



Trends in comparative genetics and their potential impacts on wheat and barley research

David A. Laurie and Katrien M. Devos

John Innes Centre, Norwich Research Park, Colney, Norwich NR4 7UH, UK (e-mail david.laurie@bbsrc.ac.uk)

Received 9 February 2001; accepted in revised form 16 August 2001

Key words: barley (*Hordeum vulgare*), comparative genetics, comparative mapping, wheat (*Triticum aestivum*)

Abstract

We review some general points about comparative mapping, the evolution of gene families and recent advances in the understanding of angiosperm phylogeny. These are considered in relation to studies of large-genome cereals, particularly barley (*Hordeum vulgare*) and wheat (*Triticum aestivum*), with reference to methods of gene isolation. The relative merits of direct map-based cloning in barley and wheat, utilization of the smaller genome of rice (*Oryza sativa*) and gene homology methods that utilize information from model species such as *Arabidopsis thaliana* are briefly discussed.

Introduction

‘The day does not seem far off when nearly all experimental biology will be driven by some kind of comparative genome suggested hypothesis.’ This quote, from Smith (1998), illustrates two general trends in genetics. The first is an increasing reliance on genomics as the starting point for investigation and the second is the integration of knowledge from different species to understand how the genetic control of traits evolves and diversifies. The latter (comparative genetics) seeks to understand characters in an evolutionary context and the extent to which similar phenotypes are the result of variation in orthologous genes. To quote from Smith (1998) again, ‘... as we begin to try to understand the regulatory circuitry of the cell, one assumes that by comparison of closely related organisms we will learn much about how these circuits have been modified and augmented to carry out distinct developmental and other programs’.

Comparative genetics focuses on three aspects of biology. The first is the conservation of gene sequence in different organisms, and investigates the extent to which conservation of sequence is paralleled by conservation of function. The second aspect utilizes the existence of conserved sequences, typically cDNA clones, for the development of genetic maps in differ-

ent species that contain shared ‘anchor’ markers (Van Deynze *et al.*, 1998). This enables the linkage groups of different species to be aligned and the number of rearrangements between them to be defined. The resulting comparative genetic maps provide information on chromosome evolution and identify cases where equivalent phenotypes in different species are likely to be controlled by orthologous genes. The third aspect, which we will not consider here, studies the overall physical organization of chromosomes and encompasses the role of repeated sequences, including transposable elements, in genome evolution.

How can comparative genetics help our understanding of barley and wheat?

Cereals are a staple food for a large proportion of the world’s population and it is therefore important to understand the genetic basis of yield, yield stability and grain quality. But, because of the cost and effort of analysis, a major issue for cereals, and other crops, is the extent to which information can be integrated between species. A common genetic basis for specific traits is a logical consequence of evolutionary divergence from common ancestors, but commonality will obviously decrease as more di-

verged species are considered. Comparative genetics seeks to utilize evolutionary conservation, avoiding unnecessary duplication of effort and exploiting the technical advantages of individual species including model plants such as *Arabidopsis thaliana*. Identification of conserved mechanisms of growth and development between species potentially offers rapid and cost-effective ways of understanding the genetic basis of performance in crop species. What is unclear at present is the extent of conservation of gene sequence and function within the grasses and within plants as a whole.

Comparative approaches are of particular interest in wheat (*Triticum aestivum*) and barley (*Hordeum vulgare*) because they have common features that hamper gene isolation. The first is genome size. Barley, although a diploid species ($2n = 2x = 14$), has a genome of about 5400 Mb compared to 130 Mb in *Arabidopsis* and 450 Mb in rice (*Oryza sativa*). Wheat is a hexaploid ($2n = 6x = 42$) with each constituent genome about the size of barley, giving a total of ca. 17000 Mb. This has discouraged map-based cloning. However, high-quality large-insert libraries now exist for barley (Yu *et al.*, 2000) and the A and D genome diploid relatives of wheat (Lijavetzky *et al.*, 1999 and Moullet *et al.*, 1999, respectively). Map-based gene cloning in barley has been proved feasible by the isolation of the *mlo* and *rar1* disease resistance genes (Büschges *et al.*, 1997 and Shirasu *et al.*, 1999, respectively).

The second major problem is that although a large proportion of the wheat and barley genomes contain retrotransposons of various kinds, they differ from maize in lacking transposable elements suitable for gene tagging. Retroelements are being used in rice (Hirochika, 1997; Sato *et al.*, 1999) but in that case the starting copy number is much lower, enabling transposition to be monitored efficiently. These difficulties, although not insurmountable, have encouraged comparative approaches that utilize the experimental advantages of other species. But to make best use of comparative methods it is advantageous to have an accurate phylogeny of plants.

Recent advances in angiosperm phylogeny and their implications for comparative analyses

Recently, extensive comparisons of nuclear and organelle DNA sequences in plants have been made with the aim of developing a robust angiosperm phylogeny

(Soltis *et al.*, 1999, 2000; Qiu *et al.*, 1999; Kuzoff and Gasser, 2000; Mathews and Donoghue, 2000; Soltis and Soltis, 2000). In some cases this is with the express aim of developing a tool for comparative biology (Soltis *et al.*, 1999).

Figure 1 illustrates current views on the phylogenetic relationships of seed plants and the positions of some important model and crop species. Recent studies show that gymnosperms are a monophyletic group that is sister to all angiosperms. Within the angiosperms the Amborellaceae (with only one species, *Amborella trichopoda*, native to New Caledonia) is the sister group to all others and the next two successive branches are the Nymphaeaceae (water lilies) and a group including the Illiciaceae, Schisandraceae and Austrobaileyaceae. Together, these three basal branches include seven families and about 177 species. The next branch is the magnoliids which include the monocots (102 families and 65 000 species) which, in turn, include the grasses, and within these the cereals. Other large groupings of species include the Ceratophyllales (26 families and 9400 species), asterids (107 families and 87 000 species) and rosids (149 families with 77 000 species; numbers of families and species quoted here are from Kuzoff and Grassler, 2000). Figure 1 also shows the phylogenetic positions of several model and crop species.

Arabidopsis is the first plant for which complete genomic sequence has been produced and sequencing of the rice genome is in progress. The complete sequences of these species will provide valuable new opportunities for comparative analyses because they represent highly diverged lineages within the angiosperms. Analyses of *Arabidopsis* and rice are complemented by extensive EST programmes that have already generated large amounts of information from a range of other species (Figure 1). This will reveal much about the commonality of gene sequences.

Within the grasses, taxonomists recognize four major lineages (Kellogg, 1998; inset in Figure 1). All contain important crop species and large-scale genomics programmes are under way in the Pooideae (wheat, barley, oats, rye and the *Festuca* and *Lolium* forage grasses), the Oryzoideae (rice) and the Panicoideae (maize, sorghum and pearl millet). Information from these projects will contribute to our knowledge on the grasses as a whole and will significantly assist analysis of the fourth group, the Chloridoideae, which contains crops such as finger millet (*Eleusine coracana*) and tef (*Eragrostis tef*). A robust phylogeny assists comparative genetic analysis in two ways. First,

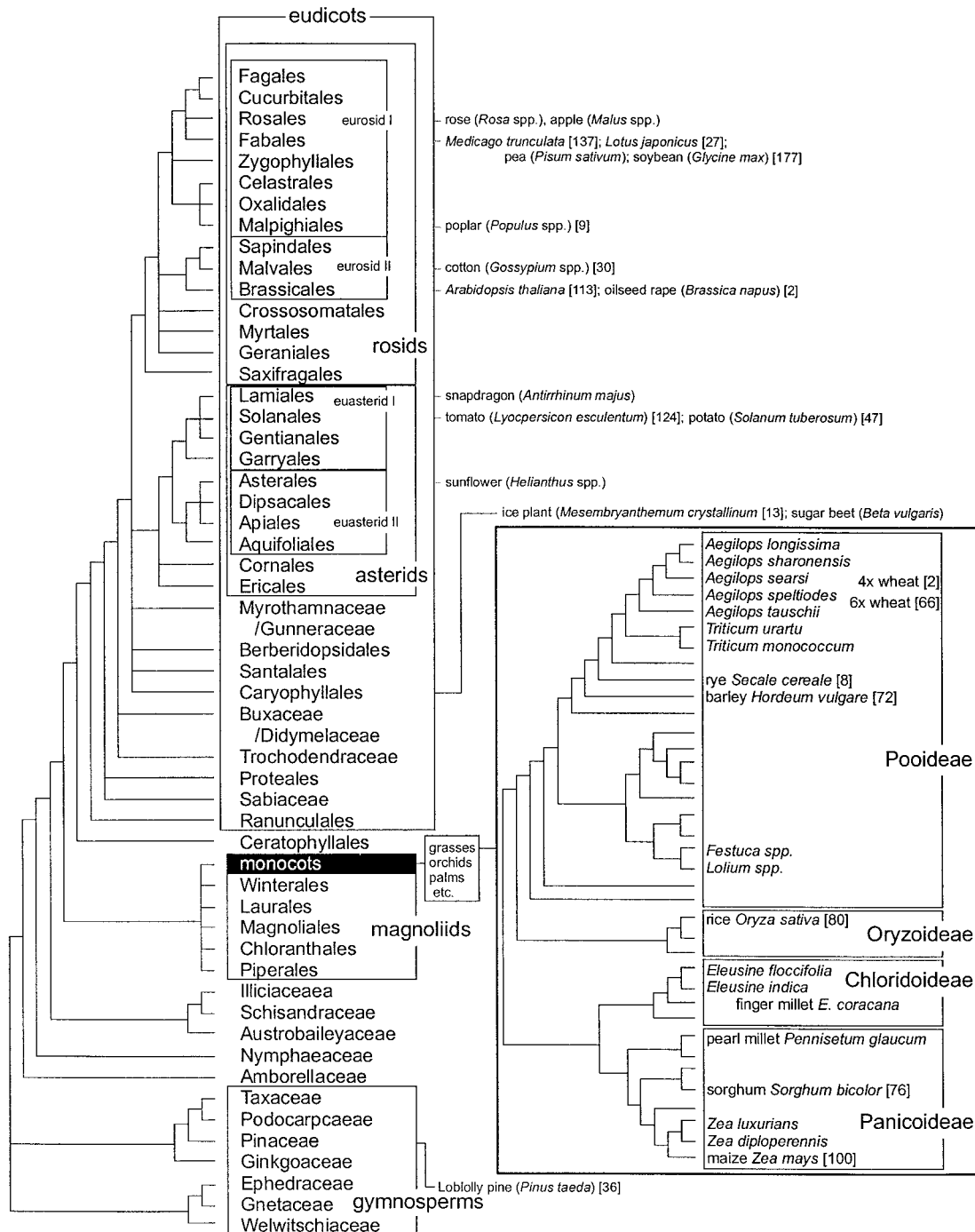


Figure 1. Phylogeny of seed plants illustrating the positions of several model and crop species. The overall phylogeny (left) is redrawn from Soltis *et al.* (1999) and that of the grasses (inset) from Kellogg (1998). Horizontal distances between branch points are all drawn to the same size and do not represent time scales. Figures in square brackets are thousands of EST sequences deposited in GenBank by mid August 2001.

it provides a logical framework for comparing gene action and chromosome evolution. Second, it helps identify the most likely origins of conserved and novel developmental and biochemical processes. Comparison of genetic information from different species will therefore provide valuable insights into cellular, developmental and biochemical processes that are deeply rooted in angiosperm evolution. An interesting question, therefore, is how much of plant biology will prove to be shared and how much will be novel to individual lineages.

Conserved genes and conserved functions

Single/low-copy genes

It is well known that many fundamental aspects of cellular metabolism are highly conserved throughout eukaryotes, reflecting their early origin and importance to cell viability. A recent example is provided by the analysis of proteins controlling translation termination (Inagaki and Doolittle, 2000). However, our interests in comparative genetics focus more on processes that are conserved, or diverged, between angiosperm groups.

An interesting example of conserved gene function comes from recent work on growth responses to gibberellin (GA) phytohormones. The GA-insensitive *GAI* gene of *Arabidopsis* was cloned by transposon tagging of the semi-dominant *gai* mutant that confers a dwarf phenotype (Peng *et al.*, 1997). The *GAI* gene was shown to have a distinctive 5' domain and a 3' region with homology to SCARECROW, a member of a family of transcription factors in *Arabidopsis*. Cereal homologues of *GAI* were cloned with the aid of sequence information from a rice EST and shown to correspond to the *Rht-1* semi-dwarfing genes of wheat and the *d8* dwarf mutant of maize (Peng *et al.*, 1999). A rice *GAI* homologue has recently been described by Ikeda *et al.* (2001). GA-insensitive semi-dwarf phenotypes in *Arabidopsis* and cereals resulted from deletions in the 5' part of the gene, showing that equivalent mutations resulted in equivalent phenotypes. This provides strong evidence for the conservation of GA-mediated growth regulation during angiosperm evolution. Interestingly, recent work in maize identified a SCARECROW homologue which, as in *Arabidopsis*, has a role in regulating root organization (Di Laurenzio *et al.*, 1996; Lim *et al.*, 2000). Thus, SCARECROW and *GAI* evolved before

the separation of monocot and dicot lineages and their respective functions have been conserved.

Gene families

A factor that significantly complicates comparative analysis is the presence of gene families. These evolve by duplication (which may be of individual genes, chromosome segments or whole genomes) and subsequent divergence. Gene families are therefore significantly affected by polyploidy, which is a common feature of angiosperm evolution. Gene duplication and divergence have long been recognized as important components of evolution, contributing to the development of increasingly complex forms (Holland, 1999; Wendel, 2000). It is estimated that at least 50% and possibly 70% of existing angiosperm species are polyploids (Wendel, 2000). The true figure is hard to estimate because ancient polyploids may be difficult to recognize and because genomes may appear polyploid due to internal duplications. This problem is illustrated by the genome sequence of *Arabidopsis* which shows extensive duplication (Blanc *et al.*, 2000; TAGI, 2000). This may be the result of ancient polyploidy but analysis of sequence divergence by Vision *et al.* (2000) favours a model in which there have been at least four large-scale internal duplication events followed by additional rearrangements and deletions.

Comparative approaches to the analysis of gene families are well illustrated by the extensive work that has been carried out on MADS-box transcription factors (for recent reviews see Hasebe, 1999; Blázquez, 2000; Lawton-Rauh *et al.*, 2000; Theissen *et al.*, 2000; Gutierrez-Cortines and Davies, 2000). Originally identified in plants as regulators of flower development they have long been known to have an ancient evolutionary origin as they are present in plants, animals and fungi (see recent review by Theissen *et al.*, 2000). In plants they have been most fully described in *Arabidopsis* where genomic sequence has allowed the family to be analysed in detail (Alvarez-Buylla *et al.*, 2000; Vergara-Silva *et al.*, 2000).

Analysis of other angiosperms shows general conservation of floral MADS-box gene roles (for example, Ambrose *et al.*, 2000; Kyojuka *et al.*, 2000), but with significant variation in some lineages (Theissen *et al.*, 2000). For example, in the ABC model of flower development (see Bowman, 1997 and Ng and Yanofsky, 2000 for recent reviews), C function (defining stamen and carpel identity and floral determinacy) is represented by a single gene in *Arabidopsis* (*AGA-*

MOUS) and two genes in maize (*ZAG1* and *ZMM2*) with partially redundant roles (Mena *et al.*, 1996). The B function genes of *Arabidopsis* (*APETALA3* and *PISTILLATA*) have overlapping expression patterns and function as a heterodimer that defines petal and stamen identity. Generally they show similar patterns of expression in other species but with some distinctive variants such as their expression in petals of plants from the Ranunculales, a group of distantly related eudicots (see Figure 1) (Kramer and Irish, 1999). It is likely that *APETALA3* and *PISTILLATA* arose by duplication of an ancestral B type gene and that *APETALA3* has been duplicated again in the evolution of the higher eudicots (Kramer *et al.*, 1998).

Some MADS-box genes that control floral organ identity in angiosperms have orthologues in gymnosperms, showing that their origins predate the evolution of flowers. For example, B function genes of angiosperms have been shown to have homologues that are expressed in reproductive tissues of gymnosperms, suggesting that their angiosperm functions represent an elaboration of a pre-existing role, in this case the development of pollen-bearing organs (Mouradov *et al.*, 1999; Sundstrom *et al.*, 1999). Becker *et al.* (2000) estimate that at least seven MADS-box genes were present in the common ancestor of angiosperms and gymnosperms and that the MADS-box family has evolved to similar levels of complexity in modern members of both groups. Recent work on club mosses (a sister group to other vascular plants) showed that the *LAMB1* gene is expressed in the strobilus, suggesting an ancient role for MADS-box genes in regulating the development of reproductive structures (Svensson *et al.*, 2000). Furthermore, analysis of MADS-box genes from the bryophyte *Physcomitrella patens* showed that they have a structure typical of higher plants and that the overall structure of the genes had evolved before mosses, ferns and seed plants diverged (Krogan and Ashton, 2000). However, the bryophyte genes are not obvious orthologues of existing higher-plant genes suggesting, as might be expected, that considerable evolution of the higher plant genes occurred after divergence.

The picture that emerges from the MADS-box studies is one of an evolutionarily ancient group in which gene duplication and divergence has allowed a progressive elaboration of existing functions. These changes are likely to be continuing in extant lineages, contributing to variation in plant form. A similar example comes from actin genes which are present as single-copy sequences in most green algae, in two

or three copies in bryophytes and in increasingly larger families in ferns and seed plants, reflecting the development of increasing numbers of tissue types (Bhattacharya *et al.*, 2000).

Studies also show, as might be expected, that different gene families (as well as regions within genes) evolve at very different rates. For example, the *CONSTANS* (*CO*) gene, which regulates flowering in response to long days in *Arabidopsis* (Putterill *et al.*, 1995), has homologues in a range of plants including barley (Griffiths, Coupland and Laurie, unpublished), rice (Song *et al.*, 1998; Yano *et al.*, 2000) and gymnosperms. These homologues are characterized by conserved 5' and 3' domains but the intervening region is much more diverged, suggesting that it is subject to fewer evolutionary constraints (Lagercrantz and Axelsson, 2000; Griffiths, Coupland and Laurie, unpublished). Thus, while genes are clearly recognizable as members of a *CO*-like family, the relationship to individual *Arabidopsis* genes is not always clear. Nevertheless, the rice heading date gene *Hd-1* has recently been shown to be a close homologue of *CO* (Yano *et al.*, 2000). Thus, *CO*-like genes have a conserved role in regulating flowering time.

The identification of *Hd-1* as a *CO* homologue is interesting because rice is a quantitative short-day plant while *Arabidopsis*, like barley, is a quantitative long-day plant. This suggests that a common photoperiod pathway promotes flowering in these species after exposure to inductive day lengths. How inductive day lengths are recognized in long-day and short-day plants remains to be established but recent work on *Arabidopsis* suggests that gene expression in relation to the circadian clock is important (Samach and Coupland, 2000; Barak *et al.*, 2000). Intriguingly, the *Arabidopsis* *TOC1* gene, which regulates circadian rhythm, has a basic C-terminal region of about 45 amino acids that is similar to a 3' region of *CO* (Strayer *et al.*, 2000), although the functional significance of this is currently unclear.

CO is one of about 80 genes so far implicated in the control of flowering time in *Arabidopsis* (Simpson *et al.*, 1999). Many of these genes can be assigned to the gibberellin, photoperiod, autonomous or vernalization promotion pathways (Simpson *et al.*, 1999). Recent work has shown that MADS-box genes play key roles in regulating timing of flowering in *Arabidopsis*, as well as in regulating floral structure. The *FLC* gene acts as a repressor of flowering whose expression is reduced by the action of genes in the autonomous pathways and by vernalization (Michaels and Amasino,

1999; Sheldon *et al.*, 1999). A second MADS-box gene (*SOCI* or *AGL20*) has been identified as a target of both *CO* and *FLC* (Samach *et al.*, 2000; Lee *et al.*, 2000). This, together with work on the meristem identity gene *LEAFY* (Blázquez and Weigel, 2000), is revealing how different pathways regulating flowering are integrated. An interesting question is whether the genetic control of flowering time will prove to be as conserved as the mechanisms regulating flower structure.

LEAFY homologues have been identified in several species and a particularly interesting analysis has been carried out in pea (*Pisum sativum*). The *UNIFOLIATA* (*UNI*) gene of pea was shown to be a *LEAFY* homologue but, in addition to effects on flower development, *uni* mutants also simplify the normal compound leaf to a single leaflet (Hofer *et al.*, 1997). A role in leaf development was not observed or predicted from the behaviour of *LEAFY* in *Arabidopsis* or the orthologous *FLORICAULA* gene in *Antirrhinum* (Coen *et al.*, 1990). Studies on *uni* show how genes can be used for additional functions in plant development and the value of comparative analyses of gene function in a range of diverse plants.

Rapidly evolving genes

Comparative mapping and sequence analysis has identified classes of genes that are evolving rapidly. A prime example of these are disease resistance genes which probably reflects the need for plants to maintain resistance to pathogens. An evolutionary 'arms race' between hosts and pathogens is a plausible reason for strong selection pressures favouring the evolution of novel sequences. An interesting question is whether the mechanisms that generate diversity are the same for all genes or whether some are specific to rapidly evolving types.

Analysis of disease resistance genes and related sequences within the genomes of diverse plants shows that gene duplication, intra- and intergenic recombination and gene conversion contribute to the formation of novel variation (Ellis *et al.*, 2000; Richter and Ronald, 2000). For example, the nucleotide-binding site/leucine-rich repeat (NBL-LRR) class of genes in several species shows a tendency to clustering, suggesting that tandem duplication and subsequent divergence of the gene copies are important factors in their evolution (Leister *et al.*, 1998; Pan *et al.*, 2000). Despite rapid sequence evolution these genes are still recognizably members of the same family, showing

that some selective constraints remain. Indeed, the presence of conserved domains was used to isolate barley and rice genes using PCR-based approaches (Leister *et al.*, 1998).

The results described in the above sections show how comparative approaches based on sequence homology can be used to isolate genes from cereals and other crop plants. This approach is most straightforward in cases where there are single genes or small gene families. For large gene families the situation is more complex. Thus, while members of families may be readily identified, as in the case of MADS-box transcription factors, it may be difficult or time-consuming to determine if homologues of individual members of the family exist in other species and, if so, whether they retain conserved roles. The interpretation of homology will be further complicated if genes have been duplicated in evolution and recruited to related but distinct functions. Homology-based gene isolation may be particularly problematic in disease resistance, where function is expected to be evolving more rapidly. In such cases, direct map-based cloning in the species of interest may be the preferred option. Whatever method is used it will be valuable to have efficient methods of integrating genetic maps and trait data across species. This is discussed in the next section.

Comparative mapping

Large-scale organization

Comparative mapping, recently reviewed by Gale and Devos (1997), Devos and Gale (2000) and Paterson *et al.* (2000), utilizes the existence of conserved sequences, usually cDNAs, to develop genetic maps that share common markers. Alignment of maps allows the structure of linkage groups in related organisms to be defined and this provides information on the taxonomic relationship between organisms and the common genetic control of specific aspects of plant phenotype in different species.

As would be expected, closely related species tend to have well conserved linkage groups, as revealed by studies of wheat and other Triticeae members. In this group, genomes are distinguished by a relatively small number of rearrangements (translocations and inversions) but the number of these varies among lineages. For example, *Aegilops longissima* and the closely related *A. umbellulata* differ from *A. tauschii* by 1 and

11 rearrangements, respectively, while *A. tauschii* differs from the more distantly related barley by only two (Devos and Gale, 2000). Variation in rates of chromosomal rearrangements have also been reported in comparisons of mammalian lineages (Burt *et al.*, 1999; Stanyon *et al.*, 1999).

Comparative maps naturally gravitate to the most extensively studied species. Rice has therefore become increasingly central to cereal comparisons because of its extensive RFLP maps and, more recently, the international commitment to genomic sequencing (Sasaki and Burr, 2000). The linkage groups of all grasses studied so far can be described in relation to a fairly small number of rice 'linkage segments', often bounded by centromere or telomere regions (Moore *et al.*, 1995; Devos and Gale, 2000) but it is not known whether this reflects functional constraints on how genes can be organized or merely a low rate of rearrangement fixation in grasses. As sequencing develops it will be interesting to analyse junction regions to determine if these show specific properties. Similar analyses in mammals have given mixed results with a variety of repeats interspersed with gene sequences in one case (Pletcher *et al.*, 2000) and a region rich in simple tandem repeats in another (Puttagunta *et al.*, 2000).

Fine-scale organization

Although the gross arrangements of chromosomes can be defined relatively easily, recent studies of cereals show that fine-scale collinearity is often disrupted (Keller and Feuillet, 2000; Tarchini *et al.*, 2000). This is most obvious for the candidate disease resistance genes in these studies but other more conserved sequences, such as alcohol dehydrogenase genes in the study of Tarchini *et al.* (2000), may also be affected. Similarly, although *Arabidopsis* genomic sequence is generally predictive of gene order in the Brassicaceae, which are closely related, comparison is complicated by duplications and divergence of gene content (Jackson *et al.*, 2000; O'Neill and Bancroft, 2000).

Not unexpectedly, loss of collinearity increases as more distantly related species are considered. Thus, a recent comparison of the gene content of a tomato BAC with *Arabidopsis* reveals seven genes collinear with a region on *Arabidopsis* chromosome IV and with the remainder on *Arabidopsis* chromosomes II, III and V. These are not likely to be the result of translocations but probably reflect patterns of duplication and sub-

sequent deletion which may be common features of plant genome evolution (Ku *et al.*, 2000). *Arabidopsis* and tomato are in the rosids and asterids, respectively (Figure 1) and, therefore, it is not surprising that comparisons between *Arabidopsis* and the more distantly related genome of rice have produced little evidence of the maintenance of extensive collinearity (van Dodeweerd *et al.* 1999; Devos *et al.* 1999; Paterson *et al.* 2000). However, these conclusions need to be seen in the light of possible biases in the regions selected for study. If the regions are rich in rapidly evolving sequences, such as resistance genes, loss of collinearity might appear more severe than in other regions of the genome. Therefore, it will be important to investigate other regions and in cereals this can be based on the emerging rice genomic sequence.

A further complication for comparative mapping is the possibility of convergent evolution, in which genes with equivalent functions have evolved independently in different lineages. It is not clear how widespread this phenomenon is in plants but it has been clearly demonstrated in secondary metabolite synthesis (Pichersky and Gang, 2000). In these cases, genes identified on the basis of similar effects on phenotype may be unrelated in sequence. However, a particular complication for comparative mapping is that in several examples enzymes with the same biochemical role have evolved from paralogous members of gene families (Alfenito *et al.*, 1998; Pichersky and Gang, 2000). In these cases the genes will be related in sequence but will probably show no collinearity.

The complications outlined above will probably discourage comparative map-based approaches to gene isolation from wheat and barley. If a trait can be mapped in a small genome species such as rice then cloning from that species is the logical route. Genes in other species can then be isolated by sequence homology. If the trait is not mapped in the small genome species then this approach becomes much riskier because the relevant gene may be absent. In such cases, map-based cloning in wheat or barley may be preferred because the target gene is certain to be present. However, even in these cases comparative mapping can provide useful markers even if it does not provide candidate genes. The use of rice genes as markers to home in on target regions of barley or wheat will become increasingly easy as rice genomic sequence becomes available.

Conventional trait mapping and comparative analysis

Comparative mapping needs to be interpreted with caution, for the reasons described above, but it should be emphasized that it remains valuable, especially in groups of related species such as the Triticeae or Brassicaceae. It identifies candidate orthologous trait loci and provides a valuable source of additional markers which may be useful even when fine-scale order is not preserved. However, an increasing proportion of comparative research is likely to be driven by knowledge of gene function in model organisms and in these cases it is gene sequence homology, not collinearity, that is the issue. The example of *GAI* illustrates this as the wheat *Rht* gene was isolated without the need for any comparison of *Arabidopsis* and cereal genetic maps. More important was the existence of segregating populations that confirmed co-segregation of the candidate gene and the trait.

This emphasizes the need to continue the development of wheat and barley populations that maximize the number of trait loci that are mapped. The need to analyse many populations will lead to an increasing use of high-throughput marker systems which, in contrast to RFLPs, may not be directly transferable between species. This difficulty can be reconciled with the desirability of aligning maps of related species by locating anonymous markers (such as simple sequence repeats derived from random genomic clones) on maps already characterized for RFLPs, and by developing additional markers within sequenced genes. Single-nucleotide polymorphisms (see Blake, this issue) are an attractive choice for the latter approach. New high-throughput marker systems will also enable much greater sampling of allelic variation within species and its relationship to trait variation (see Rafalski, this issue).

Allelic variation and its relationship to traits is attracting increasing attention in model species such as *Arabidopsis* as well as in crops. For example, many *Arabidopsis* ecotypes are winter annuals and require a period of low temperature (vernalization) to promote flowering. Early-flowering forms with no vernalization requirement occur and variation at *FLC* (see above) and *FRIGIDA* (*FRI*) are major determinants of this difference (Johanson *et al.*, 2000). Analysis of early-flowering alleles of *FRI* in a range of ecotypes revealed two mutant forms, showing that earliness has evolved at least twice. Given the number of flowering time genes that have been identified by mutagenesis of *Arabidopsis*, the involvement of just two suggests

that mutation at other loci may be incompatible with survival in the wild.

Similar conclusions come from studies of vernalization response in crosses between winter (vernalization-requiring) and spring (non-vernalization-requiring) cereals. A major gene regulating this response has been mapped in a similar position on the long arms of the homoeologous group 5 chromosomes of wheat (Galiba *et al.*, 1995), rye (Plaschke *et al.*, 1993), barley (Laurie *et al.*, 1995) and *Triticum monococcum* (Dubcovsky *et al.*, 1998). This clearly suggests the involvement of a common gene. It is not surprising that these species share a genetic pathway regulating vernalization response. The surprise is that the same step in the pathway appears to have been affected during the domestication of at least four different species. This suggests that few genes affecting vernalization can be mutated without unacceptable loss of performance in agriculture. Thus, comparative mapping of a range of crops may be useful for identifying genes that are particularly suitable for manipulation in breeding programmes.

Summary points

Comparative analysis is likely to become more important to wheat and barley as the genetic control of diverse aspects of plant biology is revealed in model systems, and will become progressively easier as genomic resources develop. Thus, wheat, barley or rye ESTs can be searched for homologues of genes identified in other species. Clones can then be mapped to determine if they are potential candidates for specific traits. Similarly, rice genomic sequence can be searched and used as a source of probes or related directly to wheat and barley using comparative maps. However, there remains the risk that homologues will not be identified in EST collections or will be absent or diverged in rice. An alternative, direct map-based approaches in wheat or barley, has the advantage that the gene of interest is certain to be present, but the disadvantage that considerable mapping, cloning and sequencing effort is required. A combined strategy, utilizing comparative information to expedite the isolation of relevant regions of the wheat or barley genome, looks like an attractive option.

To make best use of the information from model species we need to understand more about intraspecific variation in crops and to increase efforts on trait mapping. Furthermore, not all genes of interest will

be accessible via model systems, and in these cases genomic resources can be used for direct map-based cloning approaches. This makes the development of physical maps an important priority (see Morgante, this volume).

Although comparative analysis is likely to focus increasingly on genes regulating specific traits, comparative mapping will continue to be useful. As well as providing insights into species such as barley and wheat, comparative mapping will benefit crops such as pearl, foxtail and finger millets, which are important for the agriculture of developing countries but are relatively under-resourced in terms of research budgets.

Intuitively, we might expect that pathways regulating specific aspects of plant growth and development are most likely to be conserved between model species and the wild ancestors of crops. This is because ancestors often have attributes that are lost during domestication. Furthermore, crops are generally thought to have been domesticated on one to few occasions, probably from relatively small starting gene pools, and they have been subject to intense selection for yield and ease of cultivation. This has included selection for adaptation to regions well outside the ecogeographical limits of the progenitor species. These special evolutionary pressures may have resulted in novel variation. Therefore, comparative analysis would benefit from greater understanding of the mutations distinguishing domesticated forms from their wild ancestors.

Candidate genes can be identified in several ways but in all cases proof of function will be necessary. In addition to the identification of allelic variants in germplasm and mutant populations, there is a need to develop efficient ways of inactivating and over-expressing target genes. Gene function assays using transgenic lines can also be used to investigate the effect of changing levels or patterns of gene expression.

Comparative genetics is very much dependent on continued developments in bioinformatics. Good databases are needed to manage the vast amount of data produced by the sequencing and EST projects, the genetic and physical maps, the comparative data and the large sets of functional data gathered from micro-array studies. The establishment of data management systems needs to be paralleled by the development of efficient visualization and data analysis tools so that variation in gene sequence and gene expression can be correlated with variation in key agronomic traits. This will allow scientists in all areas of science to benefit from the wealth of information that is available.

Acknowledgements

This work was supported by the UK Biotechnology and Biological Sciences Research Council in the form of a David Phillips Research Fellowship to K.M.D. and by a grant-in-aid to the John Innes Centre.

References

- Alfenito, M.R., Souer, E., Goodman, C.D., Buell, R., Mol, J., Koes, R. and Walbot, V. 1998. Functional complementation of anthocyanin sequestration in the vacuole by widely divergent glutathione *S*-transferases. *Plant Cell* 10: 1135–1149.
- Alvarez-Buylla, E.R., Pelaz, S., Liljegren, S.J., Gold, S.E., Burgeff, C., Ditta, G.S., de Pouplana, L.R., Martinez-Castilla, L. and Yanofsky, M.F. 2000. An ancestral MADS-box gene duplication occurred before the divergence of plants and animals. *Proc. Natl. Acad. Sci. USA* 97: 5328–5333.
- Ambrose, B.A., Lerner, D.R., Ciceri, P., Padilla, C.M., Yanofsky, M.F. and Schmidt, R.J. 2000. Molecular and genetic analyses of the *silky1* gene reveal conservation in floral organ specification between eudicots and monocots. *Mol. Cell* 5: 569–579.
- Arabidopsis* Genome Initiative. 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408: 796–813.
- Barak, S., Tobin, E.M., Andronis, C., Sugano, S. and Green, R.M. 2000. All in good time: the *Arabidopsis* circadian clock. *Trends Plant Sci.* 5: 517–522.
- Becker, A., Winter, K.U., Meyer, B., Saedler, H. and Theissen, G. 2000. MADS-box gene diversity in seed plants 300 million years ago. *Mol. Biol. Evol.* 17: 1425–1434.
- Bhattacharya, D., Aubry, J., Twait, E.C. and Jurk, S. 2000. Actin gene duplication and the evolution of morphological complexity in land plants. *J. Phycol.* 36: 813–820.
- Blanc, G., Barakat, A., Guyot, R., Cooke, R. and Delseny, I. 2000. Extensive duplication and reshuffling in the *Arabidopsis* genome. *Plant Cell* 12: 1093–1101.
- Blázquez, M.A. 2000. Flower development pathways. *J. Cell Sci.* 113: 3547–3548.
- Blázquez, M.A. and Weigel, D. 2000. Integration of floral inductive signals in *Arabidopsis*. *Nature* 404: 889–892.
- Bowman, J.L. 1997. Evolutionary conservation of angiosperm flower development at the molecular and genetic levels. *J. Biosci.* 22: 515–527.
- Burt, D.W., Bruley, C., Dunn, I.C., Jones, C.T., Ramage, A., Law, A.S., Morrice, D.R., Paton, I.R., Smith, J., Windsor, D., Sazanov, A., Fries, R. and Waddington, D. 1999. The dynamics of chromosome evolution in birds and mammals. *Nature* 402: 411–413.
- Büsches, R., Hollricher, K., Panstruga, R., Simons, G., Wolter, M., Frijters, A., van Daelen, R., van der Lee, T., Diergaarde, P., Groenendijk, J., Töpsch, S., Vos, P., Salamini, F. and Schulze-Lefert, P. 1997. The barley *mlo* gene: a novel control element of plant pathogen resistance. *Cell* 88: 695–705.
- Coen, E.S., Romero, J.M., Doyle, S., Elliott, R., Murphy, G. and Carpenter, R. 1990. *FLORICAULA*: a homeotic gene required for flower development in *Antirrhinum majus*. *Cell* 63: 1311–1322.
- Devos, K.M. and Gale, M.D. 2000. Genome relationships: the grass model in current research. *Plant Cell* 12: 637–646.

- Devos, K.M., Beales, J., Nagamura, Y. and Sasaki, T. 1999. *Arabidopsis*-rice: will colinearity allow gene prediction across the eudicot-monocot divide? *Genome Res.* 9: 825–829.
- Di Laurenzio, L., Wysocka-Diller, J., Malamy, J.E., Pysh, L., Helariutta, Y., Freshour, G., Hahn, M.G., Feldmann, K.A. and Benfey, P.N. 1996. The *SCARECROW* gene regulates an asymmetric cell division that is essential for generating the radial organization of the *Arabidopsis* root. *Cell* 86: 423–433.
- Dubcovsky, J., Lijavetzky, D., Appendino, L. and Tranquilli, G. 1998. Comparative RFLP mapping of *Triticum monococcum* genes controlling vernalization requirement. *Theor. Appl. Genet.* 97: 968–975.
- Ellis, J., Dodds, P. and Pryor, T. 2000. Structure, function and evolution of plant disease resistance genes. *Curr. Opin. Plant Biol.* 3: 278–284.
- Gale, M.D. and Devos, K.M. 1997. Plant comparative mapping after 10 years. *Science* 282: 656–659.
- Galiba, G., Quarrie, S.A., Sutka, J., Morgounov, A. and Snape, J.W. 1995. RFLP mapping of the vernalization (*Vrn1*) and frost resistance (*Fr1*) genes on chromosome 5A of wheat. *Theor. Appl. Genet.* 90: 1174–1179.
- Gutierrez-Cortines, M.E. and Davies, B. 2000. Beyond the ABCs: ternary complex formation in the control of floral organ identity. *Trends Plant Sci.* 5: 471–476.
- Hasebe, M. 1999. Evolution of reproductive organs in land plants. *J. Plant Res.* 112: 463–474.
- Hirochika, H. 1997. Retrotransposons of rice: their regulation and use for genome analysis. *Plant Mol. Biol.* 35: 231–240.
- Hofer, J., Turner, L., Hellens, R., Ambrose, M., Matthews, P., Michael, A. and Ellis, N. 1997. *UNIFOLIATA* regulates leaf and flower morphogenesis in pea. *Curr. Biol.* 7: 581–587.
- Holland, P.W.H. 1999. Gene duplication: Past, present and future. *Semin. Cell Dev. Biol.* 10: 541–547.
- Ikeda, A., Ueguchi-Tanaka, M., Sonoda, Y., Kitano, H., Koshioka, M., Futsuhara, Y., Matsuoka, M. and Yamaguchi, J. 2001. *slender* rice, a constitutive gibberellin response mutant, is caused by a null mutation of the *SLR1* gene, an ortholog of the height-regulating gene *GAI/RGA/RHT/D8*. *Plant Cell* 13: 999–1010.
- Inagaki, Y. and Doolittle, W.F. 2000. Evolution of the eukaryotic translation termination system: origins of release factors. *Mol. Biol. Evol.* 17: 882–889.
- Jackson, S.A., Cheng, Z.K., Wang, M.L., Goodman, H.M. and Jiang, J.M. 2000. Comparative fluorescence in situ hybridization mapping of a 431-kb *Arabidopsis thaliana* bacterial artificial chromosome contig reveals the role of chromosomal duplications in the expansion of the *Brassica rapa* genome. *Genetics* 156: 833–838.
- Johanson, U., West, J., Lister, C., Michaels, S., Amasino, R. and Dean, C. 2000. Molecular analysis of *FRIGIDA*, a major determinant of natural variation in *Arabidopsis* flowering time. *Science* 290: 344–347.
- Keller, B. and Feuillet, C. 2000. Colinearity and gene density in grass genomes. *Trends Plant Sci.* 5: 246–251.
- Kellogg, E.A. 1998. Relationships of cereal crops and other grasses. *Proc. Natl. Acad. Sci. USA* 95: 2005–2010.
- Kramer, E.M. and Irish, V.F. 1999. Evolution of genetic mechanisms controlling petal development. *Nature* 399: 144–148.
- Kramer, E.M., Dorit, R.L. and Irish, V.F. 1998. Molecular evolution of genes controlling petal and stamen development: duplication and divergence within the *APETALA3* and *PISTILLATA* MADS-box gene lineages. *Genetics* 149: 765–783.
- Krogan, N.T. and Ashton, N.W. 2000. Ancestry of plant MADS-box genes revealed by bryophyte (*Physcomitrella patens*) homologues. *New Phytol.* 147: 505–517.
- Ku, H.M., Vision, T., Liu, J.P. and Tanksley, S.D. 2000. Comparing sequenced segments of the tomato and *Arabidopsis* genomes: Large-scale duplication followed by selective gene loss creates a network of synteny. *Proc. Natl. Acad. Sci. USA* 97: 9121–9126.
- Kuzoff, R.K. and Gasser, C.S. 2000. Recent progress in reconstructing angiosperm phylogeny. *Trends Plant Sci.* 5: 330–336.
- Kyozuka, J., Kobayashi, T., Morita, M. and Shimamoto, K. 2000. Spatially and temporally regulated expression of rice MADS box genes with similarity to *Arabidopsis* class A, B and C genes. *Plant Cell Physiol.* 41: 710–718.
- Lagercrantz, U. and Axelsson, T. 2000. Rapid evolution of the family of *CONSTANS* like genes in plants. *Mol. Biol. Evol.* 17: 1499–1507.
- Laurie, D.A., Pratchett, N., Bezant, J.H. and Snape, J.W. 1995. RFLP mapping of five major genes and eight quantitative trait loci controlling flowering time in a winter × spring barley (*Hordeum vulgare* L.) cross. *Genome* 38: 575–585.
- Lawton-Rauh, A.L., Alvarez-Buylla, E.R. and Purugganan, M.D. 2000. Molecular evolution of flower development. *Trends Ecol. Evol.* 15: 144–149.
- Lee, H., Suh, S.S., Park, E., Cho, E., Ahn, J.H., Kim, S.G., Lee, J.S., Kwon, Y.M. and Lee, I. 2000. The AGAMOUS-like 20 MADS domain protein integrates floral inductive pathways in *Arabidopsis*. *Genes Dev.* 14: 2366–2376.
- Leister, D., Kurth, J., Laurie, D.A., Yano, M., Sasaki, T., Devos, K., Graner, A. and Schulze-Lefert, P. 1998. Rapid reorganization of resistance gene homologues in cereal genomes. *Proc. Natl. Acad. Sci. USA* 95: 370–375.
- Lim, J., Helariutta, Y., Specht, C.D., Jung, J., Sims, L., Bruce, W.B., Diehn, S. and Benfey, P.N. 2000. Molecular analysis of the *SCARECROW* gene in maize reveals a common basis for radial patterning in diverse meristems. *Plant Cell* 12: 1307–1318.
- Lijavetzky, D., Muzzi, G., Wicker, T., Keller, B., Wing, R. and Dubcovsky, J. 1999. Construction and characterization of a bacterial artificial chromosome (BAC) library for the A genome of wheat. *Genome* 42: 1176–1182.
- Mathews, S. and Donoghue, M.J. 2000. The root of angiosperm phylogeny inferred from duplicate phytochrome genes. *Science* 286: 947–950.
- Mena, M., Ambrose, B.A., Meeley, R.B., Briggs, S.P., Yanofsky, M.F. and Schmidt, R.J. 1996. Diversification of C-function activity in maize flower development. *Science* 274: 1537–1540.
- Michaels, S.D. and Amasino, R.M. 1999. *FLOWERING LOCUS C* encodes a novel MADS domain protein that acts as a repressor of flowering. *Plant Cell* 11: 949–956.
- Moore, G., Devos, K.M., Wang, Z. and Gale, M.D. 1995. Grasses, line up and form a circle. *Curr. Biol.* 5: 737–739.
- Moulet, O., Zhang, H.B. and Lagudah, E.S. 1999. Construction and characterisation of a large DNA insert library from the D genome of wheat. *Theor. Appl. Genet.* 99: 305–313.
- Mouradov, A., Hamdorf, B., Teasdale, R.D., Kim, J.T., Winter, K.U. and Theissen, G. 1999. A DEF/GLO-like MADS-Box gene from a gymnosperm: *Pinus radiata* contains an ortholog angiosperm B class floral homeotic genes. *Dev. Genet.* 25: 245–252.
- Ng, M. and Yanofsky, M.F. 2000. Three ways to learn the ABCs. *Curr. Opin. Plant Biol.* 3: 47–52.
- O'Neill, C.M. and Bancroft, I. 2000. Comparative physical mapping of segments of the genome of *Brassica oleracea* var. *alboglabra* that are homoeologous to sequenced regions of chromosomes 4 and 5 of *Arabidopsis thaliana*. *Plant J.* 23: 233–243.
- Pan, Q.L., Liu, Y.S., Budai-Hadrian, O., Sela, M., Carmel-Goren, L., Zamir, D. and Fluhr, R. 2000. Comparative genetics of nucleotide binding site leucine-rich repeat resistance gene ho-

- mologues in the genomes of two dicotyledons: tomato and *Arabidopsis*. *Genetics* 155: 309–322.
- Paterson, A.H., Bowers, J.E., Burrow, M.D., Draye, X., Elisk, C.G., Jiang, C.X., Katsar, C.S., Lan, T.H., Lin, Y.R., Ming, R.G. and Wright, R.J. 2000. Comparative genomics of plant chromosomes. *Plant Cell* 12: 1523–1539.
- Peng, J.R., Carol, P., Richards, D.E., King, K.E., Cowling, R.J., Murphy, G.P. and Harberd, N.P. 1997. The *Arabidopsis GAI* gene defines a signalling pathway that negatively regulates gibberellin responses. *Genes Dev.* 11: 3194–3205.
- Peng, J.R., Richards, D.E., Hartley, N.M., Murphy, G.P., Devos, K.M., Flintham, J.E., Beales, J., Fish, L.J., Worland, A.J., Pelica, F., Sudhakar, D., Christou, P., Snape, J.W., Gale, M.D. and Harberd, N.P. 1999. 'Green revolution' genes encode mutant gibberellin response modulators. *Nature* 400: 256–261.
- Pichersky, E. and Gang, D.R. 2000. Genetics and biochemistry of secondary metabolites in plants: an evolutionary perspective. *Trends Plant Sci.* 5: 439–445.
- Plaschke, J., Börner, A., Xie, D.X., Koebner, R.M.D., Schlegel, R. and Gale, M.D. 1993. RFLP mapping of genes affecting plant height and growth habit in rye. *Theor. Appl. Genet.* 85: 1049–1054.
- Pletcher, M.T., Roe, B.A., Chen, F., Do, T., Do, A., Malaj, E. and Reeves, R.H. 2000. Chromosome evolution: the junction of mammalian chromosomes in the formation of mouse chromosome 10. *Genome Res.* 10: 1463–1467.
- Puttagunta, R., Gordon, L.A., Meyer, G.E., Kapfhammer, D., Lamerdin, J.E., Kantheti, P., Portman, K.M., Chung, W.K., Jenne, D.E., Olsen, A.S. and Burmeister, M. 2000. Comparative maps of human 19p13.3 and mouse chromosome 10 allow identification of sequences at evolutionary breakpoints. *Genome Res.* 10: 1369–1380.
- Putterill, J., Robson, F., Lee, K., Simon, R. and Coupland, G. 1995. The *CONSTANS* gene of *Arabidopsis* promotes flowering and encodes a protein showing similarities to zinc finger transcription factors. *Cell* 80: 847–858.
- Qiu, Y.L., Lee, J.H., Bernasconi-Quadroni, F., Soltis, D.E., Soltis, P.S., Zanis, M., Zimmer, E.A., Chen, Z.D., Savolainen, V. and Chase, M.W. 1999. The earliest angiosperms: evidence from mitochondrial, plastid and nuclear genomes. *Nature* 402: 404–407.
- Richter, T.E. and Ronald, P.C. 2000. The evolution of disease resistance genes. *Plant Mol. Biol.* 42: 195–204.
- Samach, A. and Coupland, G. 2000. Time measurement and the control of flowering in plants. *BioEssays* 22: 38–47.
- Samach, A., Onouchi, H., Gold, S.E., Ditta, G.S., Schwarz-Sommer, Z., Yanofsky, M.F. and Coupland, G. 2000. Distinct roles of *CONSTANS* target genes in reproductive development of *Arabidopsis*. *Science* 288: 1613–1616.
- Sasaki, T. and Burr, B. 2000. International Rice Genome Sequencing Project: the effort to completely sequence the rice genome. *Curr. Opin. Plant Biol.* 3: 138–141.
- Sato, Y., Sentoku, N., Miura, Y., Hirochika, H., Kitano, H. and Matsuoka, M. 1999. Loss-of-function mutations in the rice homeobox gene *OSH15* affect the architecture of internodes resulting in dwarf plants. *EMBO J.* 18: 992–1002.
- Sheldon, C.C., Burn, J.E., Perez, P.P., Metzger, J., Edwards, J.A., Peacock, W.J. and Dennis, E.S. 1999. The *FLF* MADS box gene: A repressor of flowering in *Arabidopsis* regulated by vernalization and methylation. *Plant Cell* 11: 445–458.
- Shirasu, K., Lahaye, T., Tan, M.W., Zhou, F.S., Azevedo, C. and Schulze-Lefert, P. 1999. A novel class of eukaryotic zinc-binding proteins is required for disease resistance signalling in barley and development in *C. elegans*. *Cell* 99: 355–366.
- Simpson, G.G., Gendall, A.R. and Dean, C. 1999. When to switch to flowering. *Annu. Rev. Cell Dev. Biol.* 15: 519.
- Smith, T.F. 1998. Functional genomics: bioinformatics is ready for the challenge. *Trends Genet.* 14: 291–293.
- Soltis, E.D. and Soltis, P.S. 2000. Contributions of plant molecular systematics to studies of molecular evolution. *Plant Mol. Biol.* 42: 45–75.
- Soltis, P.S., Soltis, D.E. and Chase, M.W. 1999. Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology. *Nature* 402: 402–404.
- Soltis, D.E., Soltis, P.S., Chase, M.W., Mort, M.E., Albach, D.C., Zanis, M., Savolainen, V., Hahn, W.H., Hoot, S.B., Fay, M.F., Axtell, M., Swensen, S.M., Prince, L.M., Kress, W.J., Nixon, K.C. and Farris, J.S. 2000. Angiosperm phylogeny inferred from 18S rDNA, *rbcl*, and *atpB* sequences. *Bot. J. Linn. Soc.* 133: 381–461.
- Song, J., Yamamoto, K., Shomura, A., Itadani, H., Zhong, H.S., Yano, M. and Sasaki, T. 1998. Isolation and mapping of a family of putative zinc-finger protein cDNAs from rice. *DNA Res.* 30: 95–101.
- Stanyon, R., Yang, F., Cavagna, P., O'Brien, P.C.M., Bagga, M., Ferguson-Smith, M.A. and Wienberg, J. 1999. Reciprocal chromosome painting shows that genomic rearrangement between rat and mouse proceeds ten times faster than between humans and cats. *Cytogenet. Cell Genet.* 84: 150–155.
- Strayer, C., Oyama, T., Schultz, T.F., Raman, R., Somers, D.E., Mas, P., Panda, S., Kreps, J.A. and Kay, S.A. 2000. Cloning of the *Arabidopsis* clock gene *TOC1*, an autoregulatory response regulator homolog. *Science* 289: 768–771.
- Sundstrom, J., Carlsbecker, A., Svensson, M.E., Svenson, M., Johanson, U., Theissen, G. and Engstrom, P. 1999. MADS-box genes active in developing pollen cones of Norway spruce (*Picea abies*) are homologous to the B-class floral homeotic genes in angiosperms. *Dev. Genet.* 25: 253–266.
- Svensson, M.E., Johannesson, H. and Engstrom, P. 2000. The *LAMB1* gene from the clubmoss, *Lycopodium annotinum*, is a divergent MADS-box gene, expressed specifically in sporogenic structures. *Gene* 253: 31–43.
- Tarchini, R., Biddle, P., Wineland, R., Tingey, S. and Rafalski, A. 2000. The complete sequence of 340 kb of DNA around the rice *Adh1-Adh2* region reveals interrupted colinearity with maize chromosome 4. *Plant Cell* 12: 381–391.
- Theissen, G., Becker, A., Di Rosa, A., Kanno, A., Kim, J.T., Munster, T., Winter, K.U. and Saedler, H. 2000. A short history of MADS-box genes in plants. *Plant Mol. Biol.* 42: 115–149.
- Van Deynze, A.E., Sorrells, M.E., Park, W.D., Ayres, N.M., Fu, H., Cartinhour, S.W., Paul, E. and McCouch, S.R. 1998. Anchor probes for comparative mapping of grass genera. *Theor. Appl. Genet.* 97: 356–369.
- van Dodeweerd, A.M., Hall, C.R., Bent, E.G., Johnson, S.J., Bevan, M.W. and Bancroft, I. 1999. Identification and analysis of homoeologous segments of the genomes of rice and *Arabidopsis thaliana*. *Genome* 42: 887–892.
- Vergara-Silva, F., Martinez-Castilla, L. and Alvarez-Buylla, E.R. 2000. MADS-box genes: Development and evolution of plant body plans. *J. Phycol.* 36: 803–812.
- Vision, T.J., Brown, D.G. and Tanksley, S.D. 2000. The origins of genomic duplications in *Arabidopsis*. *Science* 290: 2114–2117.
- Wendel, J.F. 2000. Genome evolution in polyploids. *Plant Mol. Biol.* 42: 225–249.
- Yano, M., Katayose, Y., Ashikari, M., Yamanouchi, U., Monna, L., Fuse, T., Baba, T., Yamamoto, K., Umehara, Y., Nagamura, Y. and Sasaki, T. 2000. *Hdl*, a major photoperiod sensitivity quan-

- titative trait locus in rice, is closely related to the *Arabidopsis* flowering time gene *CONSTANS*. *Plant Cell* 12: 2473–2483.
- Yu, Y., Tomkins, J.P., Waugh, R., Frisch, D.A., Kudrna, D., Kleinhofs, A., Bruggeman, R.S., Muehlbauer, G.J., Wise, R.P. and Wing, R.A. 2000. A bacterial artificial chromosome library for barley (*Hordeum vulgare* L.) and the identification of clones containing putative resistance genes. *Theor. Appl. Genet.* 101: 1093–1099.