Towards molecular breeding of reproductive traits in cereal crops

Sangam Dwivedi1,*, Enrico Perotti2 and Rodomiro Ortiz3

1International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502 324, Andhra Pradesh, India
2CIMMYT/ANU Apomixis Project, Plant Cell Biology, Research School of Biological Sciences, Australian National University, GPO Box 475, Canberra, ACT 2601, Australia
3International Maize and Wheat Improvement Center (CIMMYT), Apdo. Postal 6-641, 06600 Mexico, D.F., Mexico

Summary
The transition from vegetative to reproductive phase, flowering per se, floral organ development, panicle structure and morphology, meiosis, pollination and fertilization, cytoplasmic male sterility (CMS) and fertility restoration, and grain development are the main reproductive traits. Unlocking their genetic insights will enable plant breeders to manipulate these traits in cereal germplasm enhancement. Multiple genes or quantitative trait loci (QTLs) affecting flowering (phase transition, photoperiod and vernalization, flowering per se), panicle morphology and grain development have been cloned, and gene expression research has provided new information about the nature of complex genetic networks involved in the expression of these traits. Molecular biology is also facilitating the identification of diverse CMS sources in hybrid breeding. Few Rf (fertility restorer) genes have been cloned in maize, rice and sorghum. DNA markers are now used to assess the genetic purity of hybrids and their parental lines, and to pyramid Rf or tms (thermosensitive male sterility) genes in rice. Transgene(s) can be used to create de novo CMS trait in cereals. The understanding of reproductive biology facilitated by functional genomics will allow a better manipulation of genes by crop breeders and their potential use across species through genetic transformation.

Introduction
Maize, rice and wheat are the major food crops worldwide, and together they contribute 74%–86% of the world’s cereal production: from an average of 737 million tons between 1962 and 1966 to 1882 million tons between 2002 and 2006 [Food and Agriculture Organization (FAO), 2006]. The world’s population is expected to reach 8.3 billion in 2030; the Earth will have to feed an extra two billion people, 90% of whom will live in developing countries (Bruinsma, 2003). It is therefore crucial to ensure not only that sufficient food is produced to feed this expanding population, but also that it is easily accessible to all. The Green Revolution more than doubled world grain production, thereby averting large-scale famine in the ‘developing world’. Crop productivity will need to be raised further to meet the growing demand for food crops.

Hybrid cultivars have made a significant contribution to the world’s food supply. In cereals, heterosis has been exploited in maize, rice, sorghum and pearl millet to produce high-yielding hybrids that, by far, dominate the global acreage for each crop. For example, about 95% of corn acreage in the USA is planted to hybrids that exhibit a 15% yield advantage relative to the best open-pollinated cultivars (Duvick, 1999), and more than half of the total rice-growing area in China is planted with hybrid rice cultivars (Wang et al., 2005a). The three- or two-line systems for producing hybrid seeds are well established in cereal crops. Cytoplasmic male sterility (CMS) and fertility restoration systems in three-line hybrids and environmentally sensitive genetic male sterility in two-line...
hybrids are used to produce hybrid seeds. Understanding the molecular mechanism of CMS and fertility restorer (Rf) genes and the molecular basis of environmentally induced genetic male sterility could improve the efficacy of hybrid technology.

Plants are sessile, and flowering is a key adaptive trait that contributes to their fitness by ensuring that they flower at an optimum time for pollination, seed development and dispersal. In addition to flowering per se, flower structure (more specifically male and female reproductive organs), panicle morphology, meiosis, pollination, fertilization and seed development are the major reproductive traits that influence grain yield. The mechanism of meiotic recombination and gamete formation ensures genomic stability and unfolds the genetic variation for targeted alterations in plant genomes. DNA marker-aided analysis of reproductive biology is still an emerging area. This contrasts with the situation for quality traits or resistance to biotic stresses, where considerable knowledge has been generated about genes and quantitative trait loci (QTLs), and DNA markers associated with these traits have been identified (Dwivedi et al., 2007, and references cited therein). Genomics can facilitate research on several key issues: unlocking the secret of what makes plants flower, the basis of structural divergence in panicle morphology amongst cereals, whether plants control meiosis, whether or not there are key gene(s) or a complex genetic network involved in the development of male and female organs, pollination and fertilization, and the development of seeds and their nutritional quality.

Global warming as a result of climate change will affect grain yields of cereals, particularly in non-temperate regions. By 2100, average global temperatures could rise by as much as 6 °C under the business-as-usual scenario (http://www.fao.org/docrep/005/y4252e/y4252e15.htm). High temperatures impair plant reproductive processes in particular, and thus may significantly reduce grain yields. For example, grain yield declined by 10% for each 1 °C increase in the growing season minimum temperature in the dry season, whereas the maximum temperature had no significant effect on rice yield (Peng et al., 2004).

Recent advances in plant biotechnology can help to address food security concerns. The structure of the cereal genomes and genes contained within them will aid geneticists and molecular biologists to understand cereal biology and help plant breeders to develop better products. Because of the remarkable degree of collinearity amongst grasses, rice can be used as a model plant for cereal genomics. This will greatly facilitate the isolation of agronomically important genes and our understanding of genome evolution. The rice genome has been fully sequenced, and this information can be used to obtain greater insights into the syntenic relationship between rice and the major cereals (Sasaki and Antonio, 2004).

Applied genomics for rice, wheat, maize, barley and pearl millet has provided useful new tools for breeding, and marker-assisted genetic enhancement has led to the development of several improved cultivars and advanced lines that show greater yields, possess good grain quality and minimize damage caused by pests and diseases. The introgression of beneficial genes under the control of specific promoters through transgenic approaches is another targeted approach to crop improvement. Genomic sciences have already identified many genes that have exciting potential for this purpose. Many genes and QTLs for important agronomic traits in cereals have been cloned, and are good candidates for transformation into other cultivars of the same crop, or into other cereals, without additional modification (Dwivedi et al., 2007, and references cited therein).

Flowering has been researched in detail in the model plant Arabidopsis, and in some grasses (Putterill et al., 2004; Corbesier and Coupland, 2006; Roux et al., 2006; Cockram et al., 2007). Inflorescence branching is a major yield trait in crop plants. It is controlled by the developmental fate of shoot apical meristems, and variation in branching patterns leads to diversity in inflorescences and affects crop yield; a number of mutants with altered inflorescence structure have been used to study the molecular and genetic control of grass inflorescence (Vollbrecht et al., 2005; Malcomber et al., 2006; Satoh-Nagasawa et al., 2006; Kellogg, 2007; Prusinkiewicz et al., 2007). A number of mutants defective in several aspects of organ development and meiosis in Arabidopsis have been used extensively to study the genetic and molecular basis of plant meiosis (Schwarzacher, 2003; Hamant et al., 2006; Ma, 2006). Information derived from this model plant is being used to study meiosis and crossover effects in crop plants including cereals. In contrast with flowering, inflorescence and plant meiosis, to the authors’ best knowledge, there have been only a couple of thorough reviews on the development of marker-assisted selection strategies for breeding hybrid rice (Xu, 2003), and on the nature and origins of the genes that determine CMS or plant mitochondrial–nuclear interactions (Chase, 2007), but there have been some recent reports on gene expression for dissecting the molecular basis of heterosis in maize, rice and wheat (Dwivedi et al., 2007, and references cited therein).

In this review, brief descriptions are given of the following: the model explaining floral organ development; the molecular basis of plant meiosis; molecular mechanisms regulating inflorescence architecture; mapping and cloning of genes or...
QTMs associated with phase transition (vegetative to reproductive growth) and variation in flowering and spike and grain development traits; gene expression with respect to floral organ development, pollination, fertilization and grain development; the molecular basis of CMS and the mapping and cloning of Rf genes; DNA markers for maintaining the purity of CMS and maintainor lines and hybrids; marker-assisted selection to pyramid tms (thermosensitive male sterility) and Rf genes; and transgenic approaches to hybrid breeding in cereals.

### Flowering

**ABC model and MADS-box genes in flower organ development**

Flowers normally consist of sepals, petals, stamens and pistils that arise in four concentric rings or whorls. The ABC model explaining flower development as the combination of different identity factors was postulated in the early 1990s (Bowman et al., 1991; Coen and Meyerowitz, 1991). The A signals specify whorl 1 or sepal formation, A + B direct whorl 2 or petals, B + C whorl 3 or stamens, and C whorl 4 or carpel formation (Figure 1). Moreover, A and C functions are mutually repressive (Jack, 2004). By using a mutagenesis approach, the genes responsible for the identity of different whorls have been cloned. For example, in Arabidopsis, the A function is triggered by APETALA1 (AP1) and APETALA2 (AP2), the B function by AP3 and PISTILLATA (PI) and the C function by AGAMOUS (AG). Ectopic expression of these genes in flowers is sufficient to trigger differentiation of the individual organs (Jack, 2004). Although AP2 defines whorls 1 and 2, transcripts accumulate in all the floral whorls (Jofuku et al., 1994). This suggests a more complex role for AP2 in organ determination. Null mutations in AP2 induce the formation of reproductive organs in place of petals and sepals (Bowman et al., 1991). Ectopic expression of AG in the outer whorls produces a similar phenotype to AP2 mutants (Mizukami and Ma, 1992). Moreover, loss of function of AG induces petal formation instead of stamens, together with additional flower formation in place of carpels (Bowman et al., 1991). This suggests that AP2 and AG have an antagonistic effect. Recently, it has been shown that microRNAs play a role in determining floral identity. MicroRNAs induce post-transcriptional degradation of genes by recognizing complementary sequences in the transcripts. For example, mir172 appears to be a negative regulator of AP2 (Zhao et al., 2007). Moreover, in petunia, two members of the mir169 family (BL and FIS) appear to regulate inner whorl identity, as null mutations increase C-function gene expression (Zhao et al., 2007). Future work will probably demonstrate the increasing importance of microRNAs in the regulation of flowering. In addition to the ABC model, a D-function gene, which controls the specific formation of ovules, has been defined (Figure 1). Two D-function genes, FBP7 (floral binding protein) and FBP11, specify placenta and ovule identity in petunia, and, when these genes are silenced, ovules fail to develop (Angenent et al., 1995). Moreover, ectopic expression of FBP11 produces ovules on sepals and petals (Colombo et al., 1995). In Arabidopsis, the D-function gene SEEDSTICK (STK) is solely expressed in ovules and corresponds to FBP7 and FBP11 (Favaro et al., 2003).

The latest addition to the ABC model is a fifth set of genes, the E class (Figure 1), necessary for normal flower development: SEPALLATA (SEP). In Arabidopsis, triple mutants in SEP1–3 produce flowers consisting solely of sepals (Pelaz et al., 2001). This suggests that the A function alone is responsible for sepal formation, whereas petals, stamens and carpels also require the E function. This has been confirmed experimentally, given that ectopic expression of SEP3 in Arabidopsis in conjunction with AP3 and PI produces petal-like organs, and over-expression of AG, AP3, PI and SEP3 produces stamen-like organs (Honma and Goto, 2001). The SEP proteins control organ identity through the formation of multimeric complexes with other MADS-box proteins, such as STK. Genetic evidence for this role has been demonstrated: a decrease in SEP activity led to a loss of ovule development (Favaro et al., 2003).

Flower morphology in cereals differs from flower morphology in other species, making comparisons in floral organ development...
difficult. Rice spikelets contain a single bisexual floret consisting of the pistil, stamens, two lodicules, palea and lemma. Similarly, maize spikelets have florets, a pistil, stamen, lodicules, palea and lemma (Bommert et al., 2005b). The lodicules correspond to whorl 2, the stamens to whorl 3 and the pistil to whorl 4 of Arabidopsis. It is not yet clear which structures correspond to whorl 1, although it might be assumed that the lemma/palea structures correspond to the sepal. The MADS-box proteins orthologous to all organ identity classes (ABCD and E) have been found in cereals (Bommert et al., 2005b; Zhao et al., 2006a).

In Arabidopsis, the A-class genes responsible for sepal and petal identity include AP1 and AP2. The grass genes belonging to the AP1 family have been found in maize (Zap1a and Zap1b), rice (OsMADS14, OsMADS15, OsMADS18 and OsMADS20) and barley (BM3 and BM8), and the VRN1 gene complex in wheat (Whipple et al., 2004; Zhao et al., 2006a). Using RNA profiling, Mena et al. (1995) showed that Zap1a was expressed in the outer whorls of the maize floret, but not in the stamen and pistil. Similarly, OsMADS14 expression was not detected in the reproductive organs, but did occur in the palea, lemma and lodicule (Koyozuka et al., 2000). Based on sequence homology, at least six A-class proteins have been found in wheat (Zhao et al., 2006a). Although there might be some redundancy, the size of this gene family suggests that these proteins could play different roles in monocotyledons than in dicotyledons. A demonstration of this additional complexity is the finding that the VRN1 gene, a Triticum AP1 orthologue, is induced by vernalization in winter wheats. As VRN1 controls flowering time on cold treatment, it appears that this gene in wheat not only defines outer whorl identity, but plays an additional role in floral transition (Yan et al., 2003). Similarly, over-expression of OsMADS18 in rice accelerates flowering time, whereas silencing of the gene produces no visible phenotype (Fornara et al., 2004).

Although the A function appears to be conserved between certain cereal genes and Arabidopsis, the situation may be more complex. Some AP1-like cereal genes do not appear to be involved in whorl 1 identity. This is not unique to cereals: no true A-function protein has been found in Antirrhinum, for example (Huijser et al., 1992). A better characterization of the numerous AP1-like genes is required to establish the presence of a defined A-function class of proteins in cereals.

The B-class genes AP3 and PI are responsible for petals and stamen identity in Arabidopsis. SILKY1 (SI1) from maize and OsMADS16 or SUPERWOMAN1 (SPW1) from rice are AP3 orthologues (Ambrose et al., 2000; Nagasawa et al., 2003). These genes are expressed during the early development of lodicules and stamens. Mutations in these genes induce homeotic changes from stamens into carpels and lodicules into lemma/palea-type organs. These data confirm that, in cereals, lodicules correspond to whorl 2 and stamens to whorl 3; the data also suggest that the lemma/palea corresponds to whorl 1. Moreover, the data also suggest that B function is conserved between monocotyledons and dicotyledons.

OsMADS2 and OsMADS24 from rice and Zmm16, Zmm18 and Zmm29 from maize appear to be homologues of PI (Münster et al., 2001; Kater et al., 2006). RNA profiling analysis showed that OsMADS2 is expressed in the inner whorls, whereas OsMADS24 is present in organs from whorls 2 and 3 (Koyozuka et al., 2000; Yadav et al., 2007). Antisense co-suppression of OsMADS4 in rice converts lodicules to lemma/palea and stamens to carpel structures, confirming its role in whorls 2 and 3, and demonstrating a conserved function between the Arabidopsis and rice orthologues (Kang et al., 1998). Interestingly, when OsMADS2 was silenced, no modification in stamen development was found, although lodicules were modified into lemma/palea (Prasad and Vijayraghavan, 2003). Wheat has two orthologues of PI (WP11 and WP12) and one orthologue of AP3 (WAP3) (Zhao et al., 2006a). When analysing the expression of the WP11 and WAP3 genes, Hama et al. (2004) showed RNA expression in lodicule and stamen primordia. Moreover, by studying alloplasmic wheat lines presenting homeotic changes of stamen into pistil-like structures, they demonstrated the role of these B-class proteins in floral organ identity.

In Arabidopsis, the B-class proteins AP3 and PI function as a heterodimer to regulate the transcription of floral genes. Whipple et al. (2004) showed that the maize orthologue of AP3 (SI1) formed a heterodimer with Zmm16, an orthologue of PI. In addition, they demonstrated that SI1 or Zmm16 could activate transcription by forming heterodimers with PI and AP3, respectively. Arabidopsis AP3 and PI mutants could also be rescued with the maize orthologues. This finding suggests that the B function in maize and Arabidopsis is conserved and grass lodicules correspond to petals.

The stamen and carpel organ identities in Arabidopsis are controlled by the C-class gene AG (Jack, 2004). Two C-class genes were found in maize: Zag1 and Zmm2 (Schmidt et al., 1993; Theissen et al., 1995). Zag1 is highly expressed in pistils, whereas Zmm2 is mainly expressed in stamen. Knockout in Zag1 induces indeterminate cellular growth in the ovary, but does not modify tassel identity (Mena et al., 1996). The lack of a tassel phenotype suggests that Zmm2 is redundant with Zag1 in stamen determination. Similar to maize, rice has an orthologue of Zmm2 (OsMADS3) and an orthologue of Zag1 (OsMADS58). RNA profiling has shown that OsMADS3 is mainly found in the ovule primordia, whereas OsMADS58 is
expressed in the stamen and carpel (Yamaguchi and Hirano, 2006). Ectopic expression of OsMADS3 under the control of an actin promoter transformed lodicules into stamen, which is predicted for C-class genes (Kyoizuka and Shimamoto, 2002). The C-class nature was confirmed by analysis of transgenic rice expressing antisense OsMADS3. These plants had lodicule-like stamens and carpellloid and stamenoid structures instead of carpels (Kang et al., 1998). RNA interference (RNAi) knockouts of OsMADS58 showed a phenotype similar to null mutants of Zag1. Floral meristem determinacy and carpel development were affected, but the effect on stamen identity was less severe than in OsMADS3 mutants (Yamaguchi and Hirano, 2006). Thus, Zmm2/ OsMADS3 plays a predominant role in whorl 3 determinism, whereas Zag1/OsMADS58 is mainly responsible for the formation of whorl 4.

Homology searches have shown that rice and maize have homologues of the D-class genes FBP7 and FBP11 that control ovule identity in petunia. These genes are Zag2, Zmm1 and Zmm25 in maize and OsMADS13 and OsMADS21 in rice (Kater et al., 2006). Expression analysis of OsMADS13 and Zag2 showed persistent expression in the ovules (Lopez-Dee et al., 1999). It is not clear whether these genes play a role in determining ovule identity, given that ectopic expression of OsMADS13 in rice and Arabidopsis does not induce ovule formation (Favaro et al., 2002).

The E function in Arabidopsis is regulated by SEP genes that appear to work in conjunction with the A, B, C and D MADS-box genes to specify floral organ identity. Numerous SEP homologues have been found in cereals. At least 12 SEP-like MADS boxes have been found in wheat (Zhao et al., 2006a). Characterization of TaMADS1 transcripts shows accumulation in floral primordia. Moreover, ectopic expression induces early flowering and alterations in floral organ development in Arabidopsis (Zhao et al., 2006b). Maize has seven E-class MADS-box genes (Whipple and Schmidt, 2006). Expression profiling of Zmm8 and Zmm14 shows transcription mainly in the meristems of the upper florets, suggesting a role in floral induction (Cacharrón et al., 1999). In rice, at least five SEP-like genes (OsMADS1, OsMADS5, OsMADS7, OsMADS8 and OsMADS19) have been found (Kater et al., 2006). Null mutations in the OsMADS1 or leafy hull sterile 1 (Ihs1) genes produce leafy lemma and palea. In addition, the lodicules and stamens are also modified into leafy lemma and palea structures (Agrawal et al., 2005; Bommert et al., 2005b). These observations demonstrate that OsMADS1, as well as the Arabidopsis SEP proteins, play a role in the E function in rice. Confirming this role of cereal E-class genes, over-expression of OsMADS8 in tobacco induced early flowering (Kang et al., 1997). In contrast, a loss of function of OsMADS5 does not produce significant disruptions in flower identity (Agrawal et al., 2005). These data, in conjunction with the complexity of the SEP family in cereals, suggest that these proteins may play additional roles in plant development.

MADS-box genes have been characterized in Arabidopsis, petunia and Antirrhinum. Recently, more effort has been invested in understanding the function of these genes in cereals. Although the picture may be more complex for the A-class genes, it is now clear that the B and C genes act similarly in monocotyledons and dicotyledons. Current data based on RNA profiling, protein interaction studies and phylogenetic analysis suggest that the D and E functions are also conserved throughout flowering plants, although gene redundancy has complicated the functional characterization of these classes of genes. As flower development is associated with traits such as grain yield and abiotic stress tolerance, a better understanding of the function of MADS-box proteins presents tremendous potential to improve crop yields.

Phase transition from vegetative to reproductive growth

Reproduction is a function of life. The transition from vegetative growth to flowering is essential for survival, and plants normally correlate the onset of flowering with suitable environmental conditions. Flower development and seed set are greatly impeded by stresses, such as drought or frost. The control of flowering is a highly regulated process that entails many environmental and endogenous signals that induce a change of determinism of the stem cells in the shoot apical meristem to give rise to reproductive structures in plants. It has been well documented that a mobile signal produced in the leaves, termed ‘florigen’, triggers floral transition in the shoot apical meristem (Corbesier and Coupland, 2006). Recently, the nature of this floral stimulus has been elucidated in Arabidopsis thaliana (Figure 2) and in rice. FLOWERING LOCUS T (FT), which encodes a small protein similar to receptor-associated factor (RAF)-kinase inhibitors, has been shown to be expressed in the leaves and then transported to the meristem. In this tissue, FT, together with the bZIP transcription factor FD expressed solely in the shoot apical meristem, activates floral identity genes, such as AP1 (Wigge et al., 2005). Using FT–green fluorescent protein (GFP) fusions, Corbesier et al. (2007) showed that the protein was translocated from the leaves via phloem companion cells to the meristem, and that this movement correlated with flowering. Tamaki et al. (2007) have demonstrated that the Hd3a protein, an orthologue of FT, induces flowering by moving from the leaf...
Figure 2 Summary of interactions regulating the phase transition from vegetative to reproductive growth in the model plant Arabidopsis thaliana (modified from Kobayashi and Weigel, 2007). Day length regulates flowering by activating FT expression via modulation of CO by the phytochromes and the cryptochromes. Cold temperatures induce flowering by activating FT expression through repression of FLC. FT is then transported via the phloem to the shoot apex, where it activates expression of API, SOC1 and LFY, thus triggering flower development.

to the shoot apical meristem. Working with Hd3a–GFP fusions, they showed that over-expression of the gene significantly accelerated flowering. Indeed, transgenic rice flowered after about 14–44 days, depending on the promoter upstream of the Hd3a genes, whereas wild-type rice usually took more than 50 days to flower in the same conditions. This suggests that the increase in expression of the mobile flowering signal FT induces early flowering, and appears to be a major player in the transition from vegetative to reproductive growth.

FT orthologues have been found in other crops (maize, barley and wheat) and appear to be involved in floral determination (Yan et al., 2006). Apart from FT, genetic analyses of Arabidopsis mutants showing alteration in flowering time have permitted the isolation of numerous genes involved in four basic interlinked pathways in floral transition: photoperiod, vernalization, gibberellin and autonomous pathways (Bernier and Perilleux, 2005).

Photoperiod pathway

The photoperiod pathway promotes flowering in Arabidopsis under long days (LD) (Boss et al., 2004). The plant perceives the day length through photoreceptors that absorb either red/far-red light for the phytochromes or blue light for the cryptochromes (Lin, 2000; Quail, 2002).

The transcription factor CONSTANS (CO) plays a major role in stimulating flowering on LD conditions through the transcriptional activation of FT (Samach et al., 2000). The expression of CO oscillates according to a circadian rhythm and has a peak at the end of the day under LD, which correlates with FT expression, and during the night under short days (SD) (Suarez-Lopez et al., 2001). Transgenic plants constitutively expressing CO flower earlier than wild-type controls and lose photoperiod sensitivity (Onouchi et al., 2000). In these plants, FT and SOC1 (SUPPRESSOR OF OVER-EXPRESSION OF CO1) are enhanced, demonstrating that CO promotes flowering by activating the expression of these genes (Samach et al., 2000). Valverde et al. (2004) showed that CO function was post-transcriptionally regulated through cryptochromes (CRY1 and CRY2) and through phytochromes A and B (phyA and phyB). Under red light and in the mornings, mutations in phyB produced an increase in the CO protein concentration, suggesting that the flowering repression activity of phyB is partly a result of a decrease in CO abundance. In contrast, mutations in the cryptochrome genes (CRY1 and CRY2) reduced the level of CO protein in the morning and under blue light. Thus, these proteins stabilize CO and compete with phyB to induce flowering. During LD and under far-red light, phyA appears to stabilize CO similarly to the cryptochromes, given that phyA mutants have less CO (Valverde et al., 2004). Thus, it appears that, under SD, CO degradation is enhanced via phyB, whereas, under LD, flowering is induced by the stabilization of CO through the cryptochromes and phyA.

The analysis of the genomic sequence of rice and other cereals has shown that many Arabidopsis genes that play a role in the photoperiod pathway and circadian clock have cereal orthologues. This suggests that the molecular mechanisms controlling flowering time in cereals and Arabidopsis are similar. Interestingly, barley CO shows a diurnal rhythm similar to that of the Arabidopsis orthologue (Cockram et al., 2007). Moreover, the rice orthologue of CO, Heading date 1 (Hd1), appears to play a major role in day length sensitivity (Hayama and Coupland, 2004). In this SD plant, loss of Hd1 produces an increase in Hd3a expression and induces early flowering under LD and delayed flowering under SD (Yano et al., 2000). Hayama and Coupland (2004) proposed that, under LD, the HD1 protein is activated by phytochrome and, contrary to Arabidopsis, inhibits flowering through inactivation of Hd3a. Under SD, in the absence of active phytochrome, Hd1 is expressed at night and triggers Hd3a expression and flowering. Not all mechanisms involved in the photoperiod
pathway in cereals are clearly understood, but similarities between genes and functions show that data from Arabidopsis can be used for enhancement of crops.

Vernalization pathway

Certain plants, such as winter wheat and Arabidopsis, require prolonged exposure to low temperatures to flower. In these plants, the vernalization response impedes flowering whilst temperatures are too low, permitting the plant to grow in winter. Several genes involved in the regulation of the vernalization pathway in Arabidopsis have been cloned (Sung and Amasino, 2005). The two major players are the MADS-box transcription factor FLOWERING LOCUS C (FLC) and FRIGIDA (FRI). FLC is a repressor of flowering and is strongly up-regulated by FRI, a coiled-coil domain protein with unknown biochemical function (Johanson et al., 2000). The repression of flowering occurs through the inhibition of FT and SOC1 expression (Mouradov et al., 2002). FLC is irreversibly down-regulated in a time-dependent manner during cold treatment, and therefore repression of flowering is lifted when warmer temperatures occur (Sheldon et al., 2000). The cold-induced decrease in FLC expression is stable throughout plant development and remains low even in organs formed at warmer temperatures. One exception is in gametes, where FLC repression is reset on sporogonesis or gametogenesis (Mouradov et al., 2002).

The analysis of late-flowering Arabidopsis mutants on vernalization treatment permitted the isolation of VERNALIZATION 1 (VRN1) and VERNALIZATION 2 (VRN2). The latter codes for a zinc-finger protein, part of the Polycomb (PcG) multimeric protein complex. This complex represses gene expression through the modification of chromatin structure (Sung and Amasino, 2005). To date, no orthologues of FLC or FRI have been found in cereals. Although PcG proteins are present in grasses, it remains unclear whether these play a role in the regulation of vernalization.

The absence, to date, of cereal orthologues of FLC and FRI suggests that vernalization in Arabidopsis and cereals is controlled through different pathways that converge at SOC1 and FT. The discovery in Arabidopsis of an FLC-independent vernalization pathway in which AGAMOUS-LIKE 24 (AGL24) plays a key role could provide answers as to how cereals modulate SOC1 and FT expression on cold treatment (Yu et al., 2002; Michaels et al., 2003). Indeed, based on sequence homology, cereal orthologues of AGL24 have been found.

In cereals such as wheat and barley, the response to vernalization is regulated by at least three genes, VRN1, VRN2 and VRN3, which have no relationship to the Arabidopsis genes of the same name (Cockram et al., 2007). VRN1 encodes an AP1-like MADS-box and is an orthologue of VRN-A1 from Triticum monococcum (Yan et al., 2003; Shitsukawa et al., 2007b). VRN2 is a transcription factor with a zinc finger and CCT (CO, CO-LIKE and TIMING OF CAB EXPRESSION 1) domain (ZCCT domain), also found in CO, and VRN3 is an orthologue of FT (Valverde et al., 2004; Yan et al., 2004). On prolonged exposure to cold, VRN1 is up-regulated, thus promoting flowering, whereas VRN2 is repressed (Yan et al., 2004).

In spring wheat characterized by dominant VRN1 alleles, the gene is constitutively expressed in plants with dominant spring alleles, therefore explaining flowering in the absence of cold treatment. In winter wheat, a MADS-box transcription factor, VEGETATIVE TO REPRODUCTIVE TRANSITION 2 (TaVRT2), represses VRN1 transcription by binding to a site in the promoter, the CarG-Box (Kane et al., 2007). When vernalization is complete, TaVRT2 transcripts are down-regulated, permitting the accumulation of VRN1 and flowering. In addition, down-regulation of VRN2 on vernalization leads to the induction of VRN1 (Yan et al., 2004). This down-regulation can be replaced by SD treatment, demonstrating a link between the photoperiod and vernalization pathways in wheat (Dubcovsky et al., 2006). This link was also found in barley grown in LD conditions, where VRN1 was expressed despite VRN2 expression (Trevaskis et al., 2006). VRN2 appears to play an integrative role, as its transcription is regulated by day length and by vernalization. Although the Arabidopsis gene FLC has no orthologues in cereals, VRN2 might play the same role in vernalization.

Gibberellin pathway

The Arabidopsis over-expressing genes involved in the gibberellin biosynthesis pathway are early flowering, independent of photoperiod. In contrast, a decrease in gibberellin levels retards flowering mainly under SD. Although the gibberellin pathway is different from other pathways, it appears to be involved in floral induction, mainly under SD (Bos et al., 2004). Gibberellin promotes flowering under SD through the induction of the LEAFY (LFY) transcription factor (Blazquez and Weigel, 2000). This protein plays a central role in the onset of flowering, LFY is expressed in leaf primordia, and a rapid increase in expression is correlated with floral transition in Arabidopsis during LD (Araki, 2001). Although LFY is not a direct target of CO, activation of this protein triggers a rapid increase in expression (Samach et al., 2000).

The Arabidopsis LFY protein is characterized by a conserved domain found in floriculal/leafy-like proteins. Although these proteins have been well studied in dicotyledons, not much is known of their function in monocotyledons. Null mutations in FLORICUAL/LEAFY-like 1 and 2 from maize produced...
defects in inflorescence development and structure, demonstrating that the LFY proteins play a conserved role (i.e. behave in the same way) in dicots and monocots (Bombilies et al., 2003).

Although the application of bioactive gibberellins accelerates floral initiation in wheat, sorghum, barley and rice, it remains unclear what transductional pathways are involved in this process (King and Evans, 2003).

The Green Revolution in the 1960s occurred as a result of farmers adopting new wheat cultivars that were shorter, photoperiod insensitive, higher yielding and with higher harvest indices. The reduced size came from mutations in genes such as Reduced height 1–3, responsible for gibberellin signalling (Peng et al., 1999). The importance of gibberellin as an essential regulator of cereal growth has been demonstrated; however, its role in flowering is still not clear.

Autonomous pathway

This pathway was identified in a series of Arabidopsis mutants that contained higher levels of FLC than wild-type plants and that were later flowering under all photoperiods (Mouradov et al., 2002). This delay in flowering is overcome by vernalization, suggesting a link between the vernalization and autonomous pathways. To date, seven proteins are known to be involved in the autonomous pathway in Arabidopsis [FCA, FY, FPA, FVE, luminidependens (LD), flowering late KH motif (FLK) and flowering locus D (FLD)]. All seven proteins have orthologues in cereals. Although the mode of action of the RNA-binding protein FCA is unclear, over-expression under the control of a constitutive promoter results in early-flowering Arabidopsis lines (Marquardt et al., 2006). As FCA physically interacts with FY, a WD repeat polyadenylation factor, it is postulated that the complex regulates FLC transcripts through the stabilization of mRNA (Marquardt et al., 2006). Constitutive expression of rice FCA (OsFCA) in fca Arabidopsis mutants rescued the late-flowering phenotype (Lee et al., 2005a), but no significant effect on FLC expression was obtained. Moreover, as no FLC orthologue has been found in cereals, FCA and the autonomous pathway might play different roles in the regulation of flowering in monocots and dicots.

The genetic and molecular bases of meiosis in plants

Meiosis, a specialized type of cell division, is critical in the life cycle of sexually reproducing organisms. In plants, specialized reproductive cells differentiate from somatic tissue. These cells then undergo a single round of DNA replication, followed by two rounds of chromosome division to produce haploid cells, which then undergo further rounds of mitotic division to produce the pollen grain and embryo sacs (Caryl et al., 2003). Chromosome pairing, synopsis and recombination are the major events in meiosis. Synapsis and recombination are tightly linked, and there are complex networks of interactions between them. Homologous chromosomes interact with each other and form bivalents. Chromosome pairing is largely dependent on the initiation and progression of recombination. Recombination is a highly conserved feature of meiosis that results in the formation of crossovers and non-crossover products. Recombination helps to ensure chromosome segregation and promotes allelic diversity. Errors in meiosis often lead to sterility (Pawlowski and Cande, 2005; Mézard et al., 2007).

Budding yeast (Saccharomyces cerevisiae), Caenorhabditis elegans and Drosophila are the model organisms used to study meiosis in fungi and animals, whereas Arabidopsis and, to some extent, maize are the models for understanding the molecular basis of plant meiosis. A basic knowledge about the Arabidopsis genes important for meiotic recombination, and their relationship to pairing and synopsis, meiotic progression, spindle assembly, chromosome separation and meiotic cytokinesis, has been uncovered and characterized. The molecular genetic studies in this model plant are also providing insights into meiosis that have not yet been recognized elsewhere in eukaryotes, including gene functions that might be unique to plants vs. those that are potentially shared with animals and fungi (Caryl et al., 2003; Hamant et al., 2006; Ma, 2006).

Crossing over is a key process for the accurate segregation of homologous chromosomes during meiosis. It tends to decrease near the centromeres and increase towards the telomeres (Anderson and Stack, 2002). For example, crossover frequency in wheat, maize and barley tends to increase with relative physical distance from the centromere, whereas in Welsh onion (Allium fistulosum) it clusters close to the centromeres; in Arabidopsis and tomato, the crossover distribution varies between and along chromosome arms, with no apparent rule. Moreover, within chromosome regions (cold and hot, which have significantly low and high crossover frequencies, respectively), crossover rates vary enormously from one kilobase to another in Arabidopsis, indicating the presence of several levels of control, each operating at different scales: chromosomal, regional (megabase) or local (kilobase) (Drouaud et al., 2007; Mézard et al., 2007). Gene-rich regions, for example in wheat, barley and maize, are more recombinationally active than gene-poor regions (Schnable and Wise, 1998). Two pathways of crossover have been identified: interference-sensitive crossover, which inhibits the
occurrence of another crossover in a distance-dependent manner; and interference-insensitive crossover, when both pathways exist in the plant kingdom, including Arabidopsis (Mercier et al., 2005). The ratio of interference-sensitive to interference-insensitive crossover differs between species, e.g. about 30% of crossovers escape the interference-sensitive mechanism in yeast and tomato, but only 15% escape it in Arabidopsis.

Recombination nodules are protein complexes associated with meiotic (pachytene) chromosomes. Studying recombination nodule structure and function provides insights into the processes by which meiotic recombination is regulated in eukaryotes. The two types of recombination nodule identified are early nodules and late nodules, with the former appearing at leptotene and persisting into early pachytene, and the latter appearing in pachytene and remaining into early diplotene. Both can be distinguished by their time of appearance and by characteristics such as shape, size, relative numbers and association with unsynapsed or synapsed chromosomal segments. Early nodule function is not clearly understood, but it may assist in research on DNA homology, synopsis, gene conversion and/or crossing over. Late nodules are well correlated with crossing over (Anderson and Stack, 2005).

Wheat is a disomic hexaploid, i.e. a polyploid that behaves as a diploid during meiosis. Chromosome pairing is restricted to homologous chromosomes, despite the presence of homologues in the nucleus (Sears, 1976). Using the Affymetrix wheat GeneChip®, Crismani et al. (2006) identified 1350 transcripts that were temporally regulated during the early stages of meiosis, many of these associated with chromatin condensation, synaptonemal complex formation, recombination and fertility. Thirty transcripts displayed at least an eightfold expression change between and including pre-meiosis and telophase II, with more than 50% of these having no similarities to known sequences in the National Center for Biotechnology Information (NCBI) and Institute for Genomic Research (TIGR) databases that could be used to study the molecular basis of pairing and recombination control in a complex polyploid such as wheat. Furthermore, Crismani et al. (2006) detected four transcripts expressed only in early meiotic stages, three of which showed no significant similarities to sequences in the NCBI and TIGR databases. These genes are prime candidates, as they may play a role in co-ordinating early meiotic recombination in wheat. Variations in the time for completion of meiosis have also been reported: meiosis in bread wheat anthers is completed within 24 h, in barley 39 h and in rye 51 h (Crismani et al., 2006, and references cited therein). There is also a high degree of asynchrony between male and female meiosis. For example, in Arabidopsis, male meiosis is completed before female meiosis reaches prophase 1 (Caryl et al., 2003).

A number of factors make plants suited to the analysis of gamete development: late germline specification, the non-lethality of mutations affecting gamete development and the large size of their chromosomes. Many genes with roles in gamete development in Arabidopsis, principally those for meiosis, recombination and DNA repair, have yielded novel information about the processes of gamete formation in higher plants (Wilson and Yang, 2004). The angiosperm female gametophyte typically consists of one egg cell, two synergid cells, one central cell and three antipodal cells. Each of these four cell types has unique structural features and performs unique functions, essential for reproduction. The gene regulatory networks conferring these four phenotypic states are largely uncharacterized. Using the Affymetrix ATH1 genome array for differential expression, real-time reverse transcriptase-polymerase chain reaction (RT-PCR) and Arabidopsis mutant determinant infertile 1 (dif1) ovules, which lack the female gametophyte, Steffen et al. (2007) identified 71 genes that exhibit decreased expression in dif1 ovules; when further validated using promoter–GFP fusions, they detected 11 genes that expressed exclusively in the antipodal cells, another 11 genes that expressed exclusively or predominantly in the central cell, 17 genes that were exclusive to or predominant in the synergid cells, one gene that expressed exclusively in the egg cell, and three genes that expressed strongly in multiple cells of the female gametophyte. These genes offer insights into the molecular processes involved in the female gametophyte, insights that serve as starting points to dissect the gene regulatory networks functioning during the differentiation of the four female gametophyte cell types. Furthermore, Ma et al. (2007a) used three male-sterile mutants, lacking a range of normal cell types resulting from a temporal progression of anther failure, in comparison with genotypes having fertile siblings, at four equivalent stages of anther development, to profile gene expression in dissected maize anthers using oligonucleotide arrays. They detected nearly 9200 sense and antisense transcripts, with the most diverse transcripts present at the pre-meiotic stage. By combining data sets from the comparisons between individual sterile and fertile anthers, they assigned candidate genes predicted to play important roles during maize anther development to stages and probable cell types. Comparative analysis with a data set of anther-specific genes from rice highlighted remarkable quantitative similarities in gene expression between these two grasses.
Pairing between wheat homeologous chromosomes is prevented by the expression of the Ph1 locus on the long arm of chromosome 5B. Suppressors with major effects were mapped as Mendelian loci on the long arms of Aegilops speltoides chromosome 3S and 7S, with the former designated as Su1-Ph1 and the latter as Su2-Ph1. A QTL designated as QPh.ucd-5S, with a minor effect, was also mapped on the short arm of chromosome 5S. Both Su1-Ph1 and Su2-Ph1 increased homeologous chromosome pairing, the former completely epistatic to the latter. The two genes acting together increased homeologous chromosome pairing to the same level as Su1-Ph1 acting alone. QPh.ucd-5S expression was additive to the expression of Su2-Ph1 for increased homeologous chromosome pairing. Based on these observations, Dvorak et al. (2006) hypothesized that the products of Su1-Ph1 and Su2-Ph1 affect pairing between homeologous chromosomes by regulating the expression of Ph1, but the product of QPh.ucd-5S may primarily regulate recombination between homologous chromosomes. Griffiths et al. (2006) localized Ph1 to a 2.5-Mb interstitial region of wheat chromosome 5B, containing a structure consisting of a segment of subtelomeric heterochromatin inserted into a cluster of cdc2-related genes. The correlation of the presence of this structure with Ph1 activity, together with the involvement of heterochromatin with Ph1 and cdc2 genes in meiosis, makes the structure a good candidate for the Ph1 locus. Taken together, the Ph1 locus prevents recombination between homeologous chromosomes in wheat, thereby furnishing a powerful tool for manipulating alien genes for wheat improvement. Incorporation of a Ph1 suppressor into wheat would greatly facilitate introgression of alien genes into wheat chromosomes via recombination between homeologous chromosomes, thus paving the way for the introgression of beneficial alleles from distant wheat relatives. In maize, phs1 is required for pairing to occur between homologous chromosomes (Pawlowski et al., 2004). This gene encodes a putative 347-amino-acid protein with a predicted molecular mass of 38 kDa. In rice, it is the PAIR1 (HOMOLOGOUS PAIRING ABERRATION IN RICE MEIOSIS1) gene that plays an essential role in the establishment of homologous chromosome pairing (Nonomura et al., 2004). The PAIR1 gene encodes a 492-amino-acid protein, which contains putative coiled-coil motifs in the middle, two basic regions at both termini and a potential nuclear localization signal at the C-terminus. Using a comparative genetic approach, Sutton et al. (2003) identified the rice genomic region syntenous to the region deleted in the wheat chromosome pairing mutant ph2a, and detected 218 wheat expressed sequence tags (ESTs) putatively located in the region deleted in ph2a.

Genetic gains in plant breeding depend on the generation and selection of new recombinants ensuing from crosses between chosen strains. The genetic variation released in segregating populations depends on the assortment of chromosomes and the amount of recombination within chromosomes. In addition to environmental factors, genotype effects influence meiotic recombination in plants: increased recombination occurs in pollen mother cells of barley, pearl millet and wheat, whereas recombination is greater in egg mother cells in the cases of rye, Brassica oleracea and Brassica nigra (Guzy-Wróbelska et al., 2007, and references cited therein). Differences in male–female recombination frequency can be potentially exploited for the construction of genetic linkage maps, fine mapping, map-based gene cloning or the development of alien substitution lines, as well as to avoid linkage drag.

Mapping and cloning genes or QTLs associated with variation in flowering

There are a number of reports in the literature regarding the identification of putative QTLs for variation in flowering time in barley, maize, rice, sorghum and wheat (Table 1); however, few studies have succeeded in the fine mapping and cloning of these QTLs. For example, Hd1, Hs3a, Hd6, EHD1 and Se5 were cloned in rice, with the first four promoting flowering under SD, and the latter involved in rice phytochrome synthesis (chromophore), regulating flowering in response to day length (Cockram et al., 2007). Two major photoperiod loci (Ppd-H1 and Ppd-H2) have been identified in barley, but only Ppd-H1 has been cloned, whereas none of the three photoperiod loci (Ppd-A1, Ppd-B1 and Ppd-D1) identified in wheat have been cloned. Interestingly, the major loci affecting photoperiod response in wheat and barley have been mapped to collinear positions on the short arm of the group 2 chromosomes (Cockram et al., 2007, and references cited therein). An understanding of the photoperiod pathway in Arabidopsis provides a source of candidate genes for Ppd loci in cereals. Hd1, Hs3a and Hd6 encode orthologues of CO and FT, and the α subunit of casein kinase 2 (CK2), which are well-characterized factors for flowering or the circadian clock in Arabidopsis (Izawa et al., 2003; Hayama and Coupland, 2004). However, rice Hd1 promotes flowering under SD, whereas Arabidopsis CO promotes flowering under LD (Izawa et al., 2003). Furthermore, 19 of the 62 consensus QTLs in maize are syntenic to rice and Arabidopsis flowering time QTLs or genes, whereas vgt-7f/h is allelic to vgt1 in maize (Table 1). A sorghum QTL associated with photoperiod sensitivity on linkage group (LG) H is an orthologue to Hd1,
Table 1 Quantitative trait loci (QTLs) associated with flowering time in barley, maize, rice, sorghum and wheat from 1967 to 2006

<table>
<thead>
<tr>
<th>Flowering time QTL</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Barley</strong></td>
<td></td>
</tr>
<tr>
<td>2–5 QTLs for heading, one in two environments</td>
<td>Baum et al. (2003)</td>
</tr>
<tr>
<td>Two major genes, eam8 and eam10, and two QTLs determine flowering; eam8 and eam10 mapped on long arm of <em>Hordeum</em> (barley) chromosome 1 (1HL) and 3HL, respectively, whereas QTL on 1HL and short arm of <em>Hordeum</em> (barley) chromosome 7 (7HS)</td>
<td>Börner et al. (2002a)</td>
</tr>
<tr>
<td>Five major genes and eight QTLs</td>
<td>Laurie et al. (1995)</td>
</tr>
<tr>
<td><strong>Maize</strong></td>
<td></td>
</tr>
<tr>
<td>vgt-Fp mapped (6 cM) on chromosome B; probably allelic to vgt1</td>
<td>Chardon et al. (2005)</td>
</tr>
<tr>
<td>Eight QTLs for GGD heat units to pollen shedding</td>
<td>Zhang et al. (2005b)</td>
</tr>
<tr>
<td>Six of the 62 consensus QTLs displaying major effect; 19 QTLs and genes syntenic to rice and Arabidopsis flowering</td>
<td>Chardon et al. (2004)</td>
</tr>
<tr>
<td><strong>Rice</strong></td>
<td></td>
</tr>
<tr>
<td>Seven QTLs for flowering across three environments</td>
<td>Khillare et al. (2005)</td>
</tr>
<tr>
<td>Nineteen QTLs related to vegetative and reproductive growth mapped on six chromosomes</td>
<td>Zhou et al. (2001)</td>
</tr>
<tr>
<td>Fourteen QTLs control flowering; <em>Hd1, Hd2, Hg5, Hd6 and Hg9</em> mapped as Mendelian factors; two tightly linked loci, <em>Hd3a</em> and <em>Hd3b</em>, detected in <em>Hd3</em> region</td>
<td>Yano et al. (2001)</td>
</tr>
<tr>
<td><strong>Sorghum</strong></td>
<td></td>
</tr>
<tr>
<td>Two major QTLs for maturity</td>
<td>Crasta et al. (1999)</td>
</tr>
<tr>
<td>Six major loci, <em>Ma1</em> to <em>Ma6</em>, controlling flowering and maturity</td>
<td>Rooney and Aydin (1999); Quinby (1967)</td>
</tr>
<tr>
<td><em>Ma3</em> mapped on LG A encoded by phytochrome B1</td>
<td>Childs et al. (1997)</td>
</tr>
<tr>
<td>A QTL on LG D assigned to <em>Ma1</em></td>
<td>Lin et al. (1995)</td>
</tr>
<tr>
<td><strong>Wheat</strong></td>
<td></td>
</tr>
<tr>
<td><em>Eps-A1</em> on chromosome 1A* in <em>Triticum monococcum</em> flanked by VatpC and Smp</td>
<td>Valárik et al. (2006)</td>
</tr>
<tr>
<td>Four QTLs for intrinsic earliness</td>
<td>Hanoqc et al. (2004)</td>
</tr>
<tr>
<td><em>Eps-SBL1</em> and <em>Eps-SBL2</em> mapped close to the centromere on 5BL; one homologous to barley on 5H</td>
<td>Tóth et al. (2003)</td>
</tr>
<tr>
<td>4–8 QTLs for ear emergence and 5–7 for flowering in 11 environments</td>
<td>Börner et al. (2002b)</td>
</tr>
<tr>
<td><em>Nse-3A</em> and <em>Nse-3Ae</em> mapped on chromosome 3A* and 5A*, respectively; former homologous to <em>eps3L</em> in barley and the latter to <em>Qet.ocs-SAl</em> in wheat</td>
<td>Shindo et al. (2002)</td>
</tr>
<tr>
<td><em>QEt.ocs-SAl</em> for earliness per se on 5AL with little influence on grain yield</td>
<td>Kato et al. (2002)</td>
</tr>
<tr>
<td>A major QTL on chromosome 2BS for heading, co-segregated with <em>Ppd-B1</em>, and another on 7BS corresponds to a QTL for earliness per se</td>
<td>Sourdille et al. (2000)</td>
</tr>
</tbody>
</table>

Table 2 DNA markers/quantitative trait loci (QTLs) associated with response to variation in photoperiod and vernalization in barley, sorghum and wheat from 2000 to 2004

<table>
<thead>
<tr>
<th>Marker/QTL information</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Barley</strong></td>
<td></td>
</tr>
<tr>
<td>Six AFLP markers closely linked to <em>Ppd-H1</em></td>
<td>Decousset et al. (2000)</td>
</tr>
<tr>
<td><strong>Sorghum</strong></td>
<td></td>
</tr>
<tr>
<td>Two QTLs on LG C and H associated with photoperiod sensitivity <em>sensu stricto</em>; former close to earliness, <em>Ef-1</em>, and the latter orthologue to <em>Hd1</em>, a major photoperiod sensitivity gene in rice</td>
<td>Chantereau et al. (2001)</td>
</tr>
<tr>
<td><strong>Wheat</strong></td>
<td></td>
</tr>
<tr>
<td>Four QTLs each for photoperiod and vernalization; two photoperiod-sensitive QTLs located at the same position as <em>Ppd-B1</em> and <em>Ppd-D1</em></td>
<td>Hanoqc et al. (2004)</td>
</tr>
<tr>
<td><em>Vrn-B1, Vrn-D1</em> and <em>Ppd-B1</em> on 5B, 5D and 2B, respectively; two types of genes for photoperiod sensitivity are known: one dependent on and the other independent of vernalization; a QTL for narrow-sense earliness close to <em>Ppd-B1</em></td>
<td>Shindo et al. (2003)</td>
</tr>
<tr>
<td><em>Vrn-B1</em> mapped on 5B and homologous to other wheat, rye and barley vernalization response genes</td>
<td>Leonova et al. (2003)</td>
</tr>
<tr>
<td>Three markers linked to <em>Vrn-B1</em> on 5BL; Xgwm408 closest in two populations</td>
<td>Barrett et al. (2002)</td>
</tr>
</tbody>
</table>

which is a major photoperiod sensitivity gene in rice. Likewise, photoperiod-sensitive QTLs are located at the same position as the major genes *Ppd-B1* and *Ppd-D1* in wheat, and *Vrn-B1* is homologous to other vernalization response genes in wheat, rye and barley (Table 2). The wheat gene *TaGI1*, which is involved in photoperiodic flowering, is an orthologue of GIGANTEA (*GI*) in Arabidopsis. *TaGI1* is expressed in leaves in a rhythmic manner under LD and SD conditions, and its rhythmic expression is regulated by photoperiods and circadian clocks (Zhao et al., 2005). A 6-bp insertion/
deletion in the Dwarf8 (D8idp) polymorphism is associated with flowering time under LD. In maize, the deletion allele shows an earlier flowering of 29 degree-days for inbred lines and 145 degree-days for landraces. Thus, Dwarf8 could be involved in maize climatic adaptation through the diversification of selection of the flowering time (Camus-Kulandaivelu et al., 2006).

Earliness per se (Eps) and ‘narrow sense’ earliness genes, often detected as QTLs, control flowering time independently of these environmental cues (photoperiod and vernalization induction pathways promote flowering), and are responsible for the fine tuning of flowering time. Several eps QTLs and genes have been reported in wheat and barley (Table 1), with some homologous to each other, thereby representing more sources of variation to breed for earliness. Moreover, there may be temperature-eps gene interaction, as shown in wheat Eps-A1 (Appendo and Slafer, 2003), which should be further investigated to eliminate the effect of growing temperature on the expression of eps genes, thereby enabling researchers to select for earliness per se under any thermal condition.

Szücs et al. (2006) mapped six photoperiod response QTLs and determined their positional relationship with respect to the phytochrome and crypochrome photoreceptor gene families and the vernalization regulatory genes Hv8BSMA, ZCCT-H and HvVRT-2. Of the six photoreceptors mapped (HvPhyA and HvPhyB to 4HS, HvPhyC to 5HL, HvCry1a, HvCry2-6HS and HvCry1b to 2HL), only HvPhyC coincided with a photoperiod response QTL. Hv8BSMA, ZCCT-H and HvVRT-2 are located in chromosome regions determining small photoperiod response QTL effects. Previously, they had mapped candidate genes for the 5HL VRN-H1 (Hv8BSMA) and 4HL VRN-H2 (ZCCT-H) loci and, in this study, HvVRT-2, the barley TaVRT-2 orthologue (a wheat-flowering repressor regulated by vernalization and photoperiod), to 7HS. Hv8BSMA and HvPhyC are closely linked on 5HL, and therefore are positional candidates for the same photoperiodic effect. von Zitzewitz et al. (2005) had already mapped candidate genes for the 5HL VRN1 (Hv8BSMA) and 4HL VRN-H2 (ZCCT-H) loci in barley. Phytochromes are a family of red/far-red light-absorbing photoreceptors that control plant development and metabolic processes in response to changes in light. Transgenic rice lines containing Arabidopsis PhyA, together with a tissue-specific rice rbcS promoter, showed reduced plant height but more panicles per plant, and produced 6%–21% more grain than non-transgenic plants, demonstrating the potential to manipulate light signal transduction pathways to enhance grain productivity (Garg et al., 2006a).

The FT gene plays a central role in integrating flowering signals in Arabidopsis, given that its expression is regulated antagonistically by the photoperiod and vernalization pathways (Boss et al., 2004; Jack, 2004; Bäurle and Dean, 2006). FT belongs to a family of six genes characterized by a phosphatidylethanolamine-binding protein domain. Thirteen FT-like sequences, designated as OsFTL1–OsFTL13, have been reported in rice, with Hd3a corresponding to OsFTL2 (Izawa et al., 2002; Chardon and Damerval, 2005; Zhang et al., 2005a). More recently, Faure et al. (2007) found five FT-like genes in barley. HvFT1, HvFT2, HvFT3 and HvFT4 are highly homologous to OsFTL2 (the Hd3a QTL), OsFTL1, OsFTL10 and OsFTL12, respectively, but there is no rice equivalent for HvFT5. HvFT1 was highly expressed under LD at the time of the morphological switch of the shoot apex from vegetative to reproductive growth, and HvFT2 and HvFT4 were expressed later in plant development. HvFT1 was thus identified as the main barley FT-like gene involved in the switch to flowering. Mapping of the HvFT genes should therefore provide important sources of flowering time variation in barley. HvFT1 is a candidate for VRN-H3, a dominant mutation giving precocious flowering; HvFT3, a candidate for Ppd-H2, is a major QTL affecting flowering time under SD in barley. In plants, seasonal changes in day length are perceived in leaves, which initiate long-distance signalling that induces flowering at the shoot apex. Florigen, the mobile signal, moves from an induced leaf to the shoot apex and causes flowering. Understanding the nature of the mobile flowering signal would provide a key insight into the molecular mechanism of floral induction. The Arabidopsis FT gene is a candidate for encoding florigen (Corbesier et al., 2007), and the protein encoded by Hd3a, a rice orthologue of FT, moves from the leaf to the shoot apical meristem and induces flowering in rice, suggesting that the Hd3a protein may be the rice florigen (Tamaki et al., 2007).

Vernalization, the induction of flowering by prolonged exposure to low temperature, is the major determinant of flowering in wheat and barley. Vernalization-dependent types are commonly called winter wheat and winter barley, whereas vernalization-independent types are called spring wheat and spring barley. VRN-H1 (Sh2, Sgh2), VRN-H2 (Sh, Sgh) and VRN-H3 (Sh3, Sgh3) in barley, VRN-A1 (VRN-1) and VRN-A2 (VRN-2) in diploid einkorn wheat (Triticum monococcum), and VRN-A1 (Vrn1), VRN-B1 (Vrn2), VRN-D1 (Vrn3) and VRN-B3 (Vrn-B4) in bread wheat promote flowering after vernalization (Cockram et al., 2007, and references cited therein). VRN-H1, VRN-A1, VRN-A1, VRN-B1 and VRN-D1 encode MADS-box and AP1-like, VRN-H2 and VRN-A2 encode B-box and CCT domain, and VRN-H3 and VRN-B3 encode putative kinase inhibitor proteins. Vrn3 is linked completely to a gene similar to Arabidopsis FT (Yan
et al., 2006). FT induction in leaves results in a signal that promotes flowering. Transcript levels of the barley and wheat orthologues, designated as HvFT and TaFT, respectively, are significantly higher in plants that are homozygous for the dominant Vrn3 alleles (early flowering) than in plants that are homozygous for the recessive Vrn3 alleles (late flowering). In wheat, the dominant Vrn3 allele is associated with the insertion of a retro-element in the TaFT promoter, whereas, in barley, mutations in the HvFT first intron differentiate plants with dominant and recessive Vrn3 alleles. Winter wheat plants transformed with the TaFT allele carrying the promoter retro-element insertion flower earlier than non-transgenic plants, supporting the identity between TaFT and Vrn-B3. This research confirms that wheat and barley FT genes are responsible for natural allelic variation in the vernalization requirement, providing additional sources of adaptive diversity in these crops.

Dubcovsky et al. (2006) showed that the vernalization requirement in some of the photoperiod-sensitive winter wheat cultivars can be replaced by interrupting LD by 6 weeks of SD, and that this interruption is associated with SD down-regulation of the VRN2 flowering repressor. In addition, SD down-regulation of VRN2 at room temperature is not followed by the up-regulation of the meristem identity gene VRN1 until plants are transferred to LD, which contrasts with the VRN1 up-regulation observed after the VRN2 down-regulation by vernalization, suggesting the existence of a second VRN repressor. Further analysis of natural VRN1 mutants has revealed that the CArG-box located in the VRN1 promoter is the most likely regulatory site for the interaction with this second repressor. Up-regulation of VRN1 under SD in accessions carrying mutations in the CArG-box resulted in an earlier initiation of spike development, compared with other genotypes. However, even the genotypes with CArG-box mutations required LD for normal and timely spike development. SD accelerates flowering in photoperiod-sensitive winter cultivars, suggesting that wheat was initially an SD–LD plant, and that strong selection pressures during domestication and breeding resulted in the modification of this dual regulation. The down-regulation of the VRN2 repressor by SD is probably part of the mechanism associated with the SD–LD regulation of flowering in photoperiod-sensitive winter wheat.

More recently, Kane et al. (2007) have shown that a MADS-box transcription factor, TaVRT2, in wheat is also associated with the vernalization response in a manner similar to TaVRN2. Using transient expression assays in Nicotiana benthamiana, they showed that TaVRT2 acts as a repressor of TaVRN1 transcription. TaVRT2 binds the CArG motifs in the TaVRN1 promoter and represses its activity in vivo. In contrast, TaVRN2 does not bind the TaVRN1 promoter and has no direct effect on its activity, but can enhance the repression effect of TaVRT2, suggesting that a repressor complex regulates the expression of TaVRN1. In winter wheat, TaVRT2, TaVRN2 and TaVRN1 transcripts accumulate in the shoot apical meristem and young leaves, and temporal expression is consistent with TaVRT2 and TaVRN2 being repressors of floral transition, whereas TaVRN1 is an activator. Non-vernalized spring wheat grown under SD accumulates TaVRT2 and shows a delay in flowering, suggesting that TaVRT2 is regulated independently by photoperiod and low temperatures. The results demonstrate that TaVRT2, in association with TaVRN2, represses the transcription of TaVRN1.

Vgt1 is a QTL involved in the control of the transition of the apical meristem from the vegetative to the reproductive phase (flowering) in maize. Using positional cloning and association mapping, Salvi et al. (2007) resolved Vgt1 to an approximately 2-kb non-coding region positioned 70 kb upstream of an Ap2-like transcription factor involved in flowering time control. Vgt1 functions as a cis-acting regulatory element, as indicated by the correlation of the Vgt1 alleles with transcript expression levels of the downstream gene. In addition, within Vgt1, they identified evolutionary conserved non-coding sequences across the maize–sorghum–rice lineages, supporting the notion that changes in distant cis-acting regulatory regions are a key component of plant genetic adaptation throughout breeding and evolution.

WAP1 (wheat AP1) is a key gene in the regulatory pathway controlling the transition phase from vegetative to reproductive growth in common wheat, and is an orthologue of VRN1 in diploid einkorn wheat (Murai et al., 2003). More recently, Shitsukawa et al. (2007a) identified a mutant in einkorn wheat, maintained vegetative phase (mvp), which makes it unable to transit from the vegetative to the reproductive phase. This mvp mutation resulted from the deletion of the VRN1 coding and promoter regions, demonstrating that WAP1/VRN1 is an indispensable gene for phase transition in wheat. Expression analysis of flowering-related genes in mvp plants indicated that wheat Gl, CO and SOC1 genes either act upstream of or in a different pathway from WAP1/VRN1. The MADS-box gene SOC1 is an integrator of flowering pathways in Arabidopsis. Shitsukawa et al. (2007b) isolated a wheat orthologue of SOC1, wheat SOC1 (WSOC1), which is expressed in young spikes, but preferentially expressed in leaves. Expression starts before the phase transition and is maintained during the reproductive growth phase. Over-expression of WSOC1 in transgenic Arabidopsis caused early flowering under SD conditions, suggesting that WSOC1
functions as a flowering activator in *Arabidopsis*. *WSOC1* expression is not affected by either vernalization or photoperiod, whereas it is induced by gibberellin at the seedling stage. Moreover, *WSOC1* is expressed in transgenic wheat plants in which *WAP1* expression is co-suppressed, indicating that *WSOC1* acts in a pathway different from the *WAP1*-related vernalization and photoperiod pathways.

The mapping and cloning of the genes associated with flowering in the model plants *Arabidopsis* and rice should enable researchers to alter flowering for the better adaptation of cereal crops to existing environments and to new conditions predicted to emerge as a result of climate change. Some MADS-box genes activate or repress flowering in rice. Using rice *OsMADS* cDNA clones connected to the maize *ubiquitin* promoter in the sense orientation, Jeon et al. (2000) introduced the constructs into rice plants by Agrobacterium-mediated transformation. The transgenic plants exhibited early flowering and dwarf phenotypes that correlated well with the transcript levels detected in these plants. Interestingly, transgenics containing *OsMADS1* attached with the *nopaline synthase* promoter shortened the heading time, but with a mild reaction on dwarfism without any pleiotropic effects on other agronomic traits. Thus, rice MADS-box genes could be used as sources of early flowering and dwarfing traits in monocots. Tadege et al. (2003) identified a *OsSOC1* gene in rice that encodes a MADS-domain protein related to the *Arabidopsis* gene *AtSOC1*, with 97% amino acid similarity in the MADS domain. *OsSOC1* is located on top of the short arm of chromosome 3, tightly linked to the heading date locus *Hd9*. Although it expressed in vegetative tissues, its expression was elevated at the time of floral initiation and remained uniformly high thereafter, similar to the expression pattern of *AtSOC1*. The constitutive expression of *OsSOC1* in *Arabidopsis* results in early flowering, suggesting that the rice gene is a functional equivalent of *AtSOC1*. However, Tadege et al. (2003) were not able to identify FLC-like sequences in the rice genome, but found that ectopic expression of the *Arabidopsis* *FLC* delayed flowering in rice, and the up-regulation of *OsSOC1* at the onset of flowering initiation was delayed in *AtFLC* transgenic lines.

The CO gene of *Arabidopsis* plays a key role in the photoperiodic flowering pathway. Nemoto et al. (2003) isolated three kinds of *CO/Hd1* (rice orthologue of *CO*) homologues from the A, B and D genomes of hexaploid wheat: *TaHd1-1*, *TaHd1-2* and *TaHd1-3*, respectively. They are highly similar to each other and to *Hd1*, and contain two conserved regions: two zinc-finger motifs and a CCT domain like *CO/Hd1* located on the long arm of the homologous chromosome 6. When the *TaHd1-1* genomic clone was intro-

duced into a rice line deficient in *Hd1* function, the transgene complemented the functions of *rice* *Hd1*, i.e. it promoted heading under SD conditions and delayed it under LD and natural conditions, indicating that *Hd1* proteins from SD and LD plants share common structures and functions. A floral control gene *LFY* from *Arabidopsis* was introduced into rice using the cauliflower mosaic virus 35S promoter, and caused the transgenic rice plant to flower 26–34 days earlier than wild-type plants, with yield loss and panicle abnormality. This suggests that floral regulatory genes from *Arabidopsis* are useful tools for heading date improvement in cereal crops (Zuhua et al., 2000).

The reciprocal recognition and flowering time effects of genes introduced into either *Arabidopsis* or rice suggest that some components of the flowering pathways may be shared. Moreover, the wheat *TaHd1-1* gene can complement rice *Hd1* and function normally in the rice background. These genes thus provide the potential to genetically manipulate the flowering time, employing either well-characterized *Arabidopsis* genes or heterologous genes from other cereals, using comparative genomics in cereals.

**Inflorescence architecture and development**

**Genetic and molecular mechanisms regulating inflorescence architecture**

Plants undergo a series of profound developmental changes throughout their life cycle. They can be broadly grouped into changes occurring from juvenile to adult leaf formation, vegetative to inflorescence development, and inflorescence to floral meristem initiation (Chuck and Hake, 2005). Grass inflorescences and flowers have characteristic structures distinct from those in *Arabidopsis* (Bommert et al., 2005b).

The major differences are related to changes in the number of branches, the numbers of orders of branches and the amount of axis elongation (Doust and Kellog, 2002). Information on the developmental transition in plants (Chuck and Hake, 2005) and grass inflorescence diversity and genetics (Bommert et al., 2005b; Malcomber et al., 2006; Doust, 2007; Kellog, 2007) has been published elsewhere. In this section, we highlight only those developmental phenomena and genes involved that, if brought together, might result in substantial improvement in the productivity of the major cereal food crops.

Maize and rice are excellent model systems for studying the genetics and molecular characterization of inflorescence architecture and development in cereals because of the subtle differences in inflorescence, the fact that a large number of mutants affect inflorescence and floral development, the
high density of genetic linkage maps, and the fact that the rice and maize genomes have been fully or partially sequenced (McSteen et al., 2000; Vollbrecht et al., 2005; Dwivedi et al., 2007, and references cited therein). A combination of internal and external environmental signals affect, for example, the transition from vegetative to reproductive stage (see the section on ‘Flowering’ above), and control the major genetic pathways regulating inflorescence architecture and development. Using mutants with discrete phenotypes and molecular biology tools, a number of maize, rice and sorghum genes regulating the transition from vegetative to reproductive phase, or those related to inflorescence structure and development, have been mapped or cloned, and their molecular functions determined (Table 3). Interestingly, some of the genes associated with inflorescence or panicle architecture traits are known orthologues to *Arabidopsis* genes, i.e. maize *fasciated ear2* (*fea2*) and *thick tassel dwarf1* (*td1*) to CLAVATA1 (CLV1) and CLAVATA2 (CLV2), and rice *floral organ number1* (*FON1*) and *Oryza sativa leucine-rich repeat receptor like protein kinase1* (*OsLRK1*) to CLV1. Rice *frizzy panicle* (FZP) and branched floretless1 (*BFL1*) are orthologues to the maize branched silkless1 (*BD1*) gene. The *ramosa 2* (*ra2*) expression pattern is conserved in rice, barley, sorghum and maize, suggesting that this gene is critical for shaping the initial steps of grass inflorescence architecture (Bortiri et al., 2006).

Inflorescence architecture is a key agronomic trait. For example, increasing the number of branches in the tassel increases pollen yield, which influences hybrid seed production and overall yield in maize (Cassani et al., 2006). From an applied breeding viewpoint, the inflorescence architecture genes that map in the vicinity of QTLs that harbour beneficial traits are of significant importance. For example, *td1*, *fea2* and *ra* map close to the QTLs reported for seed row number and spikelet density in maize (Taguchi-Shiobara et al., 2001; Bommert et al., 2005a; Upadayya et al., 2006a). Allelic variation in *ra* (*ra1*, *ra2* and *ra3*) forms part of a network of genes that control the production of tassel lateral branches in maize. Lateral branches in the panicle and spikelet density are also of importance in barley, rice and wheat. It is therefore possible that such genes could be manipulated to improve crop yields. Multiple QTLs control inflorescence branching in foxtail millet (Doust et al., 2005), and it would be interesting to determine whether any of the *ra* genes are candidates for these loci.

**Mapping and cloning QTLs associated with spike and grain development traits**

Understanding the pathways and gene(s) involved in seed development is a major step towards dissecting reproductive trait biology in cereals. Seed development in angiosperms begins with double fertilization, which results in a diploid embryo and a triploid endosperm. Further development is marked by the rapid growth of the endosperm and embryo until seed maturation, which is accompanied by desiccation. Simultaneously, the maternal ovule undergoes regulated growth to accommodate the growing embryo and endosperm. The endosperm in monocots constitutes the major contribution to the volume of the mature seed, in which the ovule integument forms the seed coat. Seed size is the result of three different growth programmes: those of the diploid embryo, the triploid endosperm and the diploid maternal ovule. All of these are genetically regulated, defining the development of the maternal integument, the embryo and the endosperm. For example, when the relative dosages of maternal and paternal genomes are perturbed, endosperm and seed size are affected. Endosperm development is also subject to the differential expression of many genes which are dependent on the parent of origin (see section on ‘Gene expression associated with panicle development, pollination, fertilization and developing seeds’ below).

Unlike the extensive use of molecular biology to study the genetics of agronomic traits, including resistance to biotic and abiotic stresses, there have been fewer studies dissecting the QTLs for panicle architecture and development in cereals (Table 4). Several QTLs from these studies were identified in regions known to harbour candidate genes, e.g. *fea2*, *td1* and *ramosa1* (*ra1*) in maize, *barren stalk1* (*ba1*), *ra1*, *ramosa2* (*ra2*), branched silkless1 (*bd1*), indeterminate spikelet1 (*ids1*), *leafy (lfy)*, terminal flower1 (*tf1*) and Dwarf3 (*Dw3*) in sorghum, and *Q* in wheat. The genes *Dw3* and *ra2* in sorghum were co-localized precisely with QTLs of large effect for rachis length, number of primary branches and primary branch length.

In rice, the QTLs for grain number (*Gn1a*), grain size (*GS2*), grain weight (*gw3.1* and *gw8.1*) and grains per panicle (*gpa7*) have been cloned, and the molecular function has been determined in some cases (Dwivedi et al., 2007, and references cited therein). More recently, Song et al. (2007) cloned a new QTL (*GW2*) for rice grain width and weight. It encodes a previously unknown RING-type protein with E3 ubiquitin ligase activity, which is known to function in the degradation of the ubiquitin–proteasome pathway. Loss of *GW2* function increased cell numbers, resulting in a larger (wider) spikelet hull, and accelerated the grain milk-filling rate, resulting in enhanced grain width, weight and yield. This provides an insight into the mechanism of seed development, a potential tool for improving grain yield in crops.

*AP2* is best known for its role in the regulation of flower meristem and flower organ identity and development in
Arabidopsis (Jofuku et al., 1994). A major QTL for seed size and seed weight was mapped on the Arabidopsis chromosome 4 region, which bears AP2 (Alonso-Blanco et al., 1999). More recently, Jofuku et al. (2005) demonstrated that AP2 gene sequences can be used to genetically engineer significant increases in Arabidopsis seed weight and seed yield. These could serve as useful markers to identify new yield-limiting genes in Arabidopsis, which, in turn, may be used in cereal crops.

Gene expression associated with panicle development, pollination, fertilization and developing seeds

Transcription factors are crucial in controlling development. Plants exhibit massive changes in gene expression, as measured by differences in transcriptional profiles, during morphophysiological and reproductive development, as well as on exposure to a range of biotic and abiotic stresses. This

| Table 3 Genes and their functions associated with inflorescence architecture in maize, rice, sorghum and wheat from 2000 to 2007 |
|---|---|---|---|
| Gene | Phenotype associated with the gene | Encoding for | Reference |
| Maize | | | |
| bi2 | Axillary meristem and lateral primordia in inflorescence | Co-orthologue of the PINOID serine/threonine kinase | McSteen et al. (2007) |
| ra1 | Inflorescence architecture | Zinc-finger transcription factor | Cassani et al. (2006) |
| ra2 | Increased branching with short branches replaced by long, indeterminate types | LATERAL ORGAN BOUNDARY domain protein | Bortiri et al. (2006) |
| ra3 | Inflorescence branching | Trehalose-6-phosphate phosphatase (TPP) enzyme | Sato et al. (2006) |
| dLFl1 | Delays flowering | A leucine zipper protein | Murakami et al. (2006) |
| td1 | Adversely affects male and female inflorescence development | A leucine-rich repeat receptor-like kinase protein, orthologue to Arabidopsis CLAVATA1 | Bommert et al. (2005a) |
| id | Blocks transition from vegetative to reproductive phase | Zinc-finger transcription factor | Kozaki et al. (2004) |
| ba1 | Plants unable to produce tillers, female inflorescence and a tassel | A non-canonical basic helix–loop–helix protein | Gallavotti et al. (2004) |
| Bd1 | Forms indeterminate branches instead of spikes | ERF transcription factor | Chuck et al. (2002) |
| fea2 | Overproliferation of ear inflorescence meristem, with modest effect on floral meristem and organ number | A leucine-rich repeat receptor-like kinase protein, related to CLAVATA2 from Arabidopsis | Taguchi-Shiobara et al. (2001) |
| Rice | | | |
| fzp-9(1) | Transition from the spikelet meristem to floral meristem | Putative protein of 318 amino acids, homologue to maize Bd1 | Yi et al. (2005) |
| FON1 | Regulates floral meristem size | A leucine-rich repeat receptor-like kinase, orthologous to CLV1 of Arabidopsis | Suzaki et al. (2004) |
| FZP and LAX | Prevents formation of axillary meristem | ERF transcription factor, orthologous to maize BD1 | Komatsu et al. (2003a) |
| OsFOR1 | Auxiliary meristem formation | Basic helix-loop–helix transcription factor | Komatsu et al. (2003b) |
| BFL1 | Formation and/or maintenance of floral organ primordia | A protein that contains a leucine-rich repeat domain | Jang et al. (2003) |
| OsdRK1 | Transition from spikelet meristem to floret meristem | Transcription factor protein containing EREBP/AP2 domain, maize orthologue BD1 | Zhu et al. (2003) |
| Sorghum | | | |
| sb-ba1 | Primary branch initiation | Basic helix-loop–helix transcription factor | Brown et al. (2006) |
| sb-ra1 and sb-ra2 | Branch determinacy | EPF-class C2H2 zinc finger, or LOB domain | Kim et al. (2000) |
| Sb-bd1 and Sb-db1 | Spikelet determinacy | 1 AP2-domain ERF, or 2 AP2-domain ERF | Taguchi-Shiobara et al. (2003) |
| Sbrf1 and Sb-tfl1 | Reproductive transition/inflorescence branching | FLORI-CAULIFLOWER, or TFL1/CEN (PEBP) | Yi et al. (2003) |
| Dv3 | Auxin transport | P-glycoprotein | Yi et al. (2003) |
| Wheat | | | |
| TaVFT-1 | Transition from vegetative to reproductive phase, responsive to vernalization | A protein homologous to the MADS-box family of transcription factors | Danyuk et al. (2003) |
| WAP1 | Activate autonomous phase transition from vegetative to reproductive growth | APETALA1 (AP1)-like MADS-box gene, with large family of transcription factors | Murai et al. (2003) |
variation in transcript abundance can be associated with gene expression using expression quantitative trait loci (eQTLs) analysis (Dwivedi et al., 2007, and references cited therein).

Elucidating the regulatory mechanisms of plant organ development is an important component of plant developmental biology, and will be useful for crop improvement. Plant organ formation, or organogenesis, is a well-orchestrated series of events by which a group of primordial cells differentiates into an organ. The male reproductive organ, consisting of an anther, in which the male gametophyte develops, and a filament, which provides water and nutrients to the anther. Anther development occurs in two phases. In phase I, the anther structure is established, including the differentiation of cell types and meiosis of pollen mother cells. In phase II, the microspores develop into pollen grains and the anthers dehisce to release pollen grains (Zhao et al., 2002). Using organ-specific gene expression profiling, Lu et al. (2006) detected 26 genes that were preferentially up-regulated during early stamen development, and clustered into two distinct clades, suggesting that early stamen development involves two distinct phases of pattern formation and cellular differentiation. Both gibberellic acid (GA₃) and jasmonic acid (JA) play important roles in anther development and pollen fertility. Using a 10K cDNA microarray with probes derived from seedlings, meiotic anthers, mature anthers and GA₃- or JA-treated suspension cells of rice, Wang et al. (2005b) detected expression level changes in 2155 genes in anthers, compared with those in seedlings. Forty-seven genes with potential functions in cell cycle and cell structure regulation, hormone response, photosynthesis, stress resistance and metabolism were differentially expressed in meiotic and mature anthers. Of the 314 genes that responded to either GA₃ or JA applications, 24 GA₃- and 82 JA-responsive genes showed significant changes in expression between meiosis and the mature anther stage, with the gene y656d05 not only highly expressed in meiotic anthers but also induced by GA₃. These reports identified a number of candidate genes likely to be involved in both pollination and fertilization. Their detailed characterization is expected to provide a better understanding of the genetic programmes controlling pollination and fertilization in cereal crops.

The basic structure of the rice inflorescence (the panicle) is determined by the pattern of branch formation, which is established at the early stage of panicle development. Young panicle organs (YPOs) in cereals correspond to the onset of

### Table 4 Quantitative trait loci (QTLs) associated with inflorescence and grain development in maize, foxtail millet, rice and wheat from 2000 to 2007

<table>
<thead>
<tr>
<th>Trait</th>
<th>Maize</th>
<th>Foxtail millet</th>
<th>Rice</th>
<th>Wheat</th>
</tr>
</thead>
<tbody>
<tr>
<td>QTL associated with inflorescence structure and grain development</td>
<td>Five QTLs for tassel and nine QTLs for ear traits; a QTL on chromosome 7 for tassel branches near ra₁, a candidate gene for tassel branches</td>
<td>Three QTLs each for primary branches and primary branch density, 6 for spikelets, and 2 for bristle number</td>
<td>Two QTLs associated with increased spikelets per panicle; QTL on chromosome 12 accelerated grain filling during early filling stage, and QTL on chromosome 8 increased grain filling by translocating non-structural carbohydrate (NSC) from the culm and leaf sheaths to the panicle</td>
<td>Two to six QTLs for five spike-related traits; a major QTL for spikelets Yagi</td>
</tr>
<tr>
<td>QTL associated with increased spikelets per panicle</td>
<td>Forty-five QTLs for tassel inflorescence; several in regions with candidate genes fea₂, td₁ and ra₁</td>
<td>Eight QTLs associated with increased spikelets per panicle</td>
<td>QTL on chromosome 10 linked to QTL for FLW and GWE</td>
<td>Eight QTLs associated with increased spikelets per panicle</td>
</tr>
<tr>
<td>A major QTL for spikelets</td>
<td>Three QTLs for tassel branch angle (TBA) and six for tassel branch number (TBN), a QTL on chromosome 5 for TBA in the same region as a QTL for TBN</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Major QTL for spikelet length and spikelet number assigned to A- and B-genome chromosomes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Four to six QTLs for spike length, spikelet number and compactness; several affecting more than one trait; a QTL co-segregated with Q involved in ear morphology</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

© 2008 The Authors
Journal compilation © 2008 Blackwell Publishing Ltd, Plant Biotechnology Journal, 6, 529–559
meiosis and mark the transition from the vegetative to the reproductive phase, and therefore make a great contribution to grain yield. By comparing gene expression profiling of YPOS with that of rice aerial vegetative organs (AVOs), it is possible to gain further molecular insight into this developmentally and functionally important period. Comparative analysis of rice ESTs between YPOS and AVOs revealed nearly 20,000 unigenes differentially expressed in YPOS and AVOs, and about 10,000 mRNAs specifically expressed in YPOS (Tang et al., 2005). Using a cDNA microarray containing 10,000 unique rice genes, Lan et al. (2004) analysed gene expression in the pistil, which revealed that the anther has a gene expression profile more similar to that of the roots than to that of the pistil, and most pistil preferentially expressed genes respond to pollination or fertilization. A total of 253 ESTs exhibited differential expression during pollination and fertilization, and about 70% of these were assigned to a putative function. Furthermore, 136 pollination-related genes and 57 fertilization-related genes of the 253 genes of these types were also regulated by dehydration and wounding, respectively. Nearly half of the genes expressed preferentially in non-pollinated pistils were responsive to dehydration, which suggests that the genetic programmes regulating them are probably related. In addition, 100 of the 253 genes do not appear to be regulated by stress treatments, which indicates that additional genetic networks are involved in pollination and fertilization. Taken together, these observations reveal that the genetic networks regulating photosynthesis, starch metabolism, gibberellin and defence responses are involved in pollination and fertilization (Lan et al., 2005).

Furutani et al. (2005) conducted global transcriptome profiling to obtain a comprehensive view of gene expression in the early stages of rice panicle development from phase transition to floral organ differentiation. They detected 357 genes expressed differentially in the early stages of panicle development, clustered into seven groups on the basis of the temporal expression patterns. Interestingly, a small number of these genes extensively rich in transcription factors were up-regulated in the shoot apical meristems immediately after phase transition, when each gene exhibits a unique and interesting localization of mRNA. Duan et al. (2005) used a cDNA chip containing 325 rice cDNA clones, encoding known or putative transcription factors belonging to 12 different families, to study gene expression at eight continuous seed developmental stages in rice. They detected 135 transcription factor genes preferentially transcribed in seeds, grouped into 12 clusters. Each cluster contained transcription factor genes that peaked and waned at different stages, and 49

seed preferential genes playing major roles in distinct stages of seed development were detected. Many of these seed preferential transcription factors are also involved in hormone and abiotic stress effects, suggesting the existence of uncharacterized transcriptional networks, or crosstalk, between hormone and abiotic stress signalling and seed development. Dof proteins play essential roles in the hierarchical regulation of gene expression during rice seed development.

The caryopsis is the fruit of grasses in which the pericarp is fused to the seed coat at maturity, referred to as the grain in cereals. The pericarp consists of three major structures, the embryo, endosperm and seed coat, with the endosperm being the tissue of economic value. Understanding the development of the cereal caryopsis holds the future of metabolic engineering to enhance global food production. McIntosh et al. (2007) developed LongSAGE (Serial Analysis of Gene Expression) libraries at five time points post-anthesis that coincide with key processes in caryopsis development in wheat, and identified 29,261 unique tag sequences across all five libraries. The 500 most abundant tags spanned development, which highlights the array of functional groups being expressed during grain development. Furthermore, differential expression profiles of abundant tags from each library revealed the co-ordinated expression of genes responsible for the cellular events constituting caryopsis development. These provided novel sequence and expression information, including the identification of potentially useful promoter activities.

Expression analysis using the Affymetrix GeneChip® Wheat Genome Array on mRNA from developing seeds of double haploid lines, grown at two locations under field conditions, identified 1455 common probe sets differing in intensity between lines in both locations; 542 eQTLs were detected that each mapped to a single chromosome interval (with a few exceptions in which eQTLs clustered), suggesting regulatory control for many genes in small chromosome intervals. This may constitute a major gene eQTL for developing seed in wheat. Comparison of expression mapping data with physical mapping of wheat ESTs confirmed the presence of both cis- and trans-acting eQTLs. Many of the eQTL clusters were coincident with QTLs controlling dough, bread-making and seed weight traits, which are based entirely on mature seed composition (Jordan et al., 2007). Thus, by mapping patterns of gene expression as eQTLs controlling complex traits, key regulatory regions may be identified and used in molecular plant breeding. A knowledge of the genes regulated by these eQTLs could provide insight into the biochemical nature of the traits, assist in map-based cloning of the regulatory elements and help associate genotype with phenotype.
Table 5  Nuclear fertility restorer (Rf) genes, cytoplasmic male sterility (CMS) sources and DNA markers associated with Rf in barley, maize, rice, sorghum and wheat from 1994 to 2006

<table>
<thead>
<tr>
<th>Rf</th>
<th>CMS source</th>
<th>DNA markers associated with Rf</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barley</td>
<td>Rfm1</td>
<td>msm1 and msm2</td>
<td>e34m2, e46m19 and e46m17 on chromosome 6; the closest e34m2 and e46m19</td>
</tr>
<tr>
<td>Maize</td>
<td>Rf1</td>
<td>S</td>
<td>RF1 on chromosome 2; closest marker E7P6, E12M7 and E3P1 from Rf3</td>
</tr>
<tr>
<td>Maize</td>
<td>Rf3</td>
<td>S</td>
<td>RF3 on chromosome 2; closest marker E7P6, E12M7 and E3P1 from Rf3</td>
</tr>
<tr>
<td>Maize</td>
<td>RF4 and RF5</td>
<td>C</td>
<td>RF4 restores fertility in all C lines, and RF5 in C lines lacking RF-1, mapped on chromosome 1</td>
</tr>
<tr>
<td>Maize</td>
<td>RF6(t)</td>
<td>t</td>
<td>RF6(t) co-segregates with RM3150 and flanked by RM118 and RM5373, whereas RF6(t) co-segregates with RM5373 and flanked by RM6737 and SBD07</td>
</tr>
<tr>
<td>Maize</td>
<td>RF1 (D1)</td>
<td>D1</td>
<td>RF1 on chromosome 10 between RM171 and RM6100</td>
</tr>
<tr>
<td>Maize</td>
<td>RF-D1(t)</td>
<td>D1</td>
<td>RF-D1(t) on chromosome 10 between OSR33 and RM228</td>
</tr>
<tr>
<td>Maize</td>
<td>RF5</td>
<td>WA</td>
<td>OPKOS-800, OPU10-1100, OPW1-350, RGS32, R140 and RG458 on chromosome 1</td>
</tr>
<tr>
<td>Sorghum</td>
<td>rf4</td>
<td>A3</td>
<td>LW7 and LW8 mapped to rf4, whereas LW9 on the flanking side of rf4</td>
</tr>
<tr>
<td>Sorghum</td>
<td>rfl</td>
<td>A1</td>
<td>rfl on LG H close to Xtxa2582, whereas Xtxp18 and Xtxp250 flank the locus</td>
</tr>
<tr>
<td>Wheat</td>
<td>D2RF1</td>
<td>Y1 4060</td>
<td>E09-SCARfl mapped to the D2RF1 locus, whereas Xgwm11 and Xgwm18 co-segregate with E09-SCARfl</td>
</tr>
<tr>
<td>Wheat</td>
<td>RF3</td>
<td>Timopheevi-based CMS</td>
<td>Xbarc207, Xgwm131 and Xbarc61 close to RF3 on 1B and two minor QTLs on 5A and 7D</td>
</tr>
</tbody>
</table>

Male sterility and fertility restoration in hybrid breeding

Mapping and cloning nuclear Rf genes

CMS is a maternally inherited trait characterized by the inability to produce functional pollen, often associated with unusual open reading frames (ORFs) in the mitochondrial genome. Male fertility can be restored by Rf genes (Schnable and Wise, 1998). Thus, CMS/Rf systems are ideal models for studying the genetic interaction and function of mitochondrial and nuclear genomes in plants. Several CMS systems, in addition to normal male-fertile cytoplasm, which yields fertile plants in all known nuclear backgrounds, have been reported in maize (S, C and T) (Backett, 1971), pearl millet (A1, A2, A3, A4, Av and aero) (Chandra-Shekara et al., 2007), rice (WA, DA, GAM, HL and DIS in indica, and D and BO in japonica backgrounds) (Tao et al., 2004, and references cited therein) and sorghum (A1, A2, A3, A4, Av and 9E) (Scherzt, 1994). They can be differentiated by Rf gene(s) that suppress their associated male-sterile phenotype, allowing normal pollen development. To date, few CMS and Rf systems have been extensively exploited in hybrid breeding: for example, A1 (Tift23A) in pearl millet (Rai et al., 2001), A1 (milo) in sorghum (Moran and Rooney, 2003), T in maize (Ullstrup, 1972), and D and BO in japonica rice and WA, DA, GAM, DIS and HL in indica rice (Tao et al., 2004).

DNA markers associated with Rf gene(s) in barley (Rfm1), maize (rf1, rfl, RF3, RF4 and RF5), rice (Rf1, RF3, RF4, RF5, RF6(t) and RF-D1(t)), sorghum (rf1 and rf4) and wheat (D2RF1 and RF3) (Table 5), and those associated with thermosensitive male sterility (TGMS) gene(s) in maize (tms3), rice [tms1, tms4(t), tms5 and tms6] and wheat (wtms1), photoperiod-sensitive genetic male sterility (PGMS) gene in rice (pms3) and thermo-photoperiod-sensitive genetic male sterility (TPGMS) genes in wheat (wtms1 and wtms2) (Table 6) have been mapped to respective chromosome regions of each species. These DNA markers can be used to transfer Rf or tms alleles to new genotypes.

To date, only r2 in maize (Cui et al., 1996), RF1 in rice (Akagi et al., 2004, Komori et al., 2004) and RF1 in sorghum (Klein et al., 2005) have been cloned. Maize r2 encodes an...
aldehyde dehydrogenase that restores male fertility in T cytoplasm (Liu et al., 2001). Rice Rf1, delimited to a 22.4-kb region, encodes a mitochondrially targeted protein containing 16 repeats of the 35-amino-acid pentatricopeptide repeat (PPR) motif. Klein et al. (2005) cloned Rf1 of sorghum, which they resolved to a 32-kb region spanning four ORFs: a plasma membrane Ca\(^{2+}\)-ATPase, a cyclin D-1, an unknown protein and a PPR13. The first three were completely conserved between fertile and sterile plants. In the approximately 7-kb region spanning PPR13, they identified 19 sequence polymorphisms that co-segregated with the fertility restoration phenotype. PPR13 encodes a mitochondrially-targeted protein containing a single exon with 14 PPR repeats, not present in rice, and a candidate gene for Rf1 in sorghum. More recently, Wang et al. (2006) demonstrated in rice with BO cytoplasm that an abnormal mitochondrial ORF, Orf79, is co-transcribed with the duplicated atrp6 (8-atrp6) gene and encodes a cytotoxic peptide. Two Rf genes at the Rf locus, Rf1a and Rf1b, within an approximately 105-kb region, are members of a multigene cluster encoding PPR proteins. Both target mitochondria and restore male fertility by blocking Orf79 production via endonucleolytic cleavage or degradation of dicistronic B-atp6/orf79 mRNA.

Of the several TGMS genes reported in rice, Zhou et al. (2006) cloned OsAPT2, located on chromosome 4, which encodes a putative adenine phosphoribosyl transferase. This is associated with tms5 in rice, and the OsAPT2 transcript in the young panicle is down-regulated at 29°C, the critical temperature for induction of fertility conversion in the TGMS mutant ‘Annong S-1’.

Table 6 DNA markers associated with thermo-, photoperiod- and thermo-photoperiod-sensitive genetic male sterility in maize, rice and wheat from 1997 to 2006

<table>
<thead>
<tr>
<th>Male sterility gene</th>
<th>Source material</th>
<th>DNA markers associated with male sterility</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Thermosensitive genetic male sterility (TGMS)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Maize</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tms3</td>
<td>Qiong68ms</td>
<td>tms3 on chromosome 2 between umc2129 and umc1041</td>
<td>Tang et al. (2006)</td>
</tr>
<tr>
<td><strong>Rice</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tms6</td>
<td>Sokcho-MS</td>
<td>tms6 on chromosome 5 between RM3351 and E60663</td>
<td>Lee et al. (2005b)</td>
</tr>
<tr>
<td>tms6(i)</td>
<td>0A15-1</td>
<td>tms6(i) on chromosome 3, close to centromere, linked to S187770</td>
<td>Wang et al. (2004)</td>
</tr>
<tr>
<td>tms5</td>
<td>AnnongS-1</td>
<td>tms5 on chromosome 2 between C365-1 and G227-1</td>
<td>Wang et al. (2003)</td>
</tr>
<tr>
<td>rtm(s)</td>
<td>J2075</td>
<td>rtm(s) on chromosome 10 between RM222 and RG257</td>
<td>Jia et al. (2001)</td>
</tr>
<tr>
<td>tms4(t)</td>
<td>TGMS-VN1</td>
<td>tms4(t) on chromosome 2, E5/M12-600 the closest</td>
<td>Dong et al. (2000)</td>
</tr>
<tr>
<td>tms2</td>
<td>Norin PL12</td>
<td>tms2 on chromosome 2 between R643 and R1440</td>
<td>Yamaguchi et al. (1997)</td>
</tr>
<tr>
<td>tms3(i)</td>
<td>IR32364TGMS</td>
<td>OP185005, OP191990, OPAA7, OPAC3, linked to tms3(i)</td>
<td>Subudhi et al. (1997)</td>
</tr>
<tr>
<td><strong>Wheat</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>wtm6</td>
<td>BNY-S</td>
<td>wtm6 on chromosome 2 between Xgwm374 and E:AAG/M:CTA</td>
<td>Xing et al. (2003)</td>
</tr>
<tr>
<td><strong>Photoperiod-sensitive genetic male sterility (PGMS)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Rice</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pms3</td>
<td>Nongken 58</td>
<td>pms3 on chromosome 12, localized to 28.4-kb DNA fragment surrounded by 15 RFLP markers</td>
<td>Lu et al. (2005)</td>
</tr>
<tr>
<td><strong>Thermo-photoperiod-sensitive genetic male sterility (TPGMS)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Wheat</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>wptms1 and wptms2</td>
<td>337S</td>
<td>wptms1 on chromosome 5B and wptms2 on 2B, wptms1 between Xgwm374 and Xgwm371, and wptms2 between Xgwm374 and Xgwm120</td>
<td>Guo et al. (2006)</td>
</tr>
</tbody>
</table>
sterility (Dewey et al., 1987). CMS S-type cytoplasm of maize is associated with the expression of orf355-orf77 in the R region of mitochondrial DNA, such that Rf3 can regulate nuclear and mitochondrial gene expression and shows pleiotropic effects on the transcriptional level. Using a cDNA microarray and suppression subtractive hybridization, Zhang et al. (2005c) studied the global expression profile caused by Rf3 substitution of rF3 in S cytoplasm during pollen development in a set of Rf3/rF3 near-isogenic line (NIL). They identified 137 tentative unique genes and differential expression amongst S-(rF3) and S-(rF3). For example, in S-(rF3), the expression patterns of genes associated with electron or H+ conduction and of anti-apoptosis genes were distinctly different from those of S-(rF3). Rf3 may therefore regulate the accumulation of nuclear and mitochondrial gene transcripts directly or indirectly to inhibit multiple programmed cell death pathways in S-type cytoplasm, allowing the normal development pathways to unfold. Allen et al. (2007) sequenced five distinct mitochondrial genomes in maize — two fertile (NA and NB) and three cytoplasmic male-sterile (CMS-C, CMS-S, CMS-T) cytotypes that ranged from 535 825 bp in CMS-T to 739 719 bp in CMS-C, with large duplications accounting for most size differences. Plastid DNA accounts for 2.3%—4.6% of each mitochondrial genome. The genomes share a minimum of 51 genes for 33 conserved proteins, three rRNAs and 15 tRNAs. Duplicate genes and plastid-derived tRNAs vary amongst cytotypes; however, known CMS-associated ORFs were detected in CMS-S and CMS-T, but not in CMS-C.

Male-sterile sorghum carrying A3 cytoplasm (IS1112C) represents an unusual example of aberrant microgametogenesis, wherein microsprores develop into non-viable pollen that remains physically intact until anther exsertion. Using cDNA amplified fragment length polymorphism (AFLP) transcript profiling and male-fertile and male-sterile NILs at three different stages of pollen development, Pring and Tang (2004) discovered that transition from early to late stages is characterized by changes in abundance of nearly 33% of shared transcripts, and early- or late-specific expression of about 10% of transcripts. Male-sterile plants exhibit extensive changes in regulatory patterns from those characteristic of fertile plants, including premature expression of late-specific and prolonged expression of early-specific transcripts. Furthermore, genome-wide transcriptome patterns indicate the expression of an estimated 12 000 genes in early-mid-microsprores, and the abundance of at least 15% of these transcripts is altered in male-sterile plants. NIL restored to male fertility is characterized by the apparent normalized expression of most of these transcripts. The development of the microgametophyte is thus characterized by dynamic programmed changes in gene expression, and the expression of male sterility compounds these changes in a complex manner. Pring et al. (2006) examined gene expression and the abundance of protein products from the selected mitochondrial genes in near-isogenic male-sterile, fertility-restored and normal cytoplasm lines in two genetic backgrounds, in addition to recombinant inbred lines differing at two fertility loci. They characterized seven transcript-derived fragments (TDFs) unique to a BC1F3 fertility-restored line, four of which were eliminated in a marker-selected BC1F3 line, whereas the remaining three TDFs were detected in all lines examined, indicating that these unique fragments represent alleles derived from the non-recurrent male parent, in which some of the TDFs may be linked to Rf alleles. The male-sterile microspore mitochondria contained elevated levels of the nuclear-encoded alternative oxidase protein, potentially associating mitochondrial dysfunction with failed pollen development. The relative abundance of mitochondrial-encoded ATPase, ATP6 and cyclo-oxygenase II proteins and nuclear-encoded manganese superoxide dismutase, cytochrome c and aldehyde dehydrogenase proteins, in contrast, did not vary with cytoplasm or fertility status at the microspore stage.

Assessing the genetic purity of hybrids, CMS and maintainer lines

Ensuring the genetic purity of parental lines and hybrids is a prerequisite to realizing the full potential of the hybrids. It is estimated that every 1% impurity of female line seed in rice hybrid seed reduces the yield by 100 kg/ha (Mao et al., 1996). One of the common admixtures in hybrid seed production is that of maintainer lines with CMS lines. Given that both are isonuclear lines, it is not possible to distinguish between them until they flower. The purity of hybrid seed lots is normally assayed by a grow-out test on a representative sample of the seed to be marketed (Verma, 1996). The grow-out test involves growing plants to maturity, and assessing the morphological and floral characteristics that distinguish the hybrids, thus excluding the immediate cultivation of the hybrid seed produced. The grow-out test is costly and subject to high genotype–environment interaction. DNA-based markers can be used to assess more quickly and precisely the genetic purity of hybrids and their parental lines. For example, several PCR-based markers have been reported in rice that not only discriminate between CMS lines and their cognate maintainer lines, but also serve to assess the genetic purity of the hybrids (Table 7). More research is needed to develop such marker systems in other crops in which three-line hybrids are used.
Mitochondrial DNA sequence variation is also useful for fingerprinting male sterility-inducing cytoplasm, determining cytoplasmic diversity amongst germplasm accessions and identifying new sources of cytoplasm that induces male sterility (Xu et al., 1995).

Engineering CMS for hybrid breeding

Several mechanisms have been reported to create a de novo CMS trait in a species lacking CMS, or to eliminate the flaws in an existing CMS system in plants, mostly affecting tapetum and pollen development (Yui et al., 2001; Zheng et al., 2003). Severe phenotypic alterations as a result of interference with plant metabolism and development have precluded their use in agriculture (Goetz et al., 2001). Mitochondria and plastids are maternally inherited in most plant species. Transformation of chloroplasts with genes interfering in the plastid metabolic pathway essential for pollen development is the most exciting prospect for the development of new forms of CMS. Ruiz and Daniell (2005) were the first to report CMS in Nicotiana tabacum through the transformation of the chloroplast genome with a single gene, β-ketothiolase (phaA). The transgenic lines were normal except for the male-sterile phenotype, which lacked pollen. Abnormal thickening of the outer wall, enlarged endothecium and vacuolation affected pollen grains, and resulted in an irregular shape or collapsed phenotype. More importantly, reversibility of the male-sterile phenotype under continuous illumination resulted in viable pollen and a copious amount of seeds. This approach opens up a totally new route for the engineering and testing of new CMS systems in crops. Technical issues related to the use of this approach, including problems associated with plastid transformation, the design of restorer genes, stability of the trait, possible negative effects of the introduced gene (causing male sterility) on agronomic traits, and biosafety and social acceptance concerns, must be addressed before this technique may be widely used (Pelletier and Budar, 2007).

The stability of the mitochondrial genome is controlled by nuclear loci, and nuclear genes suppress mitochondrial DNA rearrangements during plant development. One such nuclear gene is Msh1. To test that Msh1 disruption leads to the type of mitochondrial DNA rearrangements associated with naturally occurring CMS in plants, Sandhu et al. (2007) used a transgenic approach for RNAi to modulate the expression of Msh1 in tobacco and tomato. This resulted in reproducible mitochondrial DNA rearrangements and a condition of male sterility (non-reversible) that was heritable, associated with normal female fertility, and maternal in its inheritance, providing a means to develop novel non-genetically modified organisms or transgenic CMS lines.

Male fertility in flowering plants is dependent on the production of viable pollen grains within the anther. Genes expressed exclusively in the anther are likely to include those that control male fertility. Bcp1, an anther-specific gene from Brassica campestris, is essential for pollen development. Inhibition of its expression in either tapetum or microspores prevents the production of fertile pollen in transgenic Arabidopsis plants (Xu et al., 1995). Using differential screening of a cDNA library from rice panicles, Luo et al. (2006) isolated a tapetum-specific gene, RTS, which encodes a putative polypeptide of 91 amino acids with a hydrophobic N-terminal region. RTS is predominantly expressed in the anther’s tapetum during meiosis. Down-regulation of its expression leads to pollen abortion, and therefore male sterility, in transgenic rice plants. The promoter region of RTS, when fused to the Bacillus amyloliquefaciens ribonuclease gene, barnase, or the antisense of the RTS gene, is able to drive tissue-specific expression of both genes in rice, creeping bentgrass (Agrostis stolonifera L.) and Arabidopsis, conferring male sterility to the transgenic plants. When fertilized with

<table>
<thead>
<tr>
<th>Diagnostic marker</th>
<th>Marker differentiating hybrid and its parental lines</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drccms</td>
<td>Distinguishes CMS lines from their cognate maintainers</td>
<td>Rajendrakumar et al. (2007a)</td>
</tr>
<tr>
<td>BF-ST5-401 and BF-ST5-420</td>
<td>Differentiates CMS lines from their corresponding maintainer lines</td>
<td>Rajendran et al. (2007b)</td>
</tr>
<tr>
<td>RM258</td>
<td>Differentiates CMS and restorer lines from hybrid</td>
<td>Garg et al. (2006b)</td>
</tr>
<tr>
<td>RM9</td>
<td>RM9386 differentiates CMS lines and their hybrids but not their cognate maintainer lines</td>
<td>Yashitola et al. (2004)</td>
</tr>
<tr>
<td>M1, M2 and M3</td>
<td>Differentiates from each other japonica hybrids and their parental lines</td>
<td>Komori and Nitta (2004)</td>
</tr>
<tr>
<td>RM206, 216, 258 and 263</td>
<td>Differentiates 11 rice hybrids grouped into three clusters, hybrids within cluster share a common CMS source</td>
<td>Nandakumar et al. (2004)</td>
</tr>
<tr>
<td>E-AAC/M-CTC</td>
<td>A fragment with six base pair (Leu-tRNA gene) difference in cpDNA differentiates five CMS lines from their corresponding maintainer lines</td>
<td>He et al. (2003)</td>
</tr>
<tr>
<td>RM 164 and pTA248</td>
<td>Detects polymorphism between CMS, hybrid and restorer lines</td>
<td>Yashitola et al. (2002)</td>
</tr>
</tbody>
</table>
wild-type pollen, these transgenic plants produce normal seed sets, confirming no adverse effect on female fertility. Thus, Bcp1 or RTS and its promoter have great potential for engineering male sterility in other crop plants.

Pyramiding TGMS and Rf genes

Using simple sequence repeat (SSR)-linked markers and TGMS donors, each possessing different genes, Nas et al. (2005) developed two-gene and three-gene pyramids (IR80775-46 with tms2 and tms5, and IR80775-21 with tms2, tgm5 and tms5) possessing the RM11 allele of Norin PL 12 (tms2), the RM257 allele of SA2 (tgm5) and the RM174 allele of DQ200047-21 (tms5), which expressed as male sterile under sterile-inducing conditions. In addition, rice SF21 is a putative pollen-specific protein (IRGSP) because of its high degree of amino acid sequence alignment to known pollen proteins in Arabidopsis and sunflower.

Sattari et al. (2007) used two sequence-tagged site (STS) markers (RG140/Pvull and S10019/BlstU) to select for two major Rf genes (Rf3 and Rf4) governing fertility restoration of CMS in rice. The combined use of markers associated with these two loci improved the efficiency of screening for putative restorer lines from a set of elite lines. Breeders, in general, identify restorers by test crossing prospective lines with available CMS lines and evaluating F1 progenies for general, identify restorers by test crossing prospective lines from a set of elite lines. Breeders, in

Acknowledgements

Sangam Dwivedi is grateful to the library staff of the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) (India) for assistance in literature searches and sourcing reprint, to Jonathan Crouch [International Maize and Wheat Improvement Center (CIMMYT), Mexico] and H. D. Upadhyaya (ICRISAT, India) for support during the development of the manuscript, and to K. J. Edwards (editor) and an anonymous reviewer of Plant Biotechnology Journal for making useful suggestions on improving the manuscript. He also gratefully acknowledges the editorial input of Mike Listman and Allison Gilles (CIMMYT, Mexico) in a previous version of the manuscript.

References


© 2008 The Authors


Upadhyayula, N., de Silva, H.S., Bohn, M.O. and Rocheford, T.R.


