Plant translational genomics: from model species to crops

**Elma M. J. Salentijn · Andy Pereira ·
Gerco C. Angenent · C. Gerard van der Linden ·
Frans Krens · Marinus J. M. Smulders ·
Ben Vosman**

Abstract Plant genomic research now faces the ultimate challenge to develop applications in crop plants which implies the translation of gene functions from a model to a crop which is the field of 'Plant translational genomics'. In this paper we discuss the perspectives of the candidate gene approach (CGA) as a tool for translational genomics in the 'whole genome' era. Factors to be considered for a successful application of the CGA in crops such as the type of crop, the complexity of the trait and the type of genes involved are discussed. Several crop traits that require improvement such as tolerance to stress, pod shatter in *Brassicaceae* and Fusarium resistance, are evaluated with regard to the potential of a CGA as a tool for crop improvement

Keywords Candidate gene approach ·
Comparative genomics · Crop improvement ·
Pathogen resistance · Pod shatter ·
Stress tolerance

E. M. J. Salentijn (✉) · A. Pereira ·
G. C. Angenent · C. G. van der Linden ·
F. Krens · M. J. M. Smulders · B. Vosman
Plant Research International, Droevendaalse steeg 1,
6708PB Wageningen, The Netherlands
e-mail: elma.salentijn@wur.nl

Introduction

The genomics revolution, heralded by the sequencing of model genomes and supported by newly developed high throughput gene characterization and function analysis technologies, now faces the ultimate challenge to provide applications for crop improvement which is the field of ''Plant translational genomics'' (Gepts et al. 2005; Stacey and VandenBosch 2005). The most important traits in this respect, such as biotic and abiotic stress tolerance, plant development and consumer quality aspects, are genetically and physiologically complex. Moreover, because of the polyploid nature of many crops, breeding for such traits is time consuming and difficult. The quickly expanding knowledge on gene function and the availability of whole genome sequences of plants such as Arabidopsis (The Arabidopsis genome initiative, 2000), rice (The International Rice Genome Sequencing Program; Yu Jun et al. 2002; Goff et al. 2002) and poplar (Tuskan et al. 2006), soon to be followed by many others (see NCBI Entrez Genome Project database), is expected to offer new perspectives to solve these complex problems in crop species as well.

The most promising tool for quick implementation of this knowledge is the candidate gene approach (CGA) (Byrne and McMullen 1996; Pflieger et al. 2001). The CGA is based on the

assumption that genes with a proven or predicted function in a 'model' species (functional candidate genes) or genes that are co-localized with a trait-locus (positional candidate genes) could control a similar function or trait in an arbitrary crop of interest (target crop). As such the CGA has been often validated in crop improvement, for instance the 'Green Revolution' dwarfing gene *Rht* of wheat is orthologous to genes conferring dwarf mutants, *Dwarf 8* in maize and GAI in Arabidopsis (Peng et al. 1999).

The pre-requisite for a CGA is a repertoire of well characterized candidate genes (CGs) for a trait. The primary way to select functionally characterized CGs (functional CGs, Fig. 1A) is to examine phenotypic, biochemical and physiological information on genes acting in the pathway of interest if this information is available. A quickly expanding amount of functional genomics information can be obtained from integrated databases such as genomic sequence data, literature, expression profiles, cellular localization of the corresponding protein, protein interactions, metabolic changes, mutant phenotypes and information from genetically modified organisms (GMO) (Bro and Nielsen 2004; Meyers et al. 2004). Furthermore, a novel and high throughput approach towards functional analysis, termed targeting induced local lesions in genomes (TILLING; McCallum et al. 2000), offers the possibility to select an allelic series of mutations for a specific gene creating an unique source for gene function analysis. If functional CGs for a trait are not known, co-location of CG polymorphisms with map positions, linkage to quantitative trait loci (QTL), association of alleles with specific traits or the identification of syntenic regions among genomes can help to select positional CGs for the trait (Positional CGs, Fig. 1A). After the selection of a set of functional-and/or positional CGs in a 'model' species, the CGs have to be translated to the 'target' crop. For this, functional orthologous genes (genes derived from a common ancestor through a speciation event) have to be identified in the 'target' crop (Fig. 1B). Finally, the CGs have to be applied and thereby validated in the target crop to result in the ultimate products of CGA: a crop with a desired trait, or in a marker that can be used for breeding. This is also the ultimate goal of 'Plant translational genomics', the application of gene functions from a 'model' to a crop (Gepts et al. 2005; Stacey and VandenBosch, 2005). Depending on the genomic organization of the CG, the complexity of the target genome and the nature of the biological function, different methods can be used for successful application of the CGA (Fig. 1C).

Benefits and limitations of a CGA

The identification of CGs is a prerequisite for a CGA. Previously, the identification of CGs underlying a genomic region linked to a trait (for instance a QTL region) involved laborious fine mapping studies and genetic complementation studies. Now, with the availability of whole genome sequence information in 'model' dicot and monocot species, the genes present within a QTL region can be selected based on genomic synteny and on putative function of the genes present in that area. With respect to this, the 'model' genome is sequenced and can be aligned with comparable marker sequences in a 'target' genome to deduce the putative CGs present in that genomic region. To investigate the possibility for comparative mapping of CGs across distant related species, comparison of the complete genomes of Arabidopsis and rice (Jaiswal et al. 2006) was undertaken but very little conservation of genomic organization (synteny) could be detected (Yu Jun et al. 2002; Goff et al. 2002; Devos et al. 1999; Huan SanWen et al. 2005). Within plant families the different genomes are often colinear which might offer the possibility to identify orthologous CGs on basis of syntenic genomic regions (i.e. Krutovsky et al. 2004). For instance, the tomato and potato genomes, both belonging to the *Solanaceae* plant family, are remarkably colinear and differ only by five paracentric inversions allowing the identification of CGs across species (Thabuis et al. 2003; Huan SanWen et al. 2005). On the other hand, caution should be taken with this approach since genomic synteny is not always reflecting a perfect colinearity. Examples of this are found in the *Brassicaceae* plant family (Yang et al. 2006, Town et al. 2006) and in the maize genome (e.g. Lai et al.
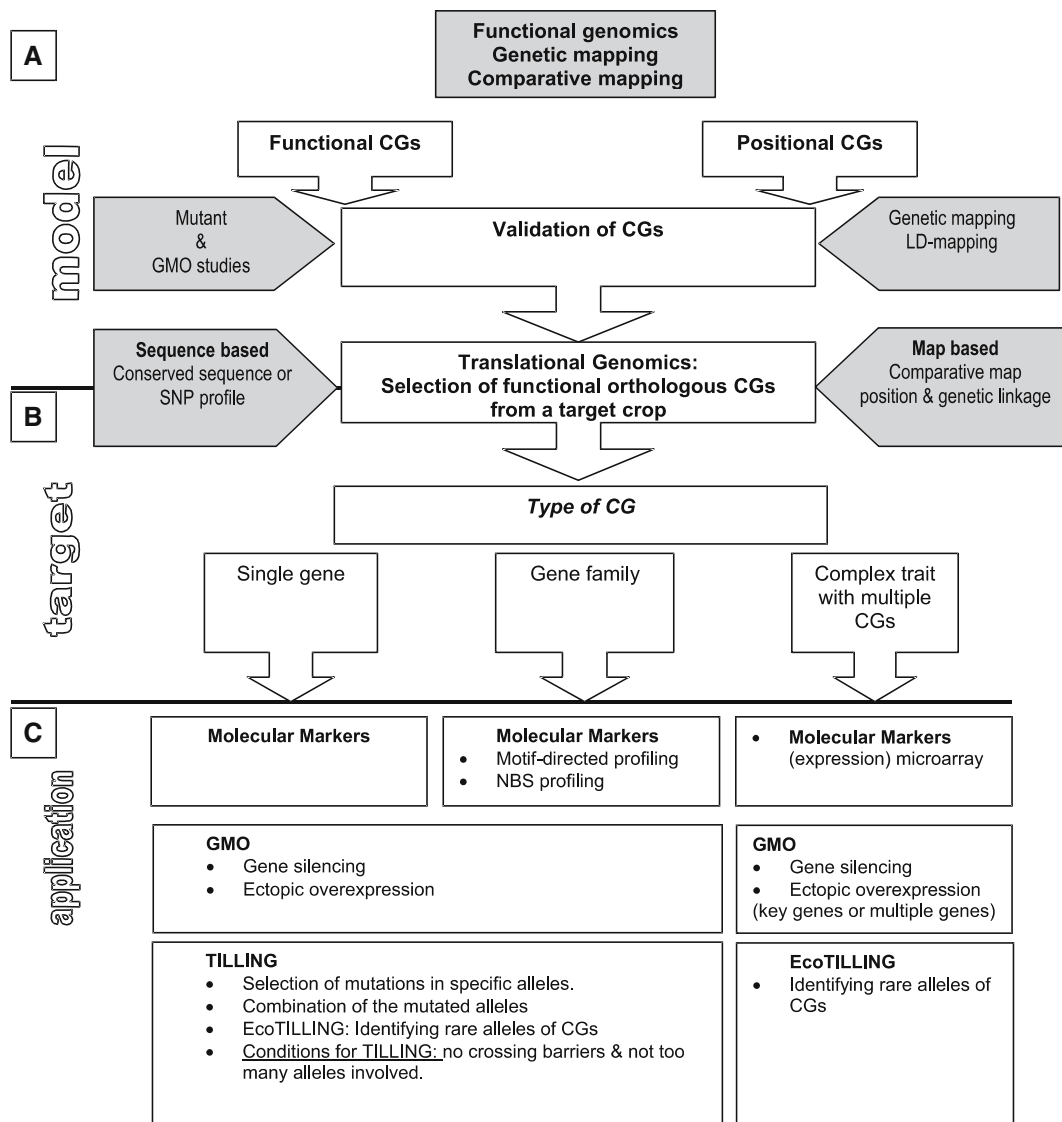
Fig. 1 Route towards the application of CGs: Starting with the identification of CGs in a model system (**A**) using functional genomics information (functional CGs) and the information of genomic mapping studies (positional CGs) the CGs are further validated in the 'model' by genetic-mapping, LD-mapping, expression studies and comple-mentation studies. Based on these CGs, orthologs are isolated from a target crop (**B**) using genome wide and sequence based comparative studies. Finally, the CGs are applied (and validated) in the target crop (**C**). Several possible methods for the application are considered with respect to the complexity of the trait, crop and the CG. The efficiency and usefulness of these methods is not guaranteed but may vary with the complexity of the trait, type of CG and complexity of the target crop genome

2006) in which extensive rearrangements and duplications have taken place during evolution that disturb the colinearity of the genomes and allow the loss of genes, consequently hampering the comparative mapping of CGs. Within a species, divergence into different haplotypes in most cases does not affect the colinearity of genes. However, there are also examples were the conserved order of genes changed, for example among breeding lines of a maize (e.g. Song and Messing 2003; Brunner et al. 2005). But even in situations of perfect genomic synteny, QTL regions appear often quite complex and approximate and may contain hundreds of genes.

Consequently, the actual involvement of the CG in most cases remains to be confirmed by genetic and physical mapping, positional cloning, expression analysis, or genetic transformation experiments. Fortunately, assistance in genetic mapping in 'models', comes from large sequencing projects. The huge expressed sequence tag (EST) databases that are generated by sequencing initiatives allow the 'in silico' identification of genetic variation such as single nucleotide polymorphisms (SNPs) and high density EST/SNP maps are becoming available now in human, Arabidopsis, rice and other organisms. This genetic variation can be linked to a trait of interest by genetic mapping or LD-mapping. With the present growth of EST resources such a targeted EST/SNP approach is becoming a powerful tool for a more accurate identification of relevant CGs (for ref. see Gutterson and Zhang 2004). Another advantage of high density SNP maps is that they increase the potential of linkage disequilibrium (LD)-mapping and association studies (Flint-Garcia et al. 2003, Feltus et al. 2004). Regarding the difficulties that maybe encountered, a CGA is always most powerful if combined with the available functional information for a trait from physiological studies, microarray expression analysis and studies of gene function via transgenics or mutants (e.g. Bro and Nielsen 2004).

Subsequently, after identification of CGs in a 'model', orthologous CGs have to be identified in the 'target' crop. Differential evolution is believed to be observed as sequence divergence, with the functionally essential genomic regions being more conserved than the non-essential ones. The alignment of whole genomes therefore should reveal the more conserved regions that potentially will prove to be of functional importance (Chervitz et al. 1998). In this way comparative genomic studies can take advantage from whole genome information and can provide information on the extrapolation of gene functions among species in relation to 'translational genomics' (e.g. Laurie et al. 2004; Stein 2004). The complete genomes of the multicellular nematode *Caenorhabditis elegans* and the unicellular yeast *Saccharomyces cerevisiae* were the first to be compared and it was shown that both organisms had a comparable number of orthologous proteins that carry out core functions including primary metabolism, protein folding, DNA and RNA metabolism, trafficking, and degradation. However, the more specialized functions that are unique to the worm, such as regulatory and signal transduction functions, are governed by proteins that have no orthologs in yeast even though they may contain domain sequences shared with yeast (Chervitz et al. 1998; Rubin et al. 2000). On an even smaller scale, the alignment of the nucleotide- or amino acid sequence of genes (e.g. in BLAST searches and subsequent sequence alignments) is an increasingly valuable starting point for the selection of CGs that share the same gene function (e.g. Brunner and Nilsson 2004). Studying the sequence variation among alleles (paralogs and orthologs) of CGs may provide conserved sequence motifs or conserved SNPs associated with a trait (Caicedo and Purugganan 2005). To assess the possibilities of extrapolation of CGs from Arabidopsis to the legumes Lupin (*Lupinus angustifolius*) and soybean, the sequences of several orthologs to lupin genes were compared. Conservation of gene structure and expression profiles suggested in some cases similar protein functions for the genes (Francki and Mullan 2004). Other examples of the extrapolation of gene function from a model crop to a more distant species are given by Laurie et al. (2004) and by Gutterson and Zhang (2004). However, similar biochemical pathways may have diverged during evolution creating a possible pitfall for the CGA. This can be exemplified by the difficulties encountered in the map based cloning of the *V*rn genes involved in vernalization requirements of wheat. Several CGs were proposed based on synteny and orthology to Arabidopsis genes involved in vernalization. Finally, detailed genetic and physical maps of a diploid wheat cultivar, combined with comparative fine-mapping studies of the colinear VRN1 and VRN2 regions in rice, sorghum and hexaploid wheat resulted in the identification of unexpected genes, respectively a homolog to APETALA1 (AP1) and another transcription factor, ZCCT1. Both genes were not among the CGs that were initially proposed for VRN1 and VRN2 on basis of comparison to Arabidopsis. Remarkably, the AP1 wheat homolog is associated with vernalization requirements in wheat but not in Arabidopsis

(Yan et al. 2003; Yan et al. 2004; Kato et al. 2002) suggesting a divergent evolution for this trait in both species resulting in a diversification of the genes involved.

Application and validation of CGs in a target crop, the ultimate goal of 'translational genomics', can be reached in several ways. A GMO approach can be applied if the function of a CG is well studied and pleiotropic effects are accounted for. However, acceptance of this elegant technique is still a subject of much discussion in Europe. In addition, a GMO approach is not feasible in all crops and cultivars. Alternatively, targeted mutagenesis (TILLING) can be applied to select induced mutations in functionally well-characterized CGs directly in the target crop (Slade et al. 2004). This is particularly useful to obtain mutants of specific members of a redundant gene family that are difficult to identify by forward genetic strategies. With this strategy an allelic series of mutations can be obtained with different effects on a trait. Limiting factors for this approach are the number of CGs involved in a trait and the additional breeding steps that are required to combine mutated alleles and purge the background mutations. In some outbreeding crops such additional breeding steps are difficult to perform without loosing the specific cultivar characteristics. TILLING further requires unique sequence motifs that allow the specific PCR amplification of the target allele if the CG belongs to a large gene family.

In case genetic variation is present in the CG, it can be used for genetic mapping or LD-mapping, to prove linkage to a trait in the 'target' crop. A growing amount of genetic variation in CGs in crops can directly be identified from databases (e.g. Rudd 2005) whereas EcoTILLING (Henikoff and Comai 2003) can be an approach to identify the more rare natural genetic variation within CGs, linked to a trait. For breeding purposes, providing its presence and linkage to a trait, such genetic variation can be converted into molecular markers which are the general product of a CGA (Feltus et al. 2004; Rudd et al. 2005; Rudd 2005). Based on conserved motifs and in combination with variable domains, CG-specific DNA profiles can be obtained that will reflect the genetic variation present in the different loci. In this way a whole gene family or subfamily can be assessed directly for associations with complex traits (QTL). As such, motifs in e.g. transcription factors (TFs) and pathogen resistance gene analogs (RGAs) are good candidates. In case of RGAs, nucleic binding site (NBS) -profiling is successful because the target group is clearly delimited by specific sequence motifs, because they are involved in a very specific process, and because they occur in clusters so that hitting the wrong gene in the right cluster also yields a useful marker (Linden van der et al. 2004; Calenge et al. 2005). In other cases the motif containing CG copies are more dispersed over the genome and functionally more diverse, so it remains to be seen how many in fact can be linked to interesting traits.

With regard to the biochemical nature of the corresponding phenotype, CGs may operate together in a coordinated way, upstream or downstream in biochemical pathways or at branch points between biochemical pathways and have either to be functionally present or absent to confer the desired trait. Furthermore, CGs may occur at single loci or as a part of a multigene family with functionally specialized or redundant alleles that are organized in clusters or dispersed over the genome. The type of crop, the complexity of the trait and the type of CGs involved therefore are important factors to be considered for a successful application of CGs in crops. Theoretically, several of these aspects can be encountered in a CGA towards the improvement of important crop traits such as tolerance to biotic stress, pathogen resistance and, pod shatter. We evaluate these possible examples of a CGA below in more detail for chance of success, as they represent complex traits that require different 'translational' approaches such as a gain of CG function, the 'translation' of a functionally specialized CG belonging to a gene family and a loss of CG function trait.

Tolerance to stress

Environmental stresses such as water, drought, heat or salt stress have adverse effects on plant growth and seed production and are limiting to

world food production. As the world food situation is expected to deteriorate in the near future, tolerance to such stresses is an important target for crop improvement (reviewed in Shinozaki et al. 2003). As a first reaction to external stress stimuli, genes coding for signal proteins and TFs are expressed (Shinozaki et al. 2003). TFs are considered as one of the major factors involved in the coordination of gene expression and in the fine-tuning of biochemical pathways. Intensive functional analysis is currently undertaken to identify TFs that control specific traits (e.g. Czechowski et al. 2004). In Arabidopsis, approximately 1,800 different TFs are known and comparative analysis among eukaryotes was performed (Riechmann et al. 2000). Despite the large number of TFs in a genome, it appears that genes coding for TFs can be recognized on basis of specific domains and motifs in their sequence and that the function of specific subfamilies of TFs are conserved among species (Gutterson and Reuber 2004; Marè et al. 2004; Reyes et al. 2004; Tian ChaoGuang et al. 2004; Dubouzet et al. 2003). The expression of stress responsive genes in Arabidopsis was shown to be controlled by the DREB/CBF type of AP2/ERF TFs and subsequently, on basis of conserved regions in the *DREB* genes from Arabidopsis, five *DREB* homologues of rice were isolated. Indeed, the overexpression of the Arabidopsis *DREB1A* gene both in Arabidopsis and in rice resulted in higher tolerance to drought, high salt and freezing (for references see Shinozaki et al. 2003; Oh-SeJun et al. 2005; Ito et al. 2006; Sakuma et al. 2006). Similarly, the overexpression of the rice *OsDREB1a* gene in transgenic Arabidopsis resulted in increased freezing and high-salt tolerance, showing that both orthologs can drive the same pathways and are, at least partly, functionally conserved among Arabidopsis and rice. However, as a side effect of *OsDREB1A* overexpression the transgenic plants exhibited significant growth retardation (Dubouzet et al. 2003; Ito et al. 2006). The use of a more specific, stress-inducible rd29A promoter instead of the constitutive 35S CaMV promoter to regulate the overexpression of *DREB1A* in transgenic Arabidopsis (Kasuga et al. 1999), tobacco (Kasuga et al. 2004) and wheat (Pellegrineschi et al. 2004)

minimized the negative effects of DREB expression on plant growth. The transgenic wheat lines obtained showed a 10-day delay in wilting upon water stress and otherwise exhibited normal plant growth. Another AP2/ERF-like transcription factor gene that induced drought tolerance upon overexpression is the *SHINE* (*SHN*) gene from Arabidopsis. Overexpression of *SHN* led to increased levels and altered composition of cuticular waxes and a reduced stomatal density (Aharoni et al. 2004). Both, *DREB1A* and *SHN* are promising CGs to accomplish drought and other abiotic stress tolerance in dicots and monocots via a GMO approach.

Pathogen resistance

Probably the most desired crop trait is resistance to plant pathogens. When not controlled chemically or biologically, pathogens may cause severe crop losses. In many cases disease resistance in plants is race-specific (vertical resistance) and determined by single dominant or semi-dominant resistance genes (*R*-genes) that are involved in the recognition of the products of avirulence (*avr*) genes from pathogens resulting in the activation of a plant defence response *R*-genes belong to large multigene families and *R*-genes acting against a broad range of pathogens including bacteria, virus, nematodes, and fungi and even to aphids have been cloned from different plant species (reviewed in e.g. Bent 1996; Hammond-Kosack and Jones 1997; Hulbert et al. 2001; Dangl and Jones 2001; He et al. 2004). Mapping studies revealed that the RGAs were often localized in clusters and near major QTL for resistance (e.g. Kanazin et al. 1996; Mago et al. 1999; Pan et al. 2000; Ramalingam et al. 2003; Linden van der et al. 2004). For this reason RGAs can also be considered as *R*-gene candidates (Pflieger et al. 2001). RGAs are present in both dicots and monocots and their action is often pathogen or even species (strain) specific. RGAs constitute about 0.6% of the genome in Arabidopsis whereas in rice more than 600 RGAs of the NBS-LRR class are present (The Arabidopsis Genome Initiative 2000; Goff et al. 2002). Functional specifications may have occurred after

the monocot–dicot divergence or even relatively recent in populations under attack by a particular pathogen (Bai et al. 2002).

An example of an important pathogen is the filamentous ascomycete of the genus *Fusarium*. The genus includes a number of economically important plant pathogenic species such as for instance *Fusarium oxysporum lycopersici* in tomato and *F. graminearum* and *F. culmorum*, the causative agents of head blight (scab) in cereals and grasses. Management of these pathogens is difficult due to their endophytic growth and persistence in soil, making genetic resistance to Fusarium a demanded alternative. Fusarium resistance in tomato and other *Solanaceae* species is race-cultivar specific (vertical resistance) whereas resistance to head blight in monocots is of a completely different type (horizontal resistance) governed by yet unknown genes (Eeuwijk et al. 1995, Mesterhazy et al. 1999, Paillard et al. 2004).

For the interaction between *F. oxysporum lycopersici* and tomato three host-specific races of *F. oxysporum lycopersici* have been described. The *I2* gene conferring resistance to race 2 was positionally cloned and is a typical *R*-gene containing coiled coil (CC)—nucleotide binding side (NBS)—leucine rich repeat (LRR) motifs (Simons et al. 1998). Until now this is the only isolated gene conferring resistance to *F. oxysporum* and as such may assist the identification of CGs conferring race- specific Fusarium resistance in other crops. The *I2* gene is situated on the long arm of chromosome 11 in a cluster of seven similar genes. To 'translate' Fusarium resistance governed by the *I2* gene from tomato to potato, the *I2* locus of tomato was compared to its syntenic region in potato, the *R3* locus. This comparison resulted in the isolation of the *R3a* late blight (*Phytophthora infestans*) resistance gene (Huan SanWen et al. 2005) but not in a potato gene conferring resistance to Fusarium. Another locus, *I3*, conferring resistance to race 3 in tomato, has been mapped on the long arm of chromosome 7 but the *I3* gene itself is not identified yet. Its syntenic region in potato harbors the *Gro1* gene, conferring resistance to the root cyst nematode. One orthologous tomato fragment to the *Gro1* locus in fact co-segregated with the *I3* resistance. However, this co-segregating marker was found to

be a putative pseudogene and was excluded as a candidate for *I3* (Hemming et al. 2004) leaving the identity of *I3* as yet unknown. It was suggested that tomato and potato R-genes, as in the case of the orthologous loci *R3a/I2* and *Gro1/I3* loci, have evolved from ancient loci conferring respectively resistance to oomycete and fungal pathogens (*R3a/I2*) (Huan SanWen et al. 2005) and to different soil born pathogens that enter their hosts through the vascular tissue of the root system (*Gro1/I3*) (Paal et al. 2004; Hemming et al. 2004). However, after the divergence of tomato and potato these loci may have evolved further resulting in a diversification of resistance genes based on co-evolution with the respective pathogens of tomato and potato. These examples of R-genes show that comparative mapping of functionally proven R-genes may give a lead towards new candidate R-genes. But, even in case the complete genome sequence is available and QTLs for resistance are located, the identification of the specific R-gene copy within a cluster of RGAs may require laborious fine mapping and synteny studies. Amplification of RGAs from specific chromosomes, isolated by microdissection (Huang et al. 2004) or flow sorting (Safar et al. 2004) could reduce the number of RGAs to be screened and speed up the identification of target RGAs. Linden van der et al. (2004) developed an interesting tool for RGA mapping termed NBS-profiling. They used the common motifs that are present in the NBS regions of R-genes in combination with nearby restriction sites in more variable regions for the PCR-amplification of a large collection of RGA-fragments which at the same time provide molecular markers that are tightly linked to R-genes. Because RGA-clusters are often linked to QTLs for R-genes, this method can be applied for R-gene mapping in a wide range of crops (e.g. Calenge et al. 2005). This profiling approach to detect RGAs linked to (and segregating with) resistance is a valuable method to obtain markers for e.g. Fusarium resistance in potato.

Whereas the resistance to Fusarium in tomato is race-specific, resistance to *F. graminearum* and *F. culmorum*, the causative agents of head blight (scab) in wheat and other cereals and grasses is a quantitative trait (horizontal resistance), with

relative high heritability and controlled by a few genes with major effects (Yang ZhuPing et al. 2005) which renders the breeding for this trait very complex and it remains to be seen what type of genes are involved. Complete resistance has not been discovered yet but a major QTL (*Qfhs.ndsu-3BS* locus) is located on the short arm of chromosome 3B (3BS) in different populations (for references Paillard et al. 2004; Snijders 2004; Yang ZhuPing et al. 2005). A fine map spanning 0.2 to 1.5 cm of this QTL locus of wheat was generated by Liu et al. (2005) and compared with syntenic regions in rice (1S) and barley (3HS). However, the synteny studies in barley and rice for this genomic region were complicated by micro-rearrangements such as inversions and insertion/deletions which hampered the direct comparative map based cloning of the CGs (Brunner et al. 2003; Liu et al. 2005). As soon as the genomic sequence of wheat and *Brachy-podium*, the syntenic species for *Triticeae*, are available the identification of genes underlying the QTL and the development of genetic markers for breeding is expected to make a fast progress.

Early pod shatter

Early pod shatter is an undesired trait that still can cause serious seed yield losses in Brassica species like cabbage (*Brassica oleracea*), oilseed rape (*B. napus, B. rapa, and B. juncea*) and Crambe (*C. hispanica and C. abyssinica*). Control of pod shatter is therefore a target of many breeding programs. Seeds included in a pod, as in oilseed rape and other Brassica species, disperse by opening of the silique (pod) at the dehiscence zone whereas seeds included in a mono-seeded pod such as in Crambe disperse by breakage at the dehiscence zone located between pod and pedicel. In Crambe, pod shatter behavior is most likely controlled by one or two loci with brittle dominant over non-brittle, but no corresponding genes are cloned as yet (personal communication D Mastenbroek, Crambe breeder). There is little genetic variation for resistance to pod shatter within the *B. napus* gene pool but interspecific crosses of wild relatives provided newly synthesized *B. napus* lines with useful variation for the

trait (ref e.g. Morgan et al. 2003, 1998). However, these plant hybrids are often related to unfavorable characteristics that must be regained by backcrossing. In the 'model' Arabidopsis, pod shattering behavior has been studied extensively and at the moment, several CGs for the trait have been identified, including *SHP1*, *SHP2*, *IND* and *ALC*. The *SHATTERPROOF* genes *SHP1* and *SHP2* are TFs belonging to the MADS-box gene family. The two genes are functionally redundant since only pods of lines that carry mutations in both genes fail to dehisce (reviewed in Liljegren et al. 2000, 2004). The *ALCATRAZ* (*ALC*) (Rajani and Sundaresan, 2001) and *INDEHIS-CENT* (*IND1*) (Liljegren et al. 2004) genes, both basic helix loop (bHLH) TFs, promote the differentiation of specific cells that are needed for pod opening. Another MADS transcription factor, FRUITFULL (FUL) mediates pod shattering by inhibiting the *SHP* genes (Ferrándiz et al. 2000). Recently, it was shown that ectopic expression of the Arabidopsis *FRUITFULL* gene in *B. juncea* is sufficient to produce pod shatter-resistant Brassica fruit (Østergaard et al. 2006). Furthermore, Arabidopsis protein GARGOYLE (GGL), identified by upregulation of the gene in activation tagging, was associated with a significant reduction in pod shattering due to an alteration of the lignification of the silique (Aharoni and Pereira 2006). In contrast to the *FUL* and *GGL* genes, expression of the *SHP* genes are correlated with the unwanted phenotype and a specific loss of or change in this CG's function is required for crop improvement. The chance of finding the functional orthologs of these other CGs in Brassica crop species is high because Brassica species are close relatives of the 'model' plant Arabidopsis (Snowdon and Friedt 2004) and because the genetic pathway leading to specification of the dehiscence zone seemed to be conserved between Arabidopsis and Brassica (Østergaard et al. 2006). Recently two INDl orthologs, *Bn IND1* and *Bn IND2*, were isolated from *B. napus* that were able to complement the Arabidopsis ind1 mutant phenotype demonstrating that Bn INDl and Bn IND2 carry out the same basic functions as IND1. Based on these genes several GM approaches have been proposed to improve the pod shattering trait (Yanofsky and Kempin 2006). Alternatively, the

expression of these CGs may be regulated via a targeted mutagenesis approach. In case of a targeted mutagenesis approach in a polyploid crop (Slade et al. 2004) like *B. napus*, redundancy of gene function may mask phenotypic changes related to a mutation. However, independently from a detectable phenotype, a series of mutants in the putative CG alleles obtained by targeted mutagenesis and the subsequent combination of the putative effective mutations (loss- or change of function mutations, mutations affecting conserved aminoacids motifs or splice sites) by breeding may help to 'translate' the CGs into a reduced, non-GM, shattering trait in Brassica spp.

Concluding remarks

From the examples presented above it becomes clear that ongoing genomic research provides an increasing body of information on gene functions in 'model' organisms (Bro and Nielsen 2004; Meyers et al. 2004; Aharoni et al. 2000). How we can use this expanding source of genomic knowledge, with the highest chance on success, for crop improvement via a CGA depends on several factors such as the colinearity of the genomes that are compared to deduce orthologous CGs for a trait, the availability of a closely related sequenced 'model' for translational genomics and the presence of much as possible genetic variation for genetic mapping studies. Comparative mapping can help to identify CGs underlying QTL and to find orthologs in target crops. However, even within species the colinearity of genomes is not always perfect (e.g. Song and Messing 2003, Brunner et al. 2005) and apparently similar biochemical pathways may have diverged during evolution (e.g. Yan et al. 2003, 2004; Kato et al. 2002). Therefore, validation of CGs by proper genetics, comparative and physical mapping and mutant studies is still recommended to prove linkage with a trait in both, 'model' and 'target'.

If a CG is identified in a 'model', the translation of this information to a 'target' crop is crucial for implementation into practice. However, as our knowledge grows it becomes apparent that many traits are more complex than previously suspected, with complex regulation of gene expression and interactions between regulatory pathways being just a few of the causes (Borevitz and Ecker 2004). In addition, in human it was shown that besides SNPs, duplications and deletions, large scale copy number polymorphisms or variations (CNPs/LCVs) may underlie a diverse range of phenotypes from body weight to cancer susceptibility (Sebat et al. 2004; Iafrate et al. 2004). In plants such differences in expression level among orthologs seemed to cause differences in flowering time via a 'retuning' of the conserved photoperiod pathway (reviewed by Laurie et al. 2004; Koornneef et al. 2004) and also yield in rice seemed to be controlled by allelic variation in the expression and structure of a gene cluster associated with a quantitative trait locus for improved yield in rice (He et al. 2006). In such complex situations genetic variation in combination with high throughput and sensitive SNP detection methods are important to offer the possibility to screen for allelic differences at the expression level (Meyers et al. 2004; Schaart et al. 2005) and to discriminate allelic forms (haplotypes) of a CG within the complete germplasm pool of a species. Also, the emerging concept exists that it is good to have for all important crops and/or plant families a good sequenced model. Considering all these aspects of a CGA and despite the complex nature of many crop traits we expect that, with the increased possibilities at the technical level and in the field of data integration, genomic research creates indispensable tools for breeding in crop species.

References

Aharoni A, Pereira A (2006) Method for preventing dehiscence and altering plant lignification. Patent Appl no. EP1621630

Aharoni A, Dixit S, Jetter R, Thoenes E, Arkel G van, Pereira A (2004) The SHINE clade of AP2 domain transcription factors activates wax biosynthesis, alters cuticle properties, and confers drought tolerance when overexpressed in Arabidopsis. Plant Cell 16(9):2463–2480

Aharoni A, Keizer LCP, Bouwmeester HJ, Sun Zhong Kui, Alvarez Huerta M, Verhoeven HA, Blaas J, Houwelingen AMML van, Vos RCH de, Voet Hvander, Jansen RC, Guis M, Mol J, Davis RW, Schena M,

Tunen AJ van, O'Connell AP (2000) Identification of the SAAT gene involved in strawberry flavor biogenesis by use of DNA microarrays. Plant Cell 12(5):647–661

Bai JF, Pennill LA, Ning JC, Lee SW, Ramalingam J, Webb CA, Zhao BY, Sun Q, Nelson JC, Leach JE, Hulbert SH (2002) Diversity in nucleotide binding site-leucine-rich repeat genes in cereals. Genome Res 12(12):1871–1884

Bent A (1996) Function meets structure in the study of plant disease resistance genes. Plant Cell 8:1757–1771

Borevitz JO, Ecker JR (2004) Plant genomics: the third wave. Annu rev Genomics Hum Genet 5:443–477

Bro C, Nielsen J (2004) Impact of 'ome' analyses on inverse metabolic engineering. Metab Eng 6:204–211

Brunner S, Fengler K, Morgante M, Tingey S, Rafalski A (2005) Evolution of DNA sequence nonhomologies among maize inbreds. Plant Cell 17:343–360

Brunner AM, Nilsson O (2004) Revisiting tree maturation and floral initiation in the poplar functional genomics era. New Phytol 164:43–51

Brunner S, Keller B, Feuillet C (2003) A large rearrangement involving genes and low-copy DNA interrupts the microcollinearity between rice and barley at the Rph7 locus. Genetics 164(2):673–683

Byrne PF, McMullen MD (1996) Defining genes for agricultural traits: QTL analysis and the candidate gene approach. Probe 7:24–27

Caicedo AL, Purugganan MD (2005) Comparative plant genomics. Frontiers and prospects. Plant Physiol 138(2):545–547

Calenge F, Linden CG van der, Weg E van de, Schouten HJ, Arkel G van, Denance C, Durel CE (2005) Resistance gene analogues identified through the NBS-profiling method map close to major genes and QTL for disease resistance in apple. Theor Appl Gen 110(4):660–668

Chervitz SA, Aravind L, Sherlock G, Ball CA, Koonin EV, Dwight SS, Harris MA, Dolinski K, Mohr S, Smith T, Shuai-Weng, Cherry JM, Botstein D (1998) Comparison of the complete protein sets of worm anfd yeast: orthology and divergence. Science 282:2022–2028

Czechowski T, Bari RP, Stitt M, Scheible WR, Udvardi MK (2004) Real-time RT-PCR profiling of over 1400 Arabidopsis transcription factors: unprecedented sensitivity reveals novel root- and shoot-specific genes. Plant J 38:366–379

Dangl JL, Jones JDG (2001) Plant pathogens and integrated defence responses to infection. Nature 411(6839):826–833

Devos KM, Beales J, Nagamura-Yoshiaki, Sasaki-Takuji (1999) Arabidospsis-rice: will colinearity allow gene prediction across the eudicot-monocot divide? Genome Res 9:825–829

Dubouzet JG, Sakuma Y, Ito Y, Kasuga M, Dubouzet EG, Miura S, Seki M, Shinozaki K, Yamaguchi Shinozaki K (2003) OsDREB genes in rice, Oryza sativa L., encode transpcrition activators that function in drought-, high-salt-, and cold-responsive gene expression. Plant J 33:751–763

Eeuwijk FA van, Mesterhazy A, Kling CI, Ruckenbauer P, Saur L, Burstmayr H, Lemmens M, Keizer LCP, Maurin N, Snijders CHA (1995) Assessing non-specificity of resistance in wheat to head blight caused by inoculation with European strains of Fusarium culmorum, F. graminearum and F. nivale using a multiplicative model for interaction. Theor Appl Gen 90(2):221–228

Feltus FA, Wan Jun, Schulze SR, Estill JC, Jiang Ning, Paterson AH (2004) An SNP Resource for Rice genetics and breeding based on subspecies Indica and Japonica genome alignments. Genome Res 14:1812–1819

Ferrándiz C, Liljegren SJ, Yanofsky MF (2000) Negative regulation of the SHATTERPROOF genes by FRUITFULL during Arabidopsis fruit development Science 289(5478):436–438

Flint-Garcia SA, Thornsberry JM, Buckler ES (2003) Structure of linkage disequilibrium in plants. Ann Rev Plant Biol 54:357–374

Francki MG, Mullan DJ (2004) Application of comparative genomics to narrow-leafed lupin (Lupinus angustifolius L.) using sequence information from soybean and Arabidopsis. Genome 47:623–632

Gepts P, Beavis WD, Brummer EC, Shoemaker RC, Stalker HT, Weeden NF, Young ND (2005) Legumes as a Model Plant Family. Genomics for food and feed report of the cross-legume advances through genomics conference. Plant Physiol 137:1228–1235

Goff SA, Ricke D, Lan TH, et al (2002) A draft sequence of the rice genome (Oryza sativa L. ssp. japonica). Science 296:92–100

Gutterson N, Reuber TL (2004) Regulation of disease resistance pathways by AP2/ERF transcription factors. Curr Opinion Plant Biol 7:465–471

Gutterson N, Zhang JZ (2004) Genomics applications to biotech traits: a revolution in progress? Curr Opin Plant Biol 7:226–230

Hammond-Kosack KE, Jones JDG (1997) Plant disease resistance genes. Ann Rev Plant Physiol Plant Mol Biol 48:575–607

He G, Luo X, Tian F, Li K, Zhu Z, Su W, Qian X, Fu Y, wang X, Sun C, Yang J (2006) Haplotype variation in structure and expression of a gene cluster associated with a quantitative trait locus for improved yield in rice. Genome Res 6:618–626

He LM, Du CG, Covaleda L, Xu ZY, Robinson AF, Yu JZ, Kohel RJ, Zhang HB (2004) Cloning, characterization, and evolution of the NBS-LRR-encoding resistance gene analogue family in polyploid cotton (Gossypium hirsutum L). Mol Plant Microbe Inter 17(11):1234–1241

Hemming MN, Basuki S, McGrath DJ, Carroll BJ, Jones DA (2004) Fine mapping of the tomato I-3 gene for fusarium wilt resistance and elimination of a co-segregating resistance gene analogue as a candidate for I-3. Theor Appl Gen 109(2):409–418

Henikoff S, Comai L (2003) single-nucleotide mutations for plant functional genomics. Annu Rev Biol 54:375–401

Huan SanWen, Vossen EAG van der, Kuang HH, Vleeshouwers VGAA, Zhang NingWen, Borm TJA, Eck

HJ van, Baker B, Jacobsen E, Visser RGF (2005) Comparative genomics enabled the isolation of the *R3a* late blight resistance gene in potato. The Plant J 42:251–261

Huang D, Wu W, Lu L (2004) Microdissection and molecular manipulation of single chromosomes in woody fruit trees with small chromosomes using pomelo (*Citrus grandis*) as a model. II. Cloning of resistance gene analogs from single chromosomes. Theor Appl Gen 108(7):1371–1377

Hulbert SH, Webb CA, Smith SM, Sun Q (2001) Resistance gene complexes: evolution and utilization. Annu Rev Phytopathol 39:285–312

Iafrate AJ, Feuk L, Rivera MN et al (2004) Detection of large-scale variation in the human genome. Nature Genet 36(9):949–951

ItoY, Katsura K, Maruyama K, Taji T, Kobayashi M, Seki M, Shinozaki K, Yamaguchi-Shinozaki K (2006) Functional analysis of rice DREB1/CBF-type transcription factors involved in cold-responsive gene expression in transgenic rice Plant and Cell Physiol 47(1):141–153

Jaiswal P, Ni J, Yap I, Ware1 D, Spooner W, Youens-Clark K, Ren L, Liang C, Zhao W, Ratnapu K, Faga B, Canaran P, Fogleman M, Hebbard C, Avraham S, Schmidt S, Casstevens TM, Buckler ES, Stein L, McCouch S (2006) Gramene: a bird's eye view of cereal genomes. Nucl Acid Res 34:D717–D723

Kanazin V, Marek LF, Shoemaker RC (1996) Resistance gene analogs are conserved and clustered in soybean. Proc Natl Acad Sci USA 93:11746–11750

Kasuga M, Liu Q, Miura S, Yamaguchi-Shinozaki K, Shinozaki K (1999) improving plant drought, salt and freezing tolerance by gene transfer of a single stress-inducible transcription factor. Nat Biotechnol 17:287–291

Kasuga M, Miura S, Shinozaki K, Yamaguchi-Shinozaki K (2004) A combination of the Arabidopsis DREB1A gene and stress-inducible rd29A promoter improved drought- and low-temperature stress tolerance in tobacco by gene transfer. Plant and Cell Physiol 45(3):346–350

Kato K, Kidou S, Miura H, Sawada S (2002) Molecular cloning of the wheat CK2a gene and detection of its linkage with Vrn-A1 on chromosome 5A. Theor Appl Gen 104:1071–1077

Koorneef M et al (2004) Naturally occurring genetic variation in Arabidopsis thaliana. Annu Rev Plant Biol 55:141–172

Krutovsky KV, Troggio M, Brown GR, Jermstad KD, Neale DB (2004) Comparative mapping in the Pinaceae. Genetics 168:447–461

Lai J, Ma J, Swigonova Z, Ramakrishna W, linton E, Llaca V, Tanyolac B, Park Y-J, Jeong O-Y, Bennetzen JL, Messing J (2006) Gene loss and movement in the maize genome. Genome Res 14:1924–1931

Laurie DA, Griffiths S, Dunford RP, Christodoulou V, Taylor SA, Cockram J, Beales J, Turner A (2004) Comparative genetic approaches to the identification of flowering time genes in temperate cereals. Field Crops Res 90:87–99

Liljegren SJ, Ditta GS, Eshed Y, Savidge B, Bowman JL, Yanofsky MF (2000) SHATTERPROOF MADS-box genes control seed dispersal in Arabidopsis. Nature 404:766–770

Liljegren SJ, Roeder AHK, Kempin SA, Gremski K, Ostergaard L, Guimil S, Reyes DK, Yanofsky MF (2004) Control of fruit patterning in Arabidopsis by INDEHISCENT. Cell 116(6):843–853

Linden CG van der, Wouters DCAE, Mihalka V, Kochieva EZ, Smulders MJM, VosmanB (2004) Efficient targeting of plant disease resistance loci using NBS profiling. Theor Appl Genet 109:384–393

Liu S, Zhang X, Pumphrey MO, Stack RW, Gill BS, Anderson JA (2005) Complex microcolinearity among wheat, rice, and barley revealed by fine mapping of the genomic region harboring a major QTL for resistance to Fusarium head blight in wheat. Funct Integr Genomics 6:83–89

Mago R, Nair S, Mohan M (1999) Resistance gene analogues from rice: Cloning, sequencing and mapping. Theor Appl Genet 99:50–57

Marè C, Mazzucotelli E, Crosatti C, Francia E Stanca AM, Cattivelli L (2004) Hv-WRKY38: a new transcription factor involved in cold- and drought-response in barley. Plant Mol Biol 55:399–416

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

Mesterhazy A, Bartok T, Mirocha CG, Komoroczy R (1999) Nature of wheat resistance to Fusarium head blight and the role of deoxynivalenol for breeding. Plant-Breeding 118(2):97–110

Meyers BC, Galbraith DW, Nelson T, Agrawal V (2004) Methods for transcriptional profiling in plants. Be fruitful and replicate. Plant Physiol 135(2):637–652

Morgan CL, Bavage A, Bancroft I, Bruce DM, Child R, Summers J (2003) Using novel variation in Brassica species to reduce agricultural inputs and improve agronomy of oilseed rape—a case study in pod shatter resistance. Plant Genet Resour 1:59–65

Morgan CL, Bruce DM, Child RD, Landbrooke ZL, Arthur AE (1998) Genetic variation for pod shatter resistance among lines of oilseed rape developed from synthetic *B. napus*. Field Crops Res 58:153–165

Oh-SeJun, Song SangIk, Kim YounShic, Jang HyunJun, Kim SooYoung, Kim MinJeongx, Kim YeonKi, Nahm BaekHie, Kim JuKon (2005) Arabidopsis CBF3/DREB1A and ABF3 in transgenic rice increased tolerance to abiotic stress without stunting growth. Plant Physiol 138(1):341–351

Østergaard L, Kempin SA, Bies D, Klee HJ and Yanofsky MF (2006) Pod shatter-resistant Brassica fruit produced by ectopic expression of the FRUITFULL gene. Plant Biotechnol J 4:45 doi:10.1111/j.1467-7652.2005.00156

Paal J, Henselewski H Muth J, Meksem K, Menendez CM, Salamini F, Ballvora A, Gebhardt C (2004) Molecular cloning of the potato Gro1-4 gene conferring resistance to pathotype Ro1 of the root cyst nematode Globodera rostochiensis, based on a candidate gene approach. Plant J 38(2):285–297

Paillard S, Schnurbusch T, Tiwari R, Messmer M, Winzeler M, Keller B, Schachermayr G (2004) QTL analysis of resistance to Fusarium head blight in Swiss winter wheat (Triticum aestivum L.). Theor Appl Genet 109(2):323–332

Pan Q, Wendel J, Fluhr W (2000) Divergent evolution of plant NBS-LRR resistance gene homologues in dicot and cereal genomes. J Mol Evol 50:203–213

Pellegrineschi A, Reynolds M, Pacheco M, Brito RM, Almeraya R, Yamaguchi- Shinozaki K, Hoisington D (2004) Stress-induced expression in wheat of the Arabidopsis thaliana DREB1A gene delays water stress symptoms under greenhouse conditions. Genome 47(3):493–500

Peng JR, Richards DE, Hartley NM et al (1999) 'Green revolution' genes encode mutant gibberellin response modulators. Nature 400(6741):256–261

Pflieger S, Lefebvre V, Causse M (2001) The candidate gene approach in plant genetics: a review. Mol Breed 7(4):275–291; Plant Physiol 138(2):567–568

Rajani S, Sundaresan V (2001) The Arabidopsis myc/bHLH gene ALCATRAZ enables cell separation in fruit dehiscence. Curr Biol 11(24):1914–1922

Ramalingam J, Vera Cruz CM, Kukreja K, Chittoor JM, Wu JL, Lee SW, Baraoidan M, George ML, Cohen MB, Hulbert SH, Leach JE, Leung H (2003) Candidate defense genes from rice, barley, and maize and their association with qualitative and quantitative resistance in rice. Mol Plant–Microbe Inter 16:14–24

Reyes JC, Muro-Pastor MI, Florencio FJ (2004) The GATA family of transcription factors in Arabidopsis and rice. Plant Physiol 134:1718–1732

Riechmann JL, Heard J, Martin G, Reuber L, Jiang CZ, Keddie J, Adam L, Pineda O, Ratcliffe OJ, Samaha RR, Creelman R, Pilgrim M, Broun P, Zhang JZ, Ghandehari D, Sherman BK, Yu GL, et al (2000) Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. Science 290:2105–2110

Rubin GM, Yandell MD, Wortman JR (2000) Comparative genomics of the eukaryotes. Science 287(5461):2204–2215

Rudd S (2005) Open Sputnik—a database to ESTablish comparative plant genomics using unsaturated sequence collections. Nucleic Acids Res 33(1):D622–D627

Rudd S, Schoof H, Mayer K (2005) PlantMarkers—a database of predicted molecular markers from plants. Nucl Acids Res 33(1):D628–D632

Safar J, Bartos J, Janda J, Bellec A, Kubalakova M, Valarik M, Pateyron S, Weiserova J, Tuskova R, ihalikova J, Vrana J, Simkova H, Faivre-Rampant P, Sourdille P, Caboche M, Bernard M, Dolezel J, Chalhoub B (2004) Dissecting large and complex genomes: flow sorting and BAC cloning of individual chromosomes from bread wheat. Plant J 39(6):960–968

Sakuma Y, Maruyama K, Osakabe Y, Qin-Feng, Seki M, Shinozaki K, Yamaguchi-Shinozaki K (2006) Functional analysis of an Arabidopsis transcription factor, DREB2A, involved in drought-responsive gene expression. Plant-Cell 18(5):1292–1309

Schaart JG, Mehli L, Schouten HJ (2005) Quantification of allele-specific expression of a gene encoding strawberry polygalacturonase-inhibiting protein (PGIP) using PyrosequencingTM. Plant J 41(3):493–500

Sebat J, Lakshmi B, Troge J, et al (2004) Large-scale copy number polymorphism in the human genome. Science 305:525–528

Shinozaki K, Dennis ES, Seki M (2003) regulatory network of gene expression in drought and cold stress responses. Curr Opinion Plant Biol 6:410–417

Shinozaki K, Dennis ES, Seki M (2003) Regulatory network of gene expression in the drought and cold stress responses. Curr Opin Plant Biol 6(5): 410–417

Simons G, Groenendijk J, Wijbrandi J, Reijans M, Groenen J, Diergaarde P, Lee T van der, Bleeker M, Onstenk J, Both M de, Haring M, Mes J, Cornelissen B, Zabeau M, Vos P (1998) Dissection of the Fusarium I2 gene cluster in tomato reveals six homologs and one active gene copy. Plant Cell 10(6):1055–1068

Slade AJ, Fuerstenberg SI, Loeffler D, Steine MN, Facciotti D (2004) A reverse genetic, nontransgenic approach to wheat crop improvement by TILLING. Nature Biotech 23:75–81

Snijders CHA (2004) Resistance in wheat to Fusarium infection and trichothecene formation. Toxicol Lett 153(1):37–46

Snowdon RJ, Friedt W (2004) Molecular markers in Brassica oilseed breeding: current status and future possibilities. Plant Breeding 123(1):1–8

Song R, Messing J (2003) Gene expression of a gene family in maize based on noncollinear haplotypes. PNAS 100(15):9055–9060

Stacey G, VandenBosch K (2005) "Translational" Legume Biology. Models to Crops. Plant Physiol 137(4):1173–1173

Stein LD (2004) Human genome: end of the beginning. Nature 431:915–916

Thabuis A, Palloix A, Pflieger S et al (2003) Comparative mapping of Phytophthora resistance loci in pepper germplasm: evidence for conserved resistance loci across Solanaceae and for a large genetic diversity. Theor Appl Gen 106(8):1473–1485

The Arabidopsis genome initiative (2000) Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature 408:796–815

Tian ChaoGuang, Wan Ping, Sun ShouHong, Li JiaYang, Chen MingSheng (2004) Genome-wide analysis of the GRAS gene family in rice and Arabidopsis. Plant Mol Biol 54:519–532

Town CD, Cheung F, Maiti R, Crabtree J, Haas BJ, Wortman JR, Hine EE, Althoff R, Arbogast TS, Tallon LJ, Vigouroux M, Trick M, Bancroft I (2006) Comparative genomics of Brassica oleracea and Arabidopsis thaliana reveal gene loss, fragmentation, and dispersal after polyploidy. Plant Cell 18(6):1348–1359

Tuskan GA, Difazio S, Jansson S, Bohlmann J, et al (2006) The genome of black cottonwood, Populus trichocarpa (Torr. & Gray). Science 15:1596–604

Yan L, Loukoianov A, Tranquilli G, Helguera M, Fahima T, Dubcovsky J (2003) Positional cloning of the wheat vernalization gene VRN1. PNAS 100:6263–6268

Yan L, Loukoianov A, Blechl A, Tranquilli G, Ramakrishna W, San Miguel P, Bennetzen JL, Echenique V, Dubcovsky J (2004) The wheat VRN2 gene is a flowering repressor down-regulated by vernalization. Science 303(5664):1640–1644

Yang ZhuPing, Gilbert J, Fedak G, Somers DJ (2005) Genetic characterization of QTL associated with resistance to Fusarium head blight in a doubled haploid spring wheat population. Genome 48:187–196

Yang T-J, Jung Sun Kim, Kwon S-J, et al (2006) Sequence-level analysis of the diploidization process in the triplicated FLOWERING LOCUS C region of *Brassica rapa*. The Plant Cell 18:1339–1347

Yanofsky MF, Kempin S (2006) Brassica INDEHISCENT1 Sequences. Patent cooperation treaty application: WO2006009649

Yu Jun, Hu Song Nian, Wang-Jun, et al (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). Science 296:79–92