs. On the

"mutant populations" page, mutants can be browsed by sample name, trait, image, and phenotype.

23.3.2 *Medicago truncatula*

The Medicago HAPMAP project (http://www.medicagohapmap.org) was spearheaded by an internal consortium. The project involves resequencing 384 inbred lines of *M. truncatula* with the purpose of discovering single-nucleotide polymorphisms (SNPs), insertions/deletions (INDELs), and copy number variants (CNVs). This information is useful for describing the population structure, identifying genome segments, and genome-wide association (GWA) mapping. The 338 germplasms can be ordered by users of the database. The genome structure of Mt3.5 and Mt3.0 can be browsed by GBrowse. The SNP data of Mt4.0, Mt3.5 and Mt3.0 can be obtained from the download page. The updated *M. truncatula* genome information is currently available from J. Craig Venter Institute (http://jcvi.org/medicago/).

23.3.3 *Lotus japonicus*

23.3.3.1 Miyakogusa.jp

Miyakogusa.jp (http://www.kazusa.or.jp/lotus/) was established in 2008 to facilitate publication of *L. japonicus* genomic information. The top page of the database is shown in Fig. 23.1. The page is divided into five sections: (1) Menus, (2) Links, (3) News, (4) Keyword search for the predicted genes, and (5) Genome browser. At the time of writing, the website is based on version 2.5 of the genome. An earlier release of the Lotus genome, version 1.0, is also available. The database and website will be updated to version 3.0 on its release. The database contains links to the related databases, such as "Legume Base" at Miyazaki University (http://www.legumebase.brc.miyazaki-u.ac.jp) and "Worldwide Lotus Science Information Desk" at the National BioResource Project (NBRP) (http://www.shigen.nig.ac.jp/infodesk/topSpeciesAction.do?speciesId=4). Users can access the information related to *Mesorhizobium loti*, the symbiont of *L. japonicus*, released from Rhizobase (http://genome.microbedb.jp/rhizobase/Mesorhizobium/), expression sequence tags (ESTs) sequenced by Kazusa DNA Research Institute (KDRI) (http://est.kazusa.or.jp/en/plant/lotus/EST/), transcript sequences released from "The Gene Index Projects" performed by Harvard University (http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/gimain.pl?gudb=l_japonicus), and the chloroplast genome sequence obtained by KDRI (http://www.kazusa.or.jp/lotus/Cp/). The menus on the top page are comprised of "Clone list," "Genetic map," "Insertion tag lines," "BLAST," "Database download," and "LjMG RI lines." The contents in each menu are shown as follows.

Clone List

The clones used for genome sequencing were anchored by genetic markers and physical linkage information, and comprised six linkage groups (chromosomes 1–6) and unanchored clones. The linkage maps can be browsed on the top page of "Clone list" as shown in Fig. 23.2. Users can find the clone from the list shown on the page or by using the search function to search by clone name. In the list, the assembly phase and sequence redundancy of each clone are also listed. In the version 2.5, 291 (SSR: 270, dCAPS: 21), 201 (SSR: 187, dCAPS: 14), 204 (SSR: 192, dCAPS: 12), 186 (SSR: 169, dCAPS: 17), 167 (SSR: 154, dCAPS: 13), and 106 (SSR: 101, dCAPS: 5) markers are located on 95, 63, 72, 58, 61, 44 contigs consisting of 540, 363, 370, 383, 321, 210 clones from chromosomes 1–6, respectively. The 174 clones on 9 contigs are not integrated into the linkage groups and are collected as unmapped clones. By clicking the marker name, users can access marker category information (SSR or dCAPS), primer sequences (Fw/Rv), EST or genome sequences, PIC values, position of markers on linkage map (Miyakojima MG-20 × Gifu B129), single-nucleotide polymorphisms (SNPs), simple sequence repeats (SSRs)

Fig. 23.1 Top page of miyakogusa.jp. Users can browse the contents from (1) Menus, (2) Links, (3) News, (4) Keyword search, (5) Genome browser

(fragment size, motif pattern, number of motif repeat), gel images, and enzyme-related markers.

Genetic Map

The location of contigs ordered by markers on each chromosome can be browsed at the top of the "Genetic map" page shown in Fig. 23.3. Users can access the marker list consisting of marker names, linkage groups, and genetic positions. The page also has clickable maps for each chromosome. As shown in Fig. 23.4, the detailed information on markers described in the "Clone list" section is displayed at the right side of the linkage map.

Insertion Tag Lines

The gene-tagging populations were established by using an endogenous long-terminal repeat (LTR) retrotransposon, Lotus Retrotransposon 1 (*LORE1*), in Denmark and Japan. The list of LORE1 tag lines of 32,272 (300001–30057983) and 1,227 (P001–P2550), established in Denmark and Japan, respectively (Fukai et al. 2012; Urbanski et al. 2012), is available on the top page of "Insertion tag lines" (http://www.kazusa.or.jp/lotus/LORE1/). In each plant line, the inserted positions related to genes and functions predicted by homology searches against TrEMBL (http://www.ebi.ac.uk/uniprot) were summarized. Through this page (example for plant line P0001: http://www.kazusa.or.jp/lotus/predict/cgi-bin/LORE1_lineinfo.cgi?id=P0001), users can order plant lines from Legume Base developed by National BioResource Project (NBRP). Users can search the query sequence against LORE1 insertion sites by using BLASTN (Altschul et al. 1997) on the top page of "Insertion tag lines."

BLAST Search

On the BLAST search page (http://www.kazusa.or.jp/lotus/blast.html), users can perform searches against all predicted genes and clones (contigs) of versions 2.5 and 1.0, and also against LORE1 insertion sites. This page was based on the Web BLAST interface distributed at NCBI (http://www.ncbi.nlm.nih.gov).

Data Download

The genome contigs and pseudomolecules, gene, and protein sequences can be downloaded from this page, and gene model and assembly were

Index

Clone list

Genetic map

Insertion tag lines

BLAST

Data download

LjMG RI lines

Links

Legume Base

Worldwide Lotus Science Information Desk

by National BioResource Project

Symbiont

Mesorhizobium Loti genome database

Transcripts

EST information

L. japonicus Gene Index

Clone Order List

View chromosome: 1|2|3|4|5|6|unmapped|

Keyword Search: SUBMIT

Chromosome 1

marker	marker type	chr	cM	contig name	clone name	phase	sequence redundancy
		1	-	CM0088	LjT12M01	1	x3
TM0088	SSR	1	0.0	CM0088	LjT15K21	1	x5
		1	-	CM0088	LjB26M21	1	x3
TM0224	dCAPS	1	0.0	CM0088	LjT13A11	3	x5
TM0358	SSR	1	0.0	CM0088	LjT48G05	3	x5
BM2431	SSR	1	0.0	CM0088	LjB04N23	3	x5
TM1703	SSR	1	0.0	CM0591	LjT48M07	3	x5
		1	-	CM0591	LjT21K10	3	x5
TM0771	SSR	1	0.0	CM0591	LjT13P11	1	x3
TM0591	SSR	1	0.0	CM0591	LjT01F11	1	x5
TM1473	SSR	1	0.0	CM0591	LjT39D07	1	x3
TM1049	SSR	1	0.0	CM0289	LjT39O15	1	x3
TM0289	SSR	1	0.0	CM0289	LjT16P18	3	x5
		1	-	CM0289	LjT45I09	1	x3
TM0454	SSR	1	0.0		LjT29N14	1	x3

Fig. 23.2 Top page of "Clone list." Users can search the clones by (1) Keyword search or (2) Clone list in each chromosome

distributed as gff3 format for release of versions 2.5 and 1.0. As for LjMG recombinant inbred lines developed by crossing between Miyakojima MG-20 and Gifu B129, the genotype data in 96 loci (chr 1: 20, chr 2: 15, chr 3: 17, chr 4: 15, chr 5: 15, chr 6: 14) for 205 lines are currently summarized in an MS-Excel file (RI_line_genotype_140325.xls). These data can be downloaded from the ftp site (ftp://ftp.kazusa.or.jp/pub/lotus/).

23.3.3.2 Lotus Base

Lotus Base (http://carb.au.dk/lotus-base/) was initially established by the Centre for Carbohydrate Recognition and Signalling (CARB) based at Aarhus University, Denmark, in 2011, as a publicly accessible online resource dedicated to *L. japonicus* plant lines derived from a single founder line in which LORE1 was activated (Madsen et al. 2005; Fukai et al. 2012; Urbański

Fig. 23.3 Top page of "Genetic map." Users can browse the genetic maps from (1) table or (2) figure of each chromosome

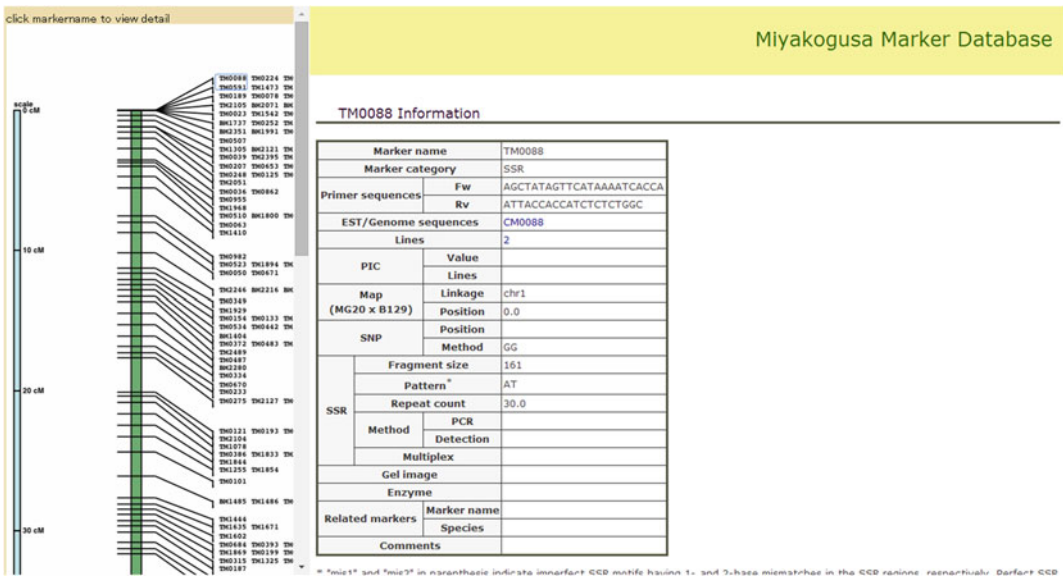


Fig. 23.4 Linkage map and marker information. The detail information of markers is available from the linkage maps

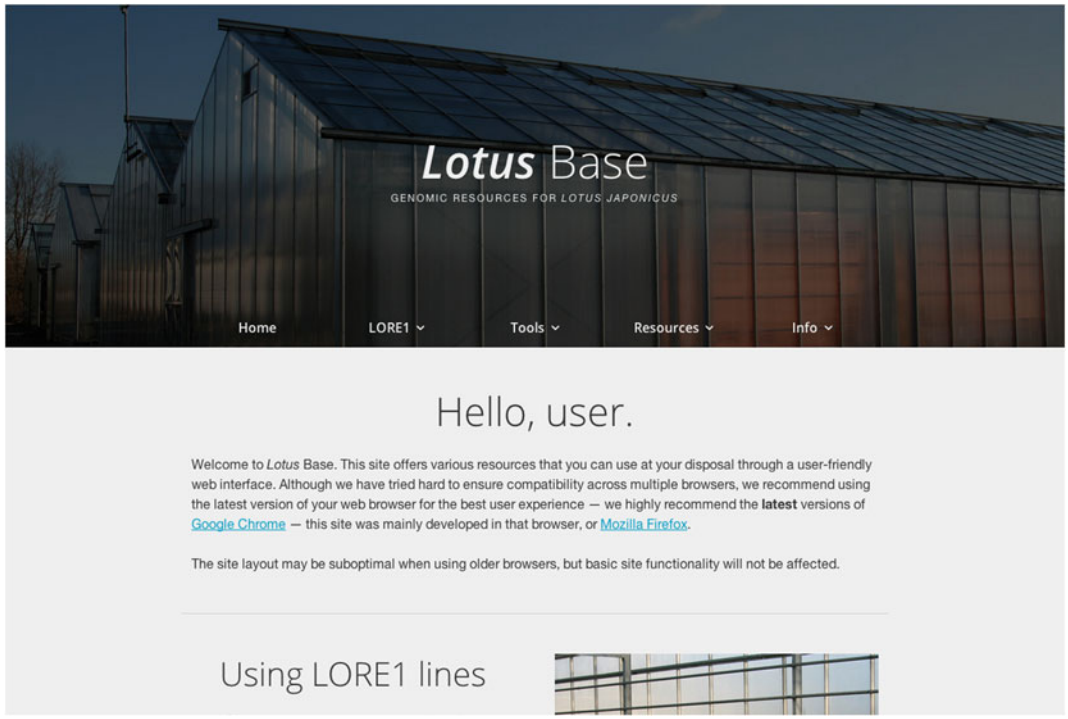


Fig. 23.5 Landing page of *Lotus* Base (http://carb.au.dk/lotus-base/). The main tools useful to the end users are grouped under navigation categories: (1) "LORE1" contains resources pertaining to LORE1 insertional mutagenesis lines, such as LORE1 line search, line order status and order form; (2) "Tools" is a collection of usability tools designed for the resource, including an in-house modified version of NCBI wwwblast suite; (3) "Resources" contains downloadable manuals, scripts and programs related to the project; and (4) "Info" hosts a repertoire of commonly asked questions on the use, application, genotyping and general information of LORE1 lines, and usability tools

et al. 2012). The resource was subsequently expanded and redesigned into a general *L. japonicus* genomics resource known as Lotus Base.

All resources available through Lotus Base are based on *L. japonicus* genome version 2.5, with version 3.0 release pending in the near future.

LORE1 Database and Line Ordering

At the time of writing, the database contained 96,905 LORE1 insertions registered in 32,272 unique plant lines (300001–30057983). Each insertion is tagged with a unique BLAST header containing the chromosome number, chromosomal position, and insertion orientation, allowing these insertions to be accessible from BLAST searches against the LORE1 insertions database. Users can retrieve plant lines of interest from the search page (http://users-mb.au.dk/pmgrp/lore1search.php) based on user-defined criteria: BLAST header, plant line/ID, chromosome, chromosomal position, and gene name or annotation, or a combination thereof (Fig. 23.6). The search results will return appropriate plant lines, each complete with information cross-referenced from database tables including gene name, annotation, forward and reverse genotyping primers, expected PCR fragment sizes of wild type and insertional mutants, and the ±1 kb flanking sequence surrounding the LORE1 insertion site. The search form is accessible via "LORE1 > LORE1 lines seach" [Fig. 23.5(1)].

From the results page, users are able to further refine their search based on alternative plant line (s) or gene(s) of interest, and access gene

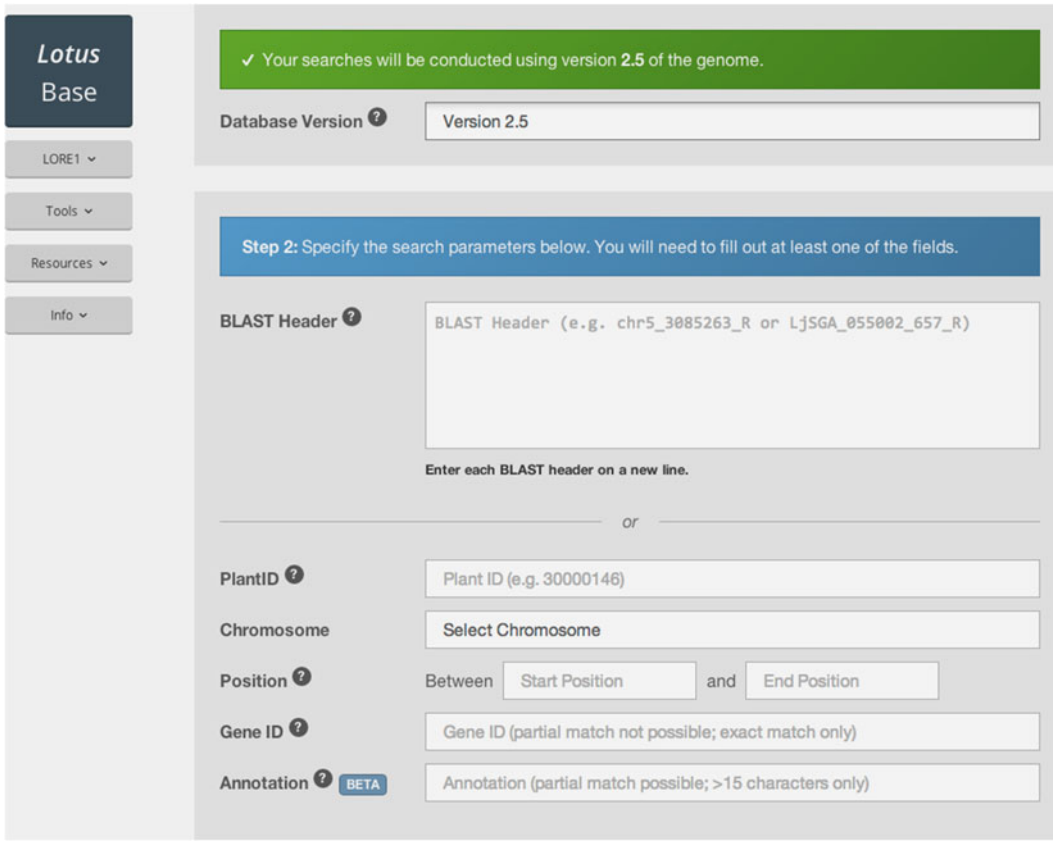


Fig. 23.6 The full set of search options made available to the end user on the LORE1 line search page, accessible via the LORE1 link [Fig. 23.5(1)]. The end user will be (1) prompted to select the appropriate database before being presented with either (2) the option to search lines by BLAST headers retrieved from a BLAST search against LORE1 flanking sequence database(s), or (3) the option to search LORE1 lines by plant ID, chromosome, position, gene ID or annotation, or a combination thereof

annotations (if available) associated with each plant line. The gene annotation table was constructed by performing BLAST search using ab initio gene predictions against NCBI nr database (http://www.ncbi.nlm.nih.gov) and filtering away statistically insignificant matches.

In order to access the full information fetched for individual plant line of interest after searches have been performed, users are recommended to download the search results, which contains the full column set of returned database results encoded in spreadsheet-program friendly file formats.

Once users have identified LORE1 lines of interest, they can place orders on these lines through the order page (http://users-mb.au.dk/pmgrp/order.php). Each order is tagged with a unique, randomly generated 32-character hexadecimal-order identification code that allows easy yet secure tracking of individual orders. Users may also access the order history of individual lines (http://users-mb.au.dk/pmgrp/order-search.php), but will not be able to search specifically for individuals who have placed orders through our resource owing to privacy restrictions.

Lotus BLAST

The Lotus basic local alignment search tool [BLAST, available at http://users-mb.au.dk/pmgrp/blast/, accessible via "Tools," Fig. 23.5 (2)] can be used to perform searches against LORE1 flanking sequences and the *L. japonicus* genome version 2.5. The tool is constructed based on the legacy wwwblast tool suite by NCBI (http://www.ncbi.nlm.nih.gov/staff/tao/

URLAPI/wwwblast/). BLAST searches using amino acid or nucleotide sequences against the LORE1 flanking sequences enables users to search for genes of interest for which our resource has insertional mutants available. Results returned from this search are tagged with a unique BLAST header, which users can subsequently query the LORE1 database with.

Usability Tools

To aid the user, we have included some tools to improve the usability of the resource [accessible via "Tools," Fig. 23.5(2)]. The sequence retrieval tool (SeqRet), based on the wwwfastacmd project and whose modified source code is publicly available (https://github.com/terrymun/wwwfastacmd), allows users to retrieve full sequences of queried identifying information, such as BLAST headers, against relevant databases. Meanwhile, the sequence processor tool (SeqPro) uses RegEx parsing to clean up data provided by users, such as large amounts of formatted information from legacy NCBI blast output, allowing for easy extraction of relevant data.

A list of commonly asked questions (http://users-mb.au.dk/pmgrp/faq.php) is provided to guide users who may have queries that have been previously addressed [accessible via "Tools > FAQ," Fig. 23.5(3)].

References

Altschul SF, Madden TL, Schäffer AA et al (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402

Asamizu E, Ichihara H, Nakaya A, Nakamura Y, Hirakawa H, Ishii T, Tamura T, Fukami-Kobayashi K, Nakajima Y, Tabata S (2014) Plant Genome DataBase Japan (PGDBj): a portal website for the integration of plant genome-related databases. Plant Cell Physiol 55:e8

Fukai E, Soyano T, Umehara Y, Nakayama S, Hirakawa H, Tabata S, Sato S, Hayashi M (2012) Establishment of a Lotus japonicus gene tagging population using the exon-targeting endogenous retrotransposon LORE1. Plant J 69:720–730

Madsen LH, Fukai E, Radutoiu S, Yost CK, Sandal N, Schauser L, Stougaard J (2005) LORE1, an active low-copy-number TY3-gypsy retrotransposon family in the model legume Lotus japonicus. Plant J 44(3):372–381

Sato S, Nakamura Y, Kaneko T et al (2008) Genome structure of the legume, Lotus japonicus. DNA Res 15:227–239

Schmutz J, Cannon SB, Schlueter J et al (2010) Genome sequence of the palaeopolyploid soybean. Nature 463:178–183

Urbanski DF, Malolepszy A, Stougaard J, Andersen SU (2012) Genome-wide LORE1 retrotransposon mutagenesis and high-throughput insertion detection in Lotus japonicus. Plant J 69:731–741

Young ND, Debellé F, Oldroyd GE et al (2011) The Medicago genome provides insight into the evolution of rhizobial symbioses. Nature 480:520–524