Gene Family Evolution in *Allium* Species

10

Jiffinvir Khosa, Robyn Lee, John McCallum and Richard Macknight

Abstract

During evolution, plant genomes have undergone duplications, deletions and rearrangements resulting in a wide variation in genome size and number of gene family members between different species. The variation in gene families is an important mechanism for adaptation to different environmental conditions. *Allium* species, such as bulb onion (*Allium cepa*), have a large unsequenced genome. However, high throughput transcriptome sequencing datasets are now available which provide an efficient way to identify the genes present in different *Allium* species. With this knowledge, strategies to accelerate physiological and genetic analysis for enhanced breeding can be developed. In this chapter, we will describe how RNA sequencing is providing a better understanding of *Allium* genetics and survey the diversity of gene families involved in bulbing, flowering, male fertility, flavonoid biosynthesis and sulphur assimilation in bulb onion. In general, we found that onion has a similar number of gene family members to other monocots, such as rice, which have much smaller genomes. This is consistent with the large genome size of *Allium* being due to a massive expansion of repetitive DNA.

10.1 Introduction

Over the last 17 years, the sequencing of plant genomes has provided tremendous resources for identifying genes underlying agriculturally important traits and equipping breeders with new tools to enhance the breeding of new cultivars (Varshney et al. 2014). The sequencing of the bulb onion genome and other *Allium* species has been hampered by their very large genome size (16 Gbp) and complex structure (McCallum 2007; Khosa et al. 2016b). The bulb onion genome is about 40 times larger than the rice genome (0.4 Gbp), which is well annotated and provides an excellent reference for comparing to other grasses (McCallum 2007; Jakse et al. 2008). Analysis of the genomes of other plant species has revealed that increases in genome size are largely due to whole-genome duplications and an increase in repetitive DNA. For example, it is thought that the ancestor of grasses had five chromosomes, and genome duplication resulted in first 10 chromosomes, then the current 12 chromosomes in rice. The large increase in the genome size of other grasses (for example,

J. Khosa · R. Lee · R. Macknight (✉)
Department of Biochemistry, University of Otago, Dunedin, New Zealand
e-mail: richard.macknight@otago.ac.nz

J. McCallum · R. Macknight
New Zealand Institute for Plant and Food Research Ltd., University of Otago, Dunedin, New Zealand

M. Shigyo et al. (eds.), *The Allium Genomes*, Compendium of Plant Genomes,
https://doi.org/10.1007/978-3-319-95825-5_10

2.4 Gbp for maize and 17 Gbp for wheat) has been the result of duplication and increase in repetitive sequences (Wendel et al. 2016).

The *Allium* genus is very large, comprising 972 accepted species (see theplantlist.org) that are widely distributed over the Holarctic region from the dry subtropics to the boreal zone (Fritsch and Friesen 2002). While most species grow in open, sunny, rather dry sites in arid and moderately humid climates, *Alliums* have also evolved to grow in diverse habitats. Ornamental and vegetable *Alliums* exhibit correspondingly wide variability in stature, branching, storage organs, and floral structures (Hanelt 1990). Bulb onion (*Allium cepa*) has been cultivated for over 4000 years, and carvings of onions exist on the pyramid walls highlighting their importance in the Egyptian diet. Crop domestication resulted in strong genetic selection for key genes (and specific variants) that control the plant structures and physiological attributes which distinguish crop plants from their wild relatives (Doebley et al. 2006; Baldwin et al. 2013). Analysing the diversity of key gene families involved in plant architecture and adaptation within and among domesticated *Allium* and their wild allies may provide insights into the key genes under selection during domestication and dispersal (Ross-Ibarra et al. 2007).

In this chapter, we compare gene families involved in the regulation of key traits of bulb onion, such as bulbing, flowering, male fertility, flavonoid biosynthesis, and sulphur assimilation with the corresponding gene families present in other plants. The comparison suggests that the large genome in genus *Allium* is not due to the widespread expansion of individual gene families. This is consistent with a recent prediction that bulb onion has about 35,000 protein-coding genes, similar to other plants that have small genomes such as *Arabidopsis* (27,416) and rice (37,544) (Wendel et al. 2016; Sohn et al. 2016). While in cases, individual gene families have undergone expansion, other gene families contain fewer members than other plant species. Thus, rather than large-scale gene or genome duplications, the large genome in genus *Allium* is likely to be the result of a massive expansion of the amount of repetitive DNA (King et al. 1998; Jakse et al. 2008).

10.2 RNA Sequencing for Gene Identification in Genus *Allium*

RNA sequencing (RNA-Seq) provides a powerful way to identify the genes in species with large genomes (Mutz et al. 2013; Martin et al. 2013). RNA-Seq has been used in *Allium* species, such as bulb onion, garlic, bunching onion and chives to identify candidate genes regulating various traits (Khosa et al. 2016b). A wide range of genotypes and tissues have been sequenced using different platforms (Table 10.1). The focus of these RNA-Seq studies has been on developing molecular markers and discovering candidate genes, involved in bulbing, flowering, flower development, fertility restoration, sulphur biosynthesis and allergenicity (Table 10.1). In addition, transcriptome data has been generated in bulb onion treated with cold temperatures in an effort to better understand freezing tolerance mechanisms (Han et al. 2016). Recently, a comparative transcriptome analysis of nine different *Allium* species has been carried out to better understand morphogenesis and evolution of fistular leaves in genus *Allium* (Zhu et al. 2017).

The number of protein-encoding genes in bulb onion has been estimated to be about 35,000. However, the number of predicted transcripts from these studies is often very high (up to 293,000 unique transcripts), due to multiple transcripts being derived from a single gene and the presence of transcripts that do not encode functional proteins (Table 10.1) (Sohn et al. 2016). These large transcriptome datasets provide an opportunity to compare the number of genes in particular gene families with those of other plant species. The number of distinct genes in each transcription factor gene family varies between plant species. However, overall bulb onion has a similar number of transcription factors to rice (Table 10.2). The variation of gene

Table 10.1 Summary of published RNA sequencing projects in different *Allium* species

Species	Sequencing platform	Genotype	Gene identification	Number of transcripts	Tissue	References
Bulb onion	Illumina HiSeq 2000 platform	CUDH2107	Genes involved in flower development	271,665	Leaves, floral buds from unexpanded umbels, unopened florets from expanded umbels, open florets with pollen, older flowers and roots	Khosa et al. (2016a)
	Illumina HiSeq 2000	36,122 and 36,101	Cold tolerance	93,637	Leaves	Han et al. (2016)
	Illumina HiSeq™ 2500	Chalinghuangpi	NAC transcription factor	117,189	Leaves	Zheng et al. (2016a)
	Illumina HiSeq 2000	Pusa Madhavi	Identification of allergens and epitopes	293,475	Bulb	Rajkumar et al. (2015)
	HisSeq™ 2500,	Utah Yellow Sweet Y1351	Carbohydrate metabolism	79,376	Bulb	Zhang et al. (2016)
	Illumina HiSeq 2000	H6, SP3B	–	165,179	Six weeks whole seedlings	Kim et al. (2014)
	Illumina HiSeq 2000	Bravo, Jumbo, Babosa, California Red, Pukekohe Longkeepe, Rio Tinto, Rumba, Sapporo Yellow Globe and South Port White Globe	–	46,596; 36,897; 99,010; 81,574; 99,761; 81,975; 103,178; 69,206 76,187	Leaves of 4–8-week-old plants	Scholten et al. (2016)
	454 sequencing	CUDH2150	Identification of *FT* gene family	24,106	Leaves and shoot meristem at the 4–5 leaf stage	Baldwin et al. (2012)
	PacBio RSII system	Eumjinara and Sinsunhwang	–	99,247	Flower, leaf, bulb and root	Sohn et al. (2016)
	ABI PRISM 3730XL analyzer	506L and H6	Male fertility restorer genes	32,674	Unopened flowers	Kim et al. (2015)
	Roche 454 FLX	OH1 and 5225	–	27,065; 33,254	Vernalized bulbs, tissue from leaves, unopened umbels, bulbs and roots	Duangjit et al. (2013)
	Illumina HiSeq 2500	–	–	117,189	Leaves	Zhu et al. (2017)
	Illumina Hiseq 2000	Orlando	PCD related genes	45,891	Bulb scales	Galsurker et al. (2017)
Shallot	Illumina HiSeq 2500	–	–	83,186	Leaves	Zhu et al. (2017)
Garlic	Illumina HiSeqTM 2000	Cangshan 15	Genes involved in organic sulphur biosynthesis	127,933	Vegetative buds	Sun et al. (2012)

(continued)

Table 10.1 (continued)

Species	Sequencing platform	Genotype	Gene identification	Number of transcripts	Tissue	References
	–	Cangshan 15	Shoot apex sprouting	45,363	Dormant and sprouting garlic shoot apex	Sun et al. (2013)
		Fertile garlic clone #87	Photoperiodic flowering; flower development and organosulfur metabolism	240,000	Root, Basal plate, Leaf, Clove, Inflorescence and Flowers	Kamenetsky et al. (2015)
		Fertile #87 (F87) and male-sterile #96 (MS96)	Genes involved in male sterility and fertility	–	Flower (Early, Mid and late)	Shemesh-Mayer et al. (2015)
	HiSeqTM 2500	ChalingZiPiSuan		135,360	Bulbs, and whole plant	Liu et al. (2015)
	Illumina HiSeq 2500	–	–	132,225	Leaves	Zhu et al. (2017)
Bunching onion	GS-FLX; Illumina HiSeq 2000	Ki," "F" and "A."	–	42,511;121,354	2-week-old seedlings, leaf, roots, basal meristem, immature flower bract, mature bract (about 1 week before anthesis), opened flowers, immature fruits and sliced pseudostem	Tsukazaki et al. (2015)
	Illumina HiSeq 2000	Bian Gan and glossy BianGan	Waxy cuticle biosynthesis	73,128	Leaves	Liu et al. (2014)
	Illumina HiSeq 2000	Zhangqiu	Genes involved in sulphur and selenium metabolism	103,286	Leaves, false stem, basal plate and root were collected from 14-day old seedlings	Sun et al. (2016)
	Illumina HiSeq 2500	–	–	128,372	Leaves	Zhu et al. (2017)
Chinese Chive	Illumina HiSeq 2000	–	–	150,154	Leaves, shoots and roots	Zhou et al. (2015)
A. ascalonicum	Illumina HiSeq 2500		–		Leaves	Zhu et al. (2017)
A. chinense	Illumina HiSeq 2500	–	–	121,008	Leaves	Zhu et al. (2017)
A. macrostemon	Illumina HiSeq 2500		–	161,681	Leaves	Zhu et al. (2017)
A. tuberosum	Illumina HiSeq 2500		–	148,715	Leaves	Zhu et al. (2017)
A. porrum	Illumina HiSeq 2500		–	189,713	Leaves	Zhu et al. (2017)
F_1 (*A. roylei* and *A. fistulosum*)	Illumina HiSeq 2000	PRI 91021–8	–	10,361	Leaves of 4–8-week-old plants	Scholten et al. (2016)

Table 10.2 Number of members in different transcription factor families in monocots

Gene family	Bulb onion[a]	Rice (*indica*)[b]	Wheat[b]	Maize[b]	Sorghum[b]	*Phalaenopsis* (orchid)[b]
bHLH	162	169	324	308	297	96
NAC	147	158	263	189	180	85
ERF	132	138	181	204	172	91
MYB	121	121	263	203	145	108
WRKY	109	109	171	161	134	67
C2H2	105	113	224	179	140	93

[a]Khosa et al. (2016a)
[b]http://planttfdb.cbi.pku.edu.cn/ on 15-3-2017

family sizes in transcription factors is most likely due to gene duplication and deletions that alter gene family sizes (Guo 2013).

10.3 Evolution of the *FLOWERING LOCUS T* (*FT*) Gene Family in Bulb Onion

The *FT* gene family is found in all taxa of plants and encodes phosphatidylethanolamine-binding domain proteins (PEBP) (Kardailsky et al. 1999; Turck et al. 2008). The number of *PEBP*-like genes vary greatly between different plant species; in model plant *Arabidopsis*, six *FT*-like genes, *FT*, *TSF* (*Twin sister of FT*), *MFT* (*Mother of FT*), *BFT* (*Brother of FT*), *TFL1* (*Terminal flower like 1*) and *ATC* (*Arabidopsis thaliana relatives* of *centroradialis*) have been found, whereas in rice, 13 *FT*-like genes have been identified (Table 10.3) (Turck et al. 2008; Zheng et al. 2016b). Phylogenetic analyses suggest that the PEBP gene family can be divided into three subfamilies: *MFT*, *TFL1*, and *FT* (Karlgren et al. 2011). The *FT*- and *TFL1*-like genes are highly conserved in sequence, but exhibit antagonistic functions: *FT* acts as an activator of flowering, whereas *TFL1* acts as a repressor (Turck et al. 2008; Wickland and Hanzawa 2015). *FT* is produced in the leaves under inductive environmental conditions, and the protein is transported to the shoot apical meristem for flowering initiation (Kardailsky et al. 1999; Turck et al. 2008). The average number of *FT*-like genes in monocots is approximately six times higher than in eudicots (Table 10.3). The expansion of *FT*-like gene families in recent lineages might be due to tandem and segmental duplication in their genomes (Zheng et al. 2016b). In bulb onion, we identified six *FT*-like genes, and their phylogenetic relationship with other monocot FTs revealed that these *FT*-like genes belong to the *FT*-like group (Lee et al. 2013).

FTs can act as universal flowering signals in plants, but some members are involved in a diverse range of functions (Pin and Nilsson 2012). *FT* genes have been found to play an important role in the regulation of poplar growth, heterosis in tomato, stomata opening, potato tuberization and bulbing in onions (Hsu et al. 2011; Krieger et al. 2010; Kinoshita et al. 2011; Navarro et al.

Table 10.3 Number of *FT*-like genes and their functions in different plant species

Species	*FT*-like genes	Functions	References
Bulb onion	6	Bulbing and flowering	Lee et al. (2013)
Rice	13	Flowering	Zheng et al. (2016b)
Maize	15	Flowering	Zheng et al. (2016b)
Sorghum	11	Flowering	Zheng et al. (2016b)
Potato	3	Tuberization and flowering	Navarro et al. (2011)
Poplar	2	Bud cessation and growth	Böhlenius et al. (2006)

2011; Lee et al. 2013). Three *FT*-like genes of bulb onion (*AcFT1*, *AcFT2* and *AcFT4*) influence flowering time in the *Arabidopsis ft-1* mutant (Lee et al. 2013). *AcFT1* acts as a mobile signal to vigorously promote flowering, whereas *AcFT4* represses flowering, and *AcFT2* only marginally alters *Arabidopsis* flowering time. Expression analysis of *FT*-like genes in bulb onion indicates that, during the juvenile stage of the bulb onion life cycle, *AcFT4* is upregulated and acts antagonistically to repress *AcFT1* under both non-inductive, short day (SD) and inductive long day (LD) length conditions. When the onion plant reaches a certain developmental stage, under inductive day length conditions, *AcFT4* is downregulated, and *AcFT1* is upregulated to induce bulbing. Further, *AcFT4*-overexpressing onion plants never form bulbs and have a similar appearance to leeks; *AcFT1*-overexpressing plantlets form bulb-like structures whilst still in tissue culture. Expression and functional studies indicate that *AcFT1* acts as a promoter and that *AcFT4* act as an inhibitor of bulbing. *AcFT2* might act as a flowering promoter as it is only expressed in vernalized and flowering onion plants (Lee et al. 2013). Among other alliums, expression of *FTs* in garlic (*Allium sativum* L.) has been studied and show that *AsFT1* is expressed at higher levels in bulbing plants, and *AsFT4* acts antagonistically to it, however, it is not known whether they are involved in bulbing regulation (Shalom et al. 2015). B-box sequences (regions which determine *FT* function) of *FT1*, *FT2* and *FT4* among different *Allium* species show a high degree of conservation (Khosa et al. unpublished). This indicates that the role of *FTs* in the regulation of storage organ formation and flowering might be conserved in different alliums.

10.4 Circadian Clock and Bulbing

There are a number of parallels between the photoperiodic control of onion bulb formation and the photoperiodic induction of flowering in other plants (Brewster 2008). The circadian clock enables plants to respond to changes in seasons and to regulate different development processes (Sanchez and Kay 2016). In the photoperiodic flowering pathway of *Arabidopsis*, expression of clock genes *GIGANTEA* (*GI*) and *FLAVIN-BINDING*, *KELCH REPEAT*, *F-BOX1* (*FKF1*) coincides in long days (LD), enabling the GI-FKF1 proteins to form a complex that binds to CYCLING DOF FACTOR 1 and 2 (CDF1 and CDF2), targeting them for ubiquitin-dependent degradation, thereby releasing *CONSTANS* (*CO*) from repression. CO then accumulates in LD and initiates transcription of *FT*, which in turn initiates flowering after the FT protein moves from the leaves to the shoot apical meristem (Sawa et al. 2007; Johansson and Staiger 2015). In the majority of plant species, single copies of *GI* and *FKF1* are found, but in plants with a highly duplicated genome, such as maize and orchid, they occur in multiple copies (Table 10.4). In bulb onion, there are two copies of *GI* and one copy of *FKF1*. This indicates that certain genomic regions in bulb onion genome have undergone duplication, leading to multiple copies of some genes, but not all (Taylor et al. 2010). Similar to bulb onion, potato responds to day length for the formation of underground storage organs (tubers), and it has been shown that *StGI* and *StFKF1* regulate circadian clock genes to activate the potato *FT* orthologue, *StSP6A*, expression for tuber initiation (Kloosterman et al. 2013). This indicates that photoperiodic pathway genes have evolved to regulate diverse developmental stages, such as underground storage organ formation. Similar to *Arabidopsis*, bulb onion *AcGI* and *AcFKF1* show maximum expression at dusk; however, it is not yet known whether they are involved in bulbing (Taylor et al. 2010; Lee et al. 2013). Furthermore, *AcFKF1* is differentially expressed in SD and LD onions, and this might be responsible for day length sensitivity and timing-of-bulbing differences between SD and LD onions, but it needs further investigation (Taylor et al. 2010).

Table 10.4 The number of *GIGANTEA* and *FKF1* homologues identified in various species

Plants	*GIGANTEA*	*FKF-1*	References
Arabidopsis	1	1	Higgins et al. (2010)
Rice	1	1	Murakami et al. (2007), Higgins et al. (2010)
Brachypodium	1	1	Higgins et al. (2010)
Barley	1	1	Higgins et al. (2010)
Maize	2	2	Dong et al. (2012)
Soybean	3	2	Li et al. (2013)
Medicago	1	–	Kim et al. (2013)
Bulb onion	2	1	Taylor et al. (2010)
Oncidium orchid	2	2	Chang et al. (2011)
Cymbidium sinense	2	–	Zhang et al. (2013)

10.5 Flower Development

The genes involved in the specification of floral organ identity and flower development have been identified and characterized in a wide range of plant species. These studies indicate that common mechanisms of flower development occur across plant species (Thomson et al. 2017). The majority of floral identity genes belong to the MADS box gene family, with a few exceptions, such as *APETALA2* (*AP2*) (Ng and Yanofsky 2001; Becker and Theißen 2003; Wellmer et al. 2014; Thomson et al. 2017). These floral organ identity genes are grouped into five classes (A, B, C, D and E) based on their role in flower development. Genes in class A + E specify sepals, A + B + E petals, B + C + E stamens, C + E carpels, and D ovules, have been identified in different plant species (Wellmer et al. 2014; Thomson et al. 2017). The diversification (number and expression pattern) of MADS box genes involved in flower development contributes towards the floral diversity in the plant kingdom (Irish and Litt 2005; Heijmans et al. 2012). Alliums as a genus are notable and highly valued for their floral diversity (Hanelt 1990).

The bulb onion flower consists of fives whorls of floral organs viz: outer perianth, inner perianth, outer stamens, inner stamens, and carpels (Jones and Emsweller 1936). Recently, we identified various genes involved in flower development including MADS box genes (Khosa et al. 2016a). The phylogenetic relationship of bulb onion transcripts encoding ABCDE-like genes with those of other monocots indicates these transcripts belong to the ABCDE model (Fig. 10.1). Transcripts encoding B class genes *AP3* and *PI* show a close relationship with each other, while C and D class transcripts cluster together in the same clade. Bulb onion E class genes, *AGAMOUS LIKE6* (*AGL6*) and *SEPALLATA3*-like (*SEP3*-like), grouped in separate clades. *AP3* and *PISTILLATA* (*PI*) are members of a paralogous gene lineage, whereas C and D class genes are thought to have arisen from gene duplications during the evolution of angiosperms. E class genes diverged from other MADS box genes most recently (Becker and Theißen 2003; Theißen et al. 2016; Chanderbali et al. 2016). The number of genes encoding class B proteins is higher in bulb onion and other Asparagales (orchids) than in grasses and *Arabidopsis* (Table 10.5). Different studies indicate that ABCDE genes underwent duplication in Asparagales to result in diverse floral structures (Mondragón-Palomino 2013; Tsai et al. 2014; Cai et al. 2015; Otani et al. 2016; Dodsworth 2017). The bulb onion MADS-box genes are highly expressed in floral organs (Khosa et al. 2016a). In the future, it would be desirable to study the expression of ABCDE genes to have a better understanding of their role in flower development and whether any of the paralogous members play a divergent role.

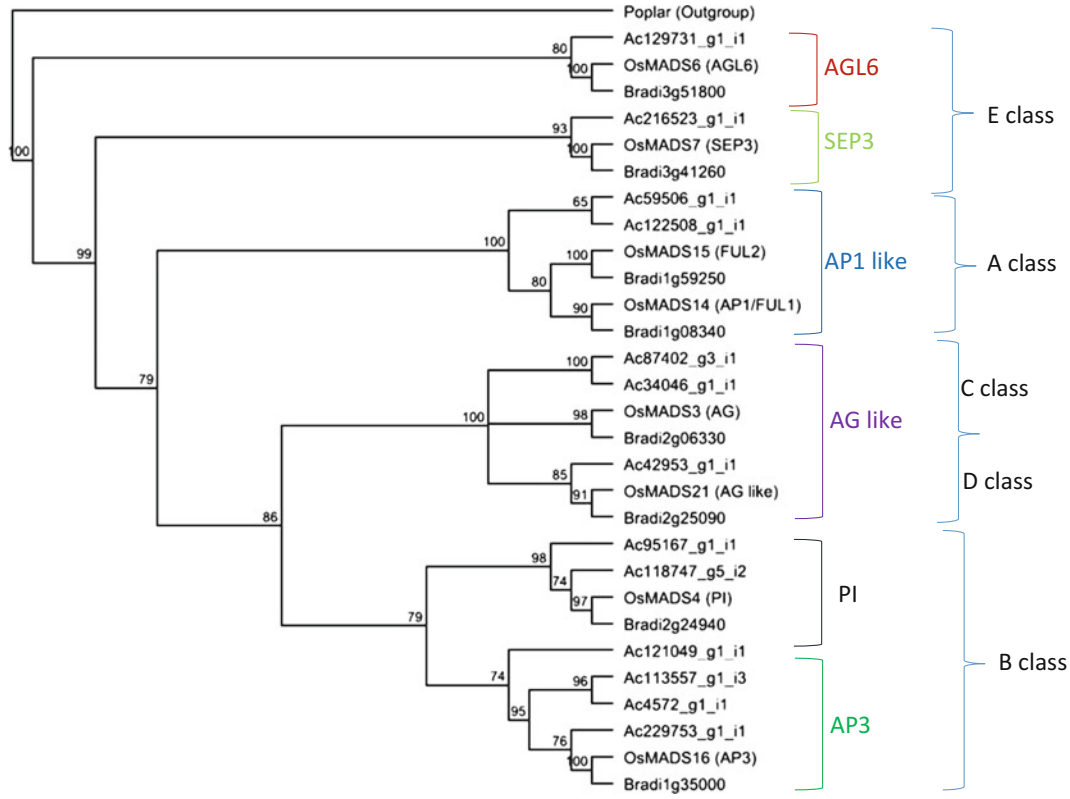

Fig. 10.1 Phylogenetic relationship of bulb onion transcripts encoding ABCDE-like genes with those of other monocots

Table 10.5 The distribution of gene homologues involved in ABCDE model of flower development in different plant species

Class	Gene name	Bulb onion	*Erycina pusilla* (orchid)	*Cymbidium sinense* (orchid)	*Arabidopsis*	Rice
A	*AP1/FUL*	2	3	1	4	4
B	*AP3*	4	3	6	1	1
B	*PI*	2	1	1	1	2
C	*AG*	3	4	3	4	5
E	*AGL6*	1	3	1	2	2
E	*SEP3*	1	2	1	1	2

Adapted from Khosa et al. (2016a), Lin et al. (2016), Xu and Kong (2007)

LEAFY (LFY) is a plant-specific transcription factor involved in floral meristem identity and in the initiation of floral development in plants (Moyroud et al. 2010). In genus *Allium*, the *LFY* orthologue has been identified in bulb onion, shallot, bunching onion and garlic (Rotem et al. 2007; Yang et al. 2016). The *LFY* homologs of different *Allium* species exhibit a high level of sequence conservation and a close phylogenetic relationship. In garlic, two alternatively spliced transcripts of *LFY* have been identified, and only the unspliced variant is associated with the transition from vegetative to floral organ differentiation (Rotem et al. 2007). Similar to other

species, *LFY* in garlic and bulb onion is highly expressed during floral transition and floral organ development (Neta et al. 2011; Yang et al. 2016). *Arabidopsis* plants overexpressing *AcLFY* exhibit early flowering and a series of morphological malformations (Yang et al. 2016). In the future, it would be interesting to determine the interaction of *LFY* with other floral integrators to activate genes involved in floral development.

10.6 Male Sterility and Fertility Restoration

Hybrid development of bulb onion varieties became economically feasible with the discovery of cytoplasmic male-sterility (CMS) systems. Various physiological and genetic studies have given us better insights into the origin, distribution and factors regulating CMS in different plant species (Chen and Liu 2014). In many species, most restorer genes (restorers of fertility, Rf or *Rf*) cloned so far belong to pentatricopeptide repeat (PPR) protein family, which contain a characteristic PPR motif of 35 amino acids (Schnable and Wise 1998; Chase 2007; Chen and Liu 2014; Islam et al. 2014). In bulb onion, a wide range of molecular markers have been identified for the identification of male-sterile and maintainer (male-fertile sister) lines (Khosa et al. 2016a). The floral parts of male-fertile and -sterile plants were used for transcriptome sequencing (RNA-Seq), and bulked segregant analysis (BSA) to identify candidate genes involved in fertility restoration (Kim et al. 2015). Candidate genes involved in male fertility restoration gene (*Rf*) in bulb onion have been identified.

A total of 483 transcripts containing the PPR motif have been identified, and among them, 41 transcripts of *Rf*-like PPR genes were found. The number of PPR and *Rf*-like genes identified in bulb onion is consistent with that in other plant species (Fujii et al. 2011; Sykes et al. 2017). However, the bulb onion *Rf*-like genes identified so far are not in linkage disequilibrium (LD) with the male sterility locus (*Ms*). Instead of a PPR-like gene, *AcPMS1* (a regulator of DNA mismatch repair, MMR) seems to be the most plausible candidate gene responsible for the restoration of male fertility in onion CMS. In *Arabidopsis* and tomato mismatch repair genes are responsible for reduced pollen development and induction of CMS (Li et al. 2009; Dion et al. 2007; Sandhu et al. 2007). Also, in other crop plants, various non-PRR genes control fertility restoration. (Chen and Liu 2014). It is possible that the real causal gene for fertility restoration has been missed during screening for homozygous SNPs and differentially expressed genes. Deep RNA, or whole genome, sequencing of mitochondrial sequences from male-sterile and fertile plants is required for better understanding of the evolution and molecular basis of CMS in onion (Kim et al. 2016).

The identification of fertile garlic flowering clones enabled researchers to identify the genetic basis of male sterility. RNA seq of male-fertile and male-sterile developing flowers led to the identification of >16,000 differentially expressed genes. The genes involved in the development of reproductive organs such as *AP3*, *MMD1*, *MS2* and *glycerol-3-phosphate acyltransferase 2* (GPAT2) show high expression in male-fertile plants, whereas genes involved in energy flow, respiration and mitochondrial functions show high expression in male-sterile plants (Shemesh-Mayer et al. 2015). Similar results have been reported during RNA-Seq of Japanese bunching onion male-sterile and fertile plants (Liu et al. 2016). These studies suggest that respiratory restrictions and/or non-regulated programmed cell death of the tapetum can lead to energy deficiency and consequent pollen abortion (Shemesh-Mayer et al. 2015; Liu et al. 2016). In future, functional characterization of the putative role of these genes in male sterility and fertility restoration will give us new clues about CMS system in genus *Allium*.

10.7 Flavonoid Biosynthesis

RNA-Seq data has been used to identify key genes in flavonoid production (Khosa et al. 2016b; Schwinn et al. 2016). Onions have a wide

variation in bulb colour due to variation in the production of different flavonoid compounds, such as red anthocyanins, pale yellow flavanols, and bright yellow chalcones. Anthocyanins, flavanols, and chalcones are all products of the flavonoid biosynthetic pathway. The key genes involved in anthocyanin biosynthesis pathway such as *anthocyanidin synthase* (*ANS*) gene, *dihydroflavonol 4-reductase* (*DFR*), chalcone synthase (*CHS*) and *chalcone isomerase* (*CHI*) have been identified, and mutation in these genes leads to different bulb colour (Masuzaki et al. 2006; Khar et al. 2008). In the bulb onion genome, *DFR* and *ANS* occur in single copies, however, in some other plants they are found in multiple copies (Table 10.6). The multiple copies of these genes also show differential expression in different organs, and different catalytic activities, which indicates that gene duplication might play a role in diversifying the functions of DFR enzymes (Huang et al. 2012).

Anthocyanin synthesis is regulated by a transcriptional activation complex consisting of *R2R3-MYB*, bHLH and a WD-repeat (WDR) proteins. This complex acts directly upon the promoters of flavonoid biosynthetic genes (Xu et al. 2015). Recently, four *R2R3-MYB* factors that putatively regulate flavonoid production have been identified in bulb onion transcriptomic datasets (Schwinn et al. 2016). The *MYB* genes from bulb onion and lilies cluster with the SG1/PAP1 clade and these sequences are absent in Poaceae and Orchidaceae indicating divergence of these transcription factors within Asparagales. The *MYB1* and anthocyanin biosynthetic genes, *DFR* and *CHS*, exhibit high expression in red onions. Transient overexpression, and knockdown experiments of the *MYB1* gene leads to induction and inhibition of anthocyanin production, respectively (Schwinn et al. 2016). Bulb onion *MYB1*, overexpressed in transgenic garlic, results in strong red pigmentation in the callus, leaves and leaf bases, but not in control plants. Onion *MYB1* is closely related to that of dicots and complements the anthocyanin *MYB* mutant of snapdragon (Schwinn et al. 2016).

10.8 Sulphur Assimilation and Metabolism

The genes involved in sulphate assimilation pathway are present in photosynthetic organisms, fungi, and many bacteria for the synthesis of sulphur-containing amino acids, and a range of other metabolites such as glucosinolates and alliins (Takahashi et al. 2011). Different *Allium* species have a high content of organosulphur compounds, and in recent years the genomic basis of sulphur assimilation has been studied, especially in bulb onion. Various key genes involved in organosulphur biosynthesis pathways have been reported in bulb onion, garlic and

Table 10.6 The distribution of key gene homologues involved in the anthocyanin biosynthesis pathway

Species	*Dihydroflavonol-4-reductase*	*Anthocyanidin synthase*	References
Arabidopsis	1	1	https://www.arabidopsis.org/
Tomato	1	1	Bongue-Bartelsman et al. (1994)
Potato	1	1	De Jong et al. (2003)
Medicago	2	1	Xie et al. (2004), Pang et al. (2007)
Lotus japonicus	6	–	Shimada et al. (2005)
Rice	1	2	Shih et al. (2008)
Populus trichocarpa	2	2	Huang et al. (2012)
Bulb onion	1	1	Kim et al. (2004a, b)

bunching onion (Brewster 2008; McManus et al. 2012; Kamenetsky et al. 2015; Sun et al. 2016).

In the sulphur assimilation pathway, sulphite reductase (SiR) is a key enzyme involved in the reduction of sulphite for the synthesis of sulphur compounds in bulb onion and other plants (Takahashi et al. 2011; McManus et al. 2012). Bulb onion have single functional *SiR* copy, whereas the rice and poplar genomes have two copies (Kopriva 2006; McManus et al. 2012). Further, phylogenetically, the *Allium* SiR groups with those of Eudicots; the SiR genes of Poales form an independent clade. This is consistent with large-scale phylogenetic relationships, based on EST and genome sequences (Kuhl et al. 2004). Recently, a large number of isoforms encoding different enzymes involved in sulphur metabolism have been identified in a garlic transcriptome dataset (Kamenetsky et al. 2015). The majority of these genes were expressed in a wide range of organs except reproductive organs. This indicates that cysteine sulphoxides are synthesized in leaves and roots, and then translocated to the underground storage organ (Kamenetsky et al. 2015).

10.9 Future Prospects

Over the last decade or so, large amount of genomic data has been generated for different plant species providing us with insights into the genome, gene family and evolution. Advances in technology have led to significant reductions in the cost of generating this data, and offer new avenues for gene discovery. While there is still limited genomic data for *Allium* species such as bulb onion. Transcriptomic datasets have now been generated for bulb onion, and other *Allium* species, leading to gene discovery and marker development. These resources have been utilized to identify the different gene families involved in regulation of bulbing, flowering, bulb colour, fertility restoration and sulphur metabolism. Insights into the evolution and neofunctionalization of genes have been garnered, and there is a lot more to discover yet. In coming years, RNA-Seq along, with whole genome sequencing, will lead to a comprehensive understanding of the different gene families in bulb onion and the *Allium* genus as a whole.

References

Baldwin S, Revanna R, Thomson S, Pither-Joyce M, Wright K, Crowhurst R, Fiers M, Chen L, Macknight R, McCallum J (2012) A toolkit for bulk PCR-based marker design from next-generation sequence data: application for development of a framework linkage map in bulb onion (*Allium cepa* L.). BMC Genom 13:637

Baldwin S, Revanna R, Pither-Joyce M, Shaw M, Wright K, Thomson S, Moya L, Lee R, Macknight R, McCallum J (2013) Genetic analyses of bolting in bulb onion (*Allium cepa* L). Theor App Genet 127:535–547

Becker A, Theißen G (2003) The major clades of MADS-box genes and their role in the development and evolution of flowering plants. Mol Phylogenet Evol 29:464–489

Böhlenius H, Huang T, Charbonnel-Campaa L, Brunner AM, Jansson S, Strauss SH, Nilsson O (2006) CO/FT regulatory module controls timing of flowering and seasonal growth cessation in trees. Science 312:1040–1043

Bongue-Bartelsman M, O'Neill SD, Tong Y, Yoder JI (1994) Characterization of the gene encoding dihydroflavonol 4-reductase in tomato. Gene 138:153–157

Brewster JL (2008) Onions and other vegetable *Alliums*. CABI, Oxfordshire, UK

Cai J, Liu X, Vanneste K, Proost S, Tsai W-C, Liu K-W, Chen L-J, He Y, Xu Q, Bian C, Zheng Z, Sun F, Liu W, Hsiao Y-Y, Pan Z-J, Hsu C-C, Yang Y-P, Hsu Y-C, Chuang Y-C, Dievart A, Dufayard J-F, Xu X, Wang J-Y, Wang J, Xiao X-J, Zhao X-M, Du R, Zhang G-Q, Wang M, Su Y-Y, Xie G-C, Liu G-H, Li L-Q, Huang L-Q, Luo Y-B, Chen H-H, Peer YV, Liu Z-J (2015) The genome sequence of the orchid *Phalaenopsis equestris*. Nat Genet 47:65–72. https://doi.org/10.1038/ng.3149

Chanderbali AS, Berger BA, Howarth DG, Soltis PS, Soltis DE (2016) Evolving ideas on the origin and evolution of flowers: new perspectives in the genomic era. Genetics 202:1255–1265. https://doi.org/10.1534/genetics.115.182964

Chang YY, Chu YW, Chen CW, Leu WM, Hsu HF, Yang CH (2011) Characterization of oncidium 'Gower Ramsey' transcriptomes using 454 GS-FLX pyrosequencing and their application to the identification of genes associated with flowering time. Plant Cell Physiol 52:1532–1545. https://doi.org/10.1093/pcp/pcr101 Epub 2011 Jul 23

Chase CD (2007) Cytoplasmic male sterility: a window to the world of plant mitochondrial–nuclear interactions. Trends Genet 23:81–90

Chen L, Liu YG (2014) Male sterility and fertility restoration in crops. Annu Rev Plant Biol 65:579–606

De Jong WS, De Jong DM, De Jong H, Kalazich J, Bodis M (2003) An allele of dihydroflavonol 4-reductase associated with the ability to produce red anthocyanin pigments in potato (*Solanum tuberosum* L.). Theor Appl Genet 107:1375–1383

Dion E, Li L, Jean M, Belzile F (2007) An *Arabidopsis MLH1* mutant exhibits reproductive defects and reveals a dual role for this gene in mitotic recombination. Plant J 51:431–440

Dodsworth S (2017) Petal, Sepal, or Tepal? B-genes and monocot flowers. Trends Plant Sci 22:8–10. https://doi.org/10.1016/j.tplants

Doebley JF, Gaut BS, Smith BD (2006) The molecular genetics of crop domestication. Cell 127:1309–1321

Dong Z, Danilevskaya O, Abadie T, Messina C, Coles N, Cooper M (2012) A gene regulatory network model for floral transition of the shoot apex in maize and its dynamic modeling. PLoS ONE 7(8):e43450. https://doi.org/10.1371/journal.pone.0043450

Duangjit J, Bohanec B, Chan AP, Town CD, Havey MJ (2013) Transcriptome sequencing to produce SNP-based genetic maps of onion. Theor Appl Genet 126:2093–2101

Fritsch RM, Friesen N (2002) Evolution, domestication and taxonomy. In: Rabinowitch HD, Currah L (eds) Allium crop science: recent advances. CABI, Wallingford, UK, pp 5–30

Fujii S, Bond CS, Small ID (2011) Selection patterns on restorer-like genes reveal a conflict between nuclear and mitochondrial genomes throughout angiosperm evolution. Proc Natl Acad Sci USA 108:1723–1728

Galsurker O, Doron-Faigenboim A, Teper-Bamnolker P, Daus A, Fridman Y, Lers A, Eshel D (2017) Cellular and molecular changes associated with onion skin formation involvement of programmed cell death. Front Plant Sci 7:2031. https://doi.org/10.3389/fpls.2016.02031

Guo YL (2013) Gene family evolution in green plants with emphasis on the origination and evolution of *Arabidopsis thaliana* genes. Plant J 73:941–951. https://doi.org/10.1111/tpj.12089

Han J, Thamilarasan SK, Natarajan S, Park J-I, Chung M-Y, Nou I-S (2016) De novo assembly and transcriptome analysis of bulb onion (*Allium cepa* L.) during cold acclimation using contrasting genotypes. PLoS ONE 11:e0161987. https://doi.org/10.1371/journal.pone.0161987

Hanelt P (1990) 'Taxonomy, evolution and history' in onions and allied crops. In: Rabinowitch HD, Brewster JL (eds) Botany, physiology, and genetics, vol 1. CRC Press, Boca Raton, pp 1–26

Heijmans K, Morel P, Vandenbussche M (2012) MADS-box genes and floral development: the dark side. J Exp Bot 63:5397–5404. https://doi.org/10.1093/jxb/ers233

Higgins JA, Bailey PC, Laurie DA (2010) Comparative genomics of flowering time pathways using *Brachypodium distachyon* as a model for the temperate grasses. PLoS ONE 5(4):e10065. https://doi.org/10.1371/journal.pone.0010065

Hsu CY, Adams JP, Kim H, No K, Ma C, Strauss SH, Drnevich J, Vandervelde L, Ellis JD, Rice BM, Wickett N, Gunter LE, Tuskan GA, Brunner AM, Page GP, Barakat A, Carlson JE, dePamphilis CW, Luthe DS, Yuceer C (2011) FLOWERING LOCUS T duplication coordinates reproductive and vegetative growth in perennial poplar. Proc Natl Acad Sci USA 108:10756–10761

Huang Y, Gou J, Jia Z, Yang L, Sun Y, Xiao X, Song F, Luo K (2012) Molecular cloning and characterization of two genes encoding dihydroflavonol-4-reductase from populus trichocarpa. PLoS ONE 7(2):e30364. https://doi.org/10.1371/journal.pone.0030364

Irish VF, Litt A (2005) Flower development and evolution: gene duplication, diversification and redeployment. Curr Opin Genet Dev 15:454–460

Islam MS, Studer S, Møller IM, Asp T (2014) Genetics and biology of cytoplasmic male sterility and its applications in forage and turf grass breeding. Plant Breeding 133:299–312

Jakse J, Meyer JDF, Suzuki G, McCallum J, Cheung F, Town CD, Havey MJ (2008) Pilot sequencing of onion genomic DNA reveals fragmented transposable elements, low gene densities, and significant gene enrichment after methy filtration. Mol Genet Genomics 280:287–292

Johansson M, Staiger D (2015) Time to flower: interplay between photoperiod and the circadian clock. J Exp Bot 66:719–730. https://doi.org/10.1093/jxb/eru441

Jones HA, Emsweller SL (1936) Development of the flower and macro-gametophyte of *Allium cepa*. J Agric Sci 10:415–428

Kamenetsky R, Faigenboim A, Mayer ES, Michael TB, Gershberg C, Kimhi S, Esquira I, Shalom SR, Eshel D, Rabinowitch HD, Sherman A (2015) Integrated transcriptome catalogue and organ-specific profiling of gene expression in fertile garlic (*Allium sativum* L.). BMC Genom 16:12

Kardailsky I, Shukla VK, Ahn JH, Dagenais N, Christensen SK, Nguyen JT, Chory J, Harrison MJ, Weigel D (1999) Activation tagging of the floral inducer FT. Science 286:1962–1965

Karlgren A, Gyllenstrand N, Källman T, Sundström JF, Moore D, Lascoux M, Lagercrantz U (2011) Evolution of the PEBP gene family in plants: functional diversification in seed plant evolution. Plant Physiol 156:1967–1977

Khar A, Jakse J, Havey MJ (2008) Segregations for onion bulb colors reveal that red is controlled by at least three loci. J Am Soc Hortic Sci 133:42–47

Khosa JS, Lee R, Bräuning S, Lord J, Pither-Joyce M, McCallum J, Macknight RC (2016a) Doubled Haploid 'CUDH2107' as a reference for bulb onion (*Allium cepa* L.) research: development of a transcriptome catalogue and identification of transcripts associated

with male fertility. PLoS ONE 11(11):e0166568. https://doi.org/10.1371/journal.pone.0166568

Khosa JS, McCallum J, Dhatt AS, Macknight R (2016b) Enhancing onion breeding using molecular tools. Plant Breed 135:9–20

Kim B, Kim K, Yang TJ, Kim S (2016) Completion of the mitochondrial genome sequence of onion (*Allium cepa* L.) containing the CMS-S male-sterile cytoplasm and identification of an independent event of the ccmF N gene split. Curr Genet 62:873–885

Kim MY, Kang YJ, Lee T, Lee S (2013) Divergence of flowering-related genes in three legume species. Plant Genome. https://doi.org/10.3835/plantgenome2013.03.0008

Kim S, Kim CW, Park M, Choi D (2015) Identification of candidate genes associated with fertility restoration of cytoplasmic male-sterility in onion (*Allium cepa* L.) using a combination of bulked segregant analysis and RNA-seq. Theor Appl Genet 128:2289–2299

Kim S, Binzel ML, Park SH, Yoo KS, Pike LM (2004a) Inactivation of DFR (dihydroflavonol 4-reductase) gene transcription results in blockage of anthocyanin production in yellow onions (*Allium cepa*). Mol Breed 14:253–263. https://doi.org/10.1023/B:MOLB.0000047770.92977.04

Kim S, Binzel ML, Yoo KS, Park S, Pike LM (2004b) Pink (P), a new locus responsible for a pink trait in onions (*Allium cepa*) resulting from natural mutations of anthocyanidin synthase. Mol Genet Genomics 272:18–27. https://doi.org/10.1007/s00438-004-1041-5

Kim S, Kim MS, Kim YM, Yeom SI, Cheong K, Kim KT, Jeon J, Kim S, Kim DS, Sohn SH, Lee YH, Choi D (2014) Integrative structural annotation of de novo RNA-Seq provides an accurate reference gene set of the enormous genome of the onion (*Allium cepa* L.). DNA Res 22:19–27

King JJ, Bradeen JM, Bark O, McCallum J, Havey MJ (1998) A low density genetic map of onion reveals a role for tandem duplication in the evolution of extremely large diploid genome. Theor Appl Genet 96:52–62

Kinoshita T, Ono N, Hayashi Y, Morimoto S, Nakamura S, Soda M, Kato Y, Ohnishi M, Nakano T, Inoue S, Shimazaki K (2011) FLOWERING LOCUST regulates stomatal opening. Curr Biol 21:1232–1238

Kloosterman B, Abelenda JA, Carretero-Gomez M, Oortwijn M, De Boer JM, Kowitwanich K, Horvath BM, Eck HJ, Smaczniak C, Prat S, Visser RGF, Bachem CWB (2013) Naturally occurring allele diversity allows potato cultivation in northern latitudes. Nature 495:246–250

Kopriva S (2006) Regulation of sulfate assimilation in *Arabidopsis* and beyond. Ann Bot 97:479–495

Krieger U, Lippman ZB, Zamir D (2010) The flowering gene SINGLE FLOWER TRUSS drives heterosis for yield in tomato. Nat Genet 42:459–463

Kuhl J, Cheung F, Yuan Q, Martin W, Zewdie Y, McCallum J, Catanach A, Rutherford P, Sink KC, Jenderek M, Prince JP, Town CD, Havey MJ (2004) A unique set of 11,008 expressed sequence tags (EST) reveals expressed sequence and genomic differences between monocot order *asparagales* and *poales*. Plant Cell 16:114–125

Lee R, Baldwin S, Kenel F, McCallum J, Macknight R (2013) FLOWERING LOCUS T genes control onion bulb formation and flowering. Nat Commun 4:2884

Li F, Zhang X, Hu R, Wu F, Ma J, Meng Y, Fu YF (2013) Identification and molecular characterization of FKF1 and GI homologous genes in soybean. PLoS ONE 8 (11):e79036. https://doi.org/10.1371/journal.pone.0079036

Li L, Dion E, Richard G, Domingue O, Jean M, Belzile F (2009) The Arabidopsis DNA mismatch repair gene PMS_1 restricts somatic recombination between homeologous sequences. Plant Mol Biol 69:675–684

Lin CS, Hsu CT, Liao DC, Chang WJ, Chou ML, Huang YT, Chen JJ, Ko SS, Chan MT, Shih MC (2016) Transcriptome-wide analysis of the MADS-box gene family in the orchid *Erycina pusilla*. Plant Biotechnol J 14:284–298. https://doi.org/10.1111/pbi.12383

Liu Q, Lan Y, Wen C, Zhao H, Wang J, Wang Y (2016) Transcriptome sequencing analyses between the cytoplasmic male sterile line and its maintainer line in welsh onion (*Allium fistulosum* L). IntJ Mol Sci 17:1058. https://doi.org/10.3390/ijms17071058

Liu Q, Wen C, Zhao H, Zhang L, Wang J, Wang Y (2014) RNA-Seq reveals leaf cuticular wax-related genes in Welsh onion. PLoS ONE 11:e113290. https://doi.org/10.1371/journal.pone.0113290

Liu T, Zeng L, Zhu S, Chen X, Tang Q, Mei S, Tang S (2015) Large-scale development of expressed sequence tag-derived simple sequence repeat markers by deep transcriptome sequencing in garlic (*Allium sativum* L.). Mol Breed 35:204

Martin LB, Fei Z, Giovannoni JJ, Rose JK (2013) Catalyzing plant science research with RNA-seq. Front Plant Sci 4:66. https://doi.org/10.3389/fpls.2013.00066/full

Masuzaki S, Shigyo M, Yamauchi N (2006) Complete assignment of structural genes involved in flavonoid biosynthesis influencing bulb color to individual chromosomes of the shallot (*Allium cepa* L.). Genes Genet Syst 81:255–263

McCallum J (2007) Onion. In: Kole C (ed) Genome mapping and molecular breeding in plants, vol 5. Springer, Heidelberg, Berlin, New York, pp 331–342

McManus MT, Joshi S, Searle B, Pither-Joyce M, Shaw M, Leung S, Albert N, Shigyo M, Jakse J, Havey MJ, McCallum J (2012) Genotypic variation in sulfur assimilation and metabolism of onion (*Allium cepa* L.) III. Characterization of sulfite reductase. Phytochem 83:34–42

Mondragón-Palomino M (2013) Perspectives on MADS-box expression during orchid flower evolution and development. Front Plant Sci 4:377. https://doi.org/10.3389/fpls.2013.00377

Moyroud E, Kusters E, Monniaux M, Koes R, Parcy F (2010) LEAFY blossoms. Trends Plant Sci 15:346–352

Murakami M, Tago Y, Yamashino T, Mizuno T (2007) Comparative overviews of clock-associated genes of *Arabidopsis thaliana* and *Oryza sativa*. Plant Cell Physiol 48:110–121

Mutz KO, Heilkenbrinker A, Lönne M, Walter JG, Stahl F (2013) Transcriptome analysis using next-generation sequencing. Curr Opin Biotechnol 24:22–30

Navarro C, Abelenda JA, Cruz-Oro E, Cuellar CA, Tamaki S, Silva J, Shimamoto K, Prat S (2011) Control of flowering and storage organ formation in potato by FLOWERING LOCUS T. Nature 478:119–122

Neta R, David-Schwartz R, Peretz Y, Sela I, Rabinowitch HD, Flaishman M, Kamenetsky R (2011) Flower development in garlic: the ups and downs of gaLFY expression. Planta 233:1063–1072. https://doi.org/10.1007/s00425-011-1361-8

Ng M, Yanofsky MF (2001) Function and evolution of the plant MADS-box gene family. Nat Rev Genet 2:186–195

Otani M, Sharifi A, Kubota S, Oizumi K, Uetake F, Hirai M, Hoshino Y, Kanno A, Nakano M (2016) Suppression of B function strongly supports the modified ABCE model in *Tricyrtis* sp. (Liliaceae). Sci Rep 6:24549. https://doi.org/10.1038/srep24549

Pang Y, Peel GJ, Wright E, Wang Z, Dixon RA (2007) Early steps in proanthocyanidin biosynthesis in the model legume *Medicago truncatula*. Plant Physiol 145:601–615

Pin P, Nilsson O (2012) The multifaceted roles of FLOWERING LOCUS T in plant development. Plant Cell Environ 35:1742–1755

Rajkumar H, Ramagoni RK, Anchoju VC, Vankudavath RN, Syed AUZ (2015) De novo transcriptome analysis of *Allium cepa* L. (onion) bulb to identify allergens and epitopes. PLoS ONE 10(8):e0135387. https://doi.org/10.1371/journal.pone.0135387

Ross-Ibarra J, Morrell PL, Gaut BS (2007) Plant domestication, a unique opportunity to identify the genetic basis of adaptation. Proc Natl Acad Sci 104(suppl 1):8641–8648

Rotem N, Shemesh E, Peretz Y, Akad F, Edelbaum O, Rabinowitch HD, Sela I, Kamenetsky R (2007) Reproductive development and phenotypic differences in garlic are associated with expression and splicing of LEAFY homologue gaLFY. J Exp Bot 58:1133–1141

Sanchez SE, Kay SA (2016) The plant circadian clock: from a simple timekeeper to a complex developmental manager. Cold Spring Harb Perspect Biol 8:12. https://doi.org/10.1101/cshperspect.a027748

Sandhu AP, Abdelnoor RV, Mackenzie SA (2007) Transgenic induction of mitochondrial rearrangements for cytoplasmic male sterility in crop plants. Proc Natl Acad Sci USA 104:1766–1770

Sawa M, Nusinow DA, Kay SA, Imaizumi T (2007) FKF1 and GIGANTEA complex formation is required for day-length measurement in *Arabidopsis*. Science 318:261–265

Schnable PS, Wise RP (1998) The molecular basis of cytoplasmic male sterility and fertility restoration. Trends Plant Sci 3:175–180

Scholten OE, van Kaauwen MPW, Shahin A, Hendrickx PM, Paul Keizer LC, Burger K, van Heusden AW, van der Linden CG, Vosman B (2016) SNP-markers in *Allium* species to facilitate introgression breeding in onion. BMC Plant Biol 16:187

Schwinn KE, Ngo H, Kenel F, Brummell DA, Albert NW, McCallum JA, Pither-Joyce M, Crowhurst NS, Eady C, Davies KM (2016) The Onion (*Allium cepa* L.) *R2R3-MYB* Gene *MYB1* regulates anthocyanin biosynthesis. Front Plant Sci 7:1865. https://doi.org/10.3389/fpls.2016.01865

Shalom SR, Gillett D, Zemach H, Kimhi S, Forer I, Zutahy Y, Tam Y, Teper-Bamnolker P, Kamenetsky R, Eshel D (2015) Storage temperature controls the timing of garlic bulb formation via shoot apical meristem termination. Planta 242:951–962

Shemesh-Mayer E, Ben-Michael T, Rotem N, Rabinowitch HD, Doron-Faigenboim A, Kosmala A, Perlikowski D, Sherman A, Kamenetsky R (2015) Garlic (*Allium sativum* L.) fertility: transcriptome and proteome analyses provide insight into flower and pollen development. Front Plant Sci 6:271. https://doi.org/10.3389/fpls.2015.00271

Shih CH, Chu H, Tang LK, Sakamoto W, Maekawa M, Chu IK, Wang M, Lo C (2008) Functional characterization of key structural genes in rice flavonoid biosynthesis. Planta 228:1043–1054

Shimada N, Sasaki R, Sato S, Kaneko T, Tabata S, Aoki T, Ayabe SA (2005) Comprehensive analysis of six dihydroflavonol 4-reductases encoded by a gene cluster of the *Lotus japonicus* genome. J Exp Bot 56:2385–2573

Sohn SH, Ahn YK, Lee TH, Lee JE, Jeong MH, Seo CH, Chandra R, Kwon YS, Kim CW, Kim DS, Won SY, Kim JS, Choi D (2016) Construction of a draft reference transcripts of onion (*Allium cepa*) using long-read sequencing. Plant Biotechnol Rep 10:383–390

Sun X, Zhou S, Meng F, Liu S (2012) De novo assembly and characterization of the garlic (*Allium sativum*) bud transcriptome by Illumina sequencing. Plant Cell Rep 31:1823–1828

Sun XD, Ma GQ, Cheng B, Li H, Qi S (2013) Identification of differentially expressed genes in shoot apex of garlic (*Allium sativum* L.) using Illumina sequencing. J Plant Stud 2:136–148

Sun XD, Yu XH, Zhou SM, Liu SQ (2016) De novo assembly and characterization of the Welsh onion (*Allium fitulosum* L.) transcriptome using Illumina technology. Mol Genet Genomics 291:647–659

Sykes T, Yates S, Nagy I, Asp T, Small I, Studer B (2017) In silico identification of candidate genes for fertility restoration in cytoplasmic male sterile perennial ryegrass (*Lolium perenne* L.). Genome Biol Evol 9:351–362

Takahashi H, Kopriva S, Giordano M, Saito K, Hell R (2011) Sulfur assimilation in photosynthetic organisms: molecular functions and regulations of transporters and assimilatory enzymes. Ann Rev of Plant Biol 62:157–184

Taylor A, Massiah AJ, Thimas B (2010) Conservation of *Arabidopsis thaliana* photoperiodic flowering genes in onion (*Allium Cepa* L.). Plant Cell Physiol 51:1638–1647

Theißen G, Melzer R, Rümpler F (2016) MADS-domain transcription factors and the floral quartet model of flower development: linking plant development and evolution. Development 143:3259–3271. https://doi.org/10.1242/dev.134080

Thomson B, Zheng B, Wellmer F (2017) Floral organogenesis: when knowing your ABCs is not enough. Plant Physiol 173:56–64

Tsai WC, Pan ZJ, Hsiao YY, Chen LJ, Liu ZJ (2014) Evolution and function of MADS-box genes involved in orchid floral development. J Syst Evol 52:397–410

Tsukazaki H, Yaguchi S, Sato S, Hirakawa H, Katayose Y, Kanamori H, Kurita K, Itoh T, Kumagai M, Mizuno S, Hamada M, Fukuoka H, Yamashita K, McCallum JA, Shigyo M, Wako T (2015) Development of transcriptome shotgun assembly-derived markers in bunching onion (*Allium fistulosum*). Mol Breed 35:55

Turck F, Fornara F, Coupland G (2008) Regulation and identity of florigen: FLOWERING LOCUS T moves center stage. Annu Rev Plant Biol 59:573–594

Varshney RK, Terauchi R, McCouch SR (2014) Harvesting the promising fruits of genomics: applying genome sequencing technologies to crop breeding. PLoS Biol 12:e1001883

Wellmer F, Graciet E, Riechmann JL (2014) Specification of floral organs in *Arabidopsis*. J Exp Bot 65:1–9

Wendel JF, Jackson SA, Meyers BC, Wing RA (2016) Evolution of plant genome architecture. Genome Biol 17:37. https://doi.org/10.1186/s13059-016-0908-1

Wickland DP, Hanzawa Y (2015) The *FLOWERING LOCUS T/TERMINAL FLOWER 1* gene family: Functional evolution and molecular mechanisms. Mol Plant 8:983–997

Xie DY, Jackson LA, Cooper JD, Ferreira D, Paiva NL (2004) Molecular and biochemical analysis of two cDNA clones encoding dihydroflavonol-4-reductase from *Medicago truncatula*. Plant Physiol 134:979–994

Xu G, Kong H (2007) Duplication and divergence of floral MADS-box genes in grasses: evidence for the generation and modification of novel regulators. J Integr Plant Biol 49:927–939

Xu W, Dubos C, Lepiniec L (2015) Transcriptional control of flavonoid biosynthesis by MYB-bHLH-WDR complexes. Trends Plant Sci 20:176–185

Yang C, Ye Y, Song C, Chen D, Jiang B, Wang Y (2016) Cloning and functional identification of the *AcLFY* gene in *Allium cepa*. Biochem Biophys Res Commun 473:1100–1105

Zhang C, Zhang H, Zhan Z, Liu B, Chen Z, Liang Y (2016) Transcriptome analysis of sucrose metabolism during bulb swelling and development in onion (*Allium cepa* L.). Front Plant Sci 7:1425. https://doi.org/10.3389/fpls.2016.01425

Zhang J, Wu K, Zeng S, Teixeira da Silva JA, Zhao X, Tian C, Xia H, Duan J (2013) Transcriptome analysis of *Cymbidium sinense* and its application to the identification of genes associated with floral development. BMC Genom 14:279

Zheng X, Tang S, Zhu S, Dai Q, Liu T (2016a) Identification of an NAC transcription factor family by deep transcriptome sequencing in onion (*Allium cepa* L.). PLoS ONE 11(6):e0157871. https://doi.org/10.1371/journal.pone.0157871

Zheng XM, Wu FQ, Zhang X, Lin QB, Wang J, Guo XP, Lei CL, Cheng ZJ, Zou C, Wan JM (2016b) Evolution of the PEBP gene family and selective signature on *FT*-like clade. J Syst Evol 54:502–510

Zhou SM, Chen LM, Liu SQ, Wang XF, Sun XD (2015) De novo assembly and annotation of the Chinese chive (*Allium tuberosum* Rottler ex Spr.) transcriptome using the Illumina platform. PLoS ONE 10(7):e0133312. https://doi.org/10.1371/journal.pone.0133312

Zhu S, Tang S, Tan Z, Yu Y, Dai Q, Liu T (2017) Comparative transcriptomics provide insight into the morphogenesis and evolution of fistular leaves in *Allium*. BMC Genom 18:60. https://doi.org/10.1186/s12864-016-3474-8