Developmental Gene Evolution and the Origin of Grass Inflorescence Diversity

SIMON T. MALCOMBER,* JILL C. PRESTON,*
RENATA REINHEIMER,† JESSIE KOSSUTH*
AND ELIZABETH A. KELLOGG*

**Department of Biology, University of Missouri-St. Louis,*
One University Boulevard, Saint Louis, Missouri 63121
†*Vegetal Morphology, Facultad de Ciencias Agrarias (UNL),*
Kreder 2805, 3080 Esperanza, Santa Fe, Argentina

Advances in Botanical Research, Vol. 44
Incorporating Advances in Plant Pathology

0065-2296/06 $35.00
DOI: 10.1016/S0065-2296(06)44011-8

ABSTRACT

Grass inflorescences are diverse, developmentally complex, and provide many of the taxonomic characters used to differentiate the estimated 10,000 grass species. Here we review grass inflorescence development in detail and discuss which genes are involved at each developmental stage. We demonstrate that grass inflorescence development is complex, with multiple structures that are not present in *Arabidopsis*. New and published phylogenetic analyses of genes involved at each developmental stage indicate that the maize *FASCIATED EAR2* (*FEA2*) and *Arabidopsis CLAVATA2* (*CLV2*) genes are the sole remaining co-orthologs following multiple rounds of whole-genome duplication. Analyses of *BARREN STALK1/LAX PANICLE* (*BA1/LAX1*), *FRUITFULL* (*FUL*), *INDETERMINATE SPIKELET1* (*IDS1*), *KNOTTED1* (*KN1*), *LEAFY HULL STERILE1* (*LHS1*), and *RICE CENTRORADIALIS1/2* (*RCN1/2*) indicate that these genes are members of grass- or monocot-specific small gene families. The complex pattern of gene relationships mirrors a complex pattern of functional evolution. Maize FEA2 and *Arabidopsis* CLV2 have nonidentical roles, whereas distantly related grass KN1-like and RCN1/2 proteins show functional convergence and conservation, respectively. Duplications near the base of grasses in *BA1/LAX, FUL, IDS1* and *LHS1* have led to diverse roles in grass inflorescence development. We conclude that developmental gene duplication followed by functional diversification appears to have played a major role in the evolution of novel morphological structures within grass inflorescences.

I. INTRODUCTION

A. TOOLS FOR EVOLUTIONARY DEVELOPMENTAL GENETICS

The study of evolutionary developmental genetics requires a good phylogeny, precise description of developmental morphology, and a set of genes connected to the morphology (Baum, 2002; Kellogg, 1996). Over the last 15 years, increasingly robust phylogenies have become available for increasing numbers of taxa (Soltis *et al.*, 2005; Stevens, 2001 onwards). Description of developmental morphology, once pursued by only a handful of talented morphologists, is now undergoing a renaissance as more and more scientists see accurate description of the phenotype as key to understanding its evolution. The *Arabidopsis thaliana* genome sequence (Arabidopsis Genome Initiative, 2000), and several rice genome sequences (Goff *et al.*, 2002; Yu

et al., 2002, 2005) provide a list of the genes, and have vastly accelerated the effort to connect genotype with phenotype.

The major model systems provide the tools for studying evolution of development and are often the only species in which gene function can be assessed rigorously. But the model systems are just that—models. To understand diversification, one needs to look at diversity (see also comments by Gewin, 2005). The model systems are thus tools for understanding evolution, just as phylogenies and morphological descriptions are (Kellogg and Shaffer, 1993).

B. THE GRASS FAMILY

The grass family (Poaceae or Gramineae) is an ideal system for the study of evolution of development. The family is large (ca. 10,000 species) and morphologically diverse (Clayton and Renvoize, 1986; Watson and Dallwitz, 1992 onwards), and the developmental morphology of many species is well known. The family includes rice, with complete genome sequences from both *indica* and *japonica* varieties (Goff *et al.*, 2002; Yu *et al.*, 2002). It also includes maize, which has an unparalleled set of well-characterized morphological mutants, plus an ongoing effort to sequence the entire genome. Sorghum, wheat, barley, ryegrass (*Lolium*), meadow fescue (*Festuca pratensis*), sugarcane, tef, foxtail millet, and pearl millet also have varying amounts of genetic resources (genome sequencing projects, EST collections, genetic maps, well-organized stock centers). Informatic tools are available for comparing genetic maps (Gramene, Ware *et al.*, 2002) and for searching mutants and genes (Maize GDB, Lawrence *et al.*, 2005).

The grass family is monophyletic (Grass Phylogeny Working Group, 2001). All members of the family share a uniquely derived embryo structure. The embryo is highly differentiated, with clearly organized shoot and root meristems and two or more leaves initiated. This represents a heterochronic change relative to the grass sister groups and presumed ancestors, in which embryo development is accelerated relative to seed maturation (Kellogg, 2000a). In the sister taxa of the grasses (Joinvilleaceae and Ecdeiocoleaceae), the embryo reaches only an undifferentiated globular stage before the seed matures (Campbell and Kellogg, 1987; Rudall *et al.*, 2005). The grass embryo also has a unique haustorial organ, the scutellum, which could be derived from a cotyledon although it is so highly modified that direct comparisons are impossible.

A robust phylogeny is available for the family and provides the basis for the classification (Grass Phylogeny Working Group, 2001). Sister to the rest of the family is the tiny subfamily Anomochlooideae; the four species in this subfamily have the characteristic grass embryo, but lack spikelets, the flowering units of most grasses, composed of glumes, lemmas, paleas, lodicules, stamens,

and pistil. Instead of spikelets, members of Anomochlooideae have "spikelet equivalents," which lack structures that are clearly homologous to glumes, lemmas, and paleas (Clark and Judziewicz, 1996). Subfamily Pharoideae is the earliest lineage with a true spikelet and is sister to all remaining spikelet-bearing grasses. Some time after divergence of Pharoideae, multiflowered spikelets arose. The earliest-diverging lineage of this group is Puelioideae. The major radiation of the grasses occurred much later, after the grasses had been around for several tens of millions of years. Bambusoideae, Pooideae, Ehrhartoideae, and a lineage leading to the PACCAD clade originated at this point, perhaps 40–50 million years ago, but the order of events is uncertain. The Grass Phylogeny Working Group (2001) phylogeny suggests that Bambusoideae, Pooideae, and Ehrhartoideae shared a common ancestor (and thus form the BEP clade), but support for this is weak. Nonetheless, each of the three subfamilies is strongly supported as monophyletic by both morphological and molecular data. The PACCAD clade includes the remaining subfamilies (Panicoideae, Arundinoideae, Centothecoideae, Chloridoideae, Aristidoideae, and Danthonioideae), relationships among which are not well supported.

Many morphological distinctions among grass taxa are based on characters of the inflorescence. In this chapter, we discuss this variation, and the current stage of knowledge about its development and genetics.

C. GRASS INFLORESCENCES VS *ARABIDOPSIS*

The grass inflorescence is a novel structure that is developmentally intricate, evolutionarily intriguing, and agronomically important. Its architecture controls both pollination and seed set, and is thus a target of natural and human selection. The major eudicot model, *Arabidopsis thaliana*, provides some useful information on the genes affecting inflorescence morphology but is not a substitute for looking directly at the grasses themselves.

The inflorescence of *Arabidopsis* (and most Brassicaceae) is simple (Fig. 1, Whipple and Schmidt, Chapter 10; this volume, Figs. 2 and 4). After receiving the appropriate environmental or endogenous signal, the shoot apical meristem produces a few leaves with inflorescence meristems in their axils, and then ceases to produce leaves entirely and begins to produce lateral floral meristems, in a phyllotactic spiral (Shannon and Meeks-Wagner, 1991; Smyth *et al.*, 1990). Unlike most angiosperms, the floral meristems are not subtended by bracts or bracteoles, which appear to be suppressed (Long and Barton, 2000) as in most Brassicaceae (Stevens, 2001 onwards). The shoot apical meristem itself (now the inflorescence meristem) eventually ceases to produce floral meristems but remains itself undifferentiated (Shannon and Meeks-Wagner, 1991). Secondary

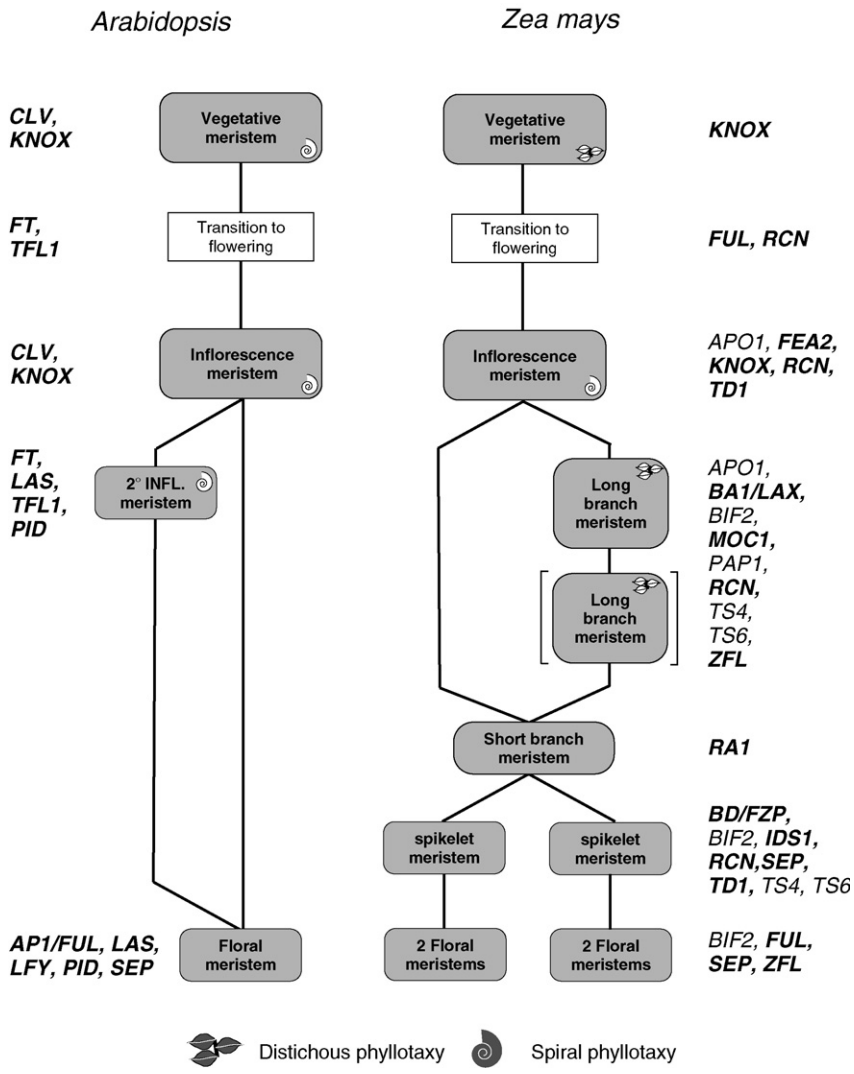

Fig. 1. Diagram comparing inflorescence development in *Arabidopsis* and maize, starting with the vegetative meristem and ending with the floral meristem, and the genes discussed in the text that are known to affect particular developmental stages. Short branch meristem is also referred to as the spikelet pair meristem in the maize literature. Boldface, cloned genes. *AP1, APETALA1; APO1, ABERRANT PANICLE ORGANIZATION1; BA1, BARREN STALK1; BD, BRANCHED SILKLESS1; BIF2, BARREN INFLORESCENCE2; CLV, CLAVATA; FEA2, FASCIATED EAR2; FON1, FLORAL ORGAN NUMBER1; FT, FLOWERING LOCUS T; FUL, FRUITFULL; FZP, FRIZZY PANICLE1; IDS1, INDETERMINATE SPIKELET1; KNOX, KNOTTED1-LIKE HOMEOBOX; LAS, LATERAL SUPPRESSOR; LAX, LAX PANICLE1; LFY, LEAFY; MOC1, MONOCULM1; PAP1, PANICLE PHYTOMER1; PID, PINOID; RA1, RAMOSA1; RCN, RICE CENTRORADIALIS; TD1, THICK TASSEL DWARF1; TFL, TERMINAL FLOWER1; TS4, TASSELSEED4; TS6, TASSELSEED6; SEP, SEPALLATA; ZFL, ZEA LEAFY.*

Fig. 2. Phylogenetic relationships of sampled plant taxa, as presented by the Angiosperm Phylogeny Web (Stevens, 2001 onwards). Grasses and *Arabidopsis* in boldface. Corn ear, grass clade; flower, *Arabidopsis*.

inflorescences in the leaf axils reiterate the developmental pattern of the main axis. The *Arabidopsis* inflorescence is thus an ideal system in which to study the transition from vegetative to inflorescence meristem, and the production of floral meristems from the inflorescence meristem. However, *Arabidopsis*

provides only limited insight into the various mechanisms underlying the formation of inflorescence branches, changes in phyllotaxis, and termination of the inflorescence axis (or axes). Processes that do not occur in *Arabidopsis* need to be studied in plants in which they occur.

The grass inflorescence is, by contrast with that of Brassicaceae, baroque (Fig. 1, Whipple and Schmidt, Chapter 10; this volume, Figs. 2 and 4). The transition to flowering is often accompanied by a change in phyllotaxis, the inflorescence meristem generally produces branches which themselves may produce branches, and the branches may simply cease to develop (as in *Arabidopsis*) or may terminate in a unique structure, a spikelet, which is morphologically a contracted, flower-bearing branch, but also itself has some similarities to a flower. In classic taxonomic literature, agrostologists (grass taxonomists) commonly make an analogy between a spikelet and a flower, as the terminal structure on a branch, and then attempt to classify grass inflorescences as: (1) spikes, with the spikelets attached directly to the inflorescence axis, as in wheat, barley, and rye; (2) racemes, with the spikelets on pedicels, as in *Brachypodium* or *Brachyelytrum*; or (3) panicles, with the inflorescence branches themselves branched, as in most familiar grasses such as oats, fescue, sorghum, or proso millet (Clark and Pohl, 1996). This set of descriptors is misleading, and implies that grass inflorescence development is quite uniform, whereas grass inflorescences vary in multiple characteristics, some of which are described in more detail in a later section. In addition, the classic terminology is of limited use for describing the inflorescences of some species, such as big bluestem (*Andropogon gerardii*) or bamboo (e.g., *Phyllostachys*), which have intricate combinations of bracts, spathes, and inflorescence branches with both sessile and pedicellate spikelets. It is common to find descriptions such as the following for the inflorescence of *Andropogon*: "Inflorescence of spicate main branches, or paniculate (usually with paired or digitate "racemes," these often spatheate and aggregated into false panicles)... Inflorescence... a complex of "partial inflorescences" and intervening foliar organs..." (Watson and Dallwitz, 1992 onwards). The work of Vegetti and Weberling (1996) provides a better description of adult inflorescences but still attempts to create a typology for the entire flower-bearing portion of the plant, which makes it difficult to hypothesize underlying developmental processes.

Grass inflorescences may be better described by components of the phenotype that can vary in a combinatorial manner (Kellogg, 2000b). Specifically, any meristem can have one of three fates: it can terminate in a spikelet, it can continue producing lateral meristems, or it can cease developing. If it produces lateral meristems, then each lateral has the same three possible fates, and similarly with the lateral meristems on the lateral meristems. Importantly, the fate of each order of branching is potentially independent

of the fate of each other order, so that over evolutionary time, virtually all conceivable combinations occur.

In this chapter, we describe the development of grass inflorescences in more detail and review current knowledge of the various genes involved at each stage of development. We show that grass inflorescences are far more complex than those of *Arabidopsis*. One could imagine multiple ways that such complexity might have evolved, but one possibility would be the origin of novel genes and gene functions. We show that gene duplication has been common among genes controlling morphogenesis. The sheer number of available genes and presumed gene functions may help explain the morphological variety that exists among grasses.

D. DUPLICATE GENES, THE RAW MATERIAL FOR EVOLUTION OF NOVEL FUNCTION

Haldane (1932, 1933), and later Ohno (1970) suggested that duplicate genes should accumulate mutations and become pseudogenes because there would be no selective pressure on them to retain function. This should happen quite rapidly, although the exact speed depends on population size (Walsh, 1995). Only rarely would a duplicate gene acquire a new function (neofunctionalization), and thus contribute to the diversity of the genome.

As gene and genomic sequences have accumulated, it has become abundantly clear that duplicated genes are common, contra the predictions of Haldane and Ohno. Most genes, in fact, belong to multigene families, which are generated by a variety of gene and genome-duplication processes (Zhang, 2003). Although duplicated genes are lost quite frequently, particularly after polyploidization events (Chantret *et al.*, 2005), an appreciable number of them persist for millions of years. The discrepancy between prediction and observation has led to a new round of theories to explain the long persistence of some duplicates.

One suggestion is that duplicate genes might divide up the ancestral function, so that, rather than acquire new functions, they simply became subfunctionalized (Force *et al.*, 1999; Hughes, 1999; Lynch and Force, 2000). Mechanisms have been proposed whereby this could occur via neutral processes, without having to invoke adaptation (positive selection) at all (Force *et al.*, 2005). Consistent with this theory, duplicate genes can be found that have apparently partitioned the ancestral expression domain or function (reviewed by Kellogg, 2003; Moore and Purugganan, 2005); much of the evidence for subfunctionalization comes from studies of gene expression.

But subfunctionalization cannot be the entire story, with new duplicates dividing up ancestral functions ever more finely. New functions must arise from time to time. Clearly, over time, a combination of subfunctionalization and neofunctionalization must occur. Models have been proposed by which

this might happen (He and Zhang, 2005). Neofunctionalization may lead to a pattern of selective sweeps such as found by Moore and Purugganan (2003) in the history of multiple genes in *Arabidopsis*.

In summary, gene duplication can lead to diversification of function in many ways. Ancestral functions can be retained by one copy, or partitioned between two copies, and new functions can evolve in one or both copies. Acquisition of new functions is one way to generate novel morphology.

In the remainder of this chapter, we review knowledge of cloned genes affecting inflorescence architecture in the grasses, and present gene trees for some of them. The pattern of duplication is such that no genes in the grasses will have an ortholog in *Arabidopsis*. The *Arabidopsis* genome bears traces of a whole-genome duplication event that occurred at or near the base of the Brassicaceae (Bowers *et al.*, 2003). Likewise, whole-genome duplications are documented near the base of eudicots (Bowers *et al.*, 2003) and at or near the base of the grasses (Paterson *et al.*, 2004; Yu *et al.*, 2005). Thus, without gene loss, no one-to-one correspondence is expected between an *Arabidopsis* gene and a related gene in grasses because of these three rounds of duplication alone. Additional duplications within particular gene families lead to even less correspondence.

The pattern of extensive gene duplication greatly limits the utility of the traditional terminology used to describe gene relationships. "Orthologs" were originally defined as two genes in two different species that derive from a speciation event, and "paralogs" as genes that derive from a single gene that was duplicated within the genome (Fitch, 1970). Three whole-genome duplication events, such as occurred near the base of the Brassicaceae, eudicots, and grasses have produced three sets of gene paralogs, none of which is truly orthologous to any individual gene in the other lineages. This is even the case when only a single (paralogous) gene remains. To overcome this problem, Sonnhammer (2002) coined the term co-orthologs. Co-orthologs are defined as paralogs that are produced by duplications of orthologs subsequent to a given speciation event (see Sonnhammer, 2002 for additional details). Using this terminology, if a clade of paralogous grass genes is sister to a clade of paralogous *Arabidopsis* genes, then the two sets of genes are considered to be co-orthologous.

In species separated by multiple rounds of whole-genome duplication, such as *Arabidopsis* and grasses, the function of co-orthologs is unlikely to be fully conserved. If gene duplication is generally followed by divergence in function, then multiple duplications should lead to multiple variants in gene-expression pattern, biochemical function, and/or developmental role. Because of this, we expect that conservation of function between *Arabidopsis* and grasses should be the exception rather than the rule.

We show in a later section that multiple rounds of duplication have led to novel genes with novel functions in some cases, whereas duplicate genes appear to have been lost and functions retained in others.

II. METHODS

A. CHOICE OF GENES AND TAXA

Genes were chosen because: (a) they represent different stages of inflorescence and floral development and (b) some information is available on function and sequence diversity (Fig. 1). Other genes have been analyzed in more detail elsewhere in the literature, and we review these only briefly here, referring interested readers to the original publications for more detail. The genes included in this chapter are known to play roles in meristem identity and maintenance, inflorescence meristem identity and size, inflorescence branching, inflorescence determinacy, and spikelet development.

Taxon sampling outside the grasses reflects largely data available in GenBank (http://www.ncbi.nlm.nih.gov/); new sequences for some genes were generated from additional grasses and near grass relatives (Fig. 2). For some gene families, such as *KNOTTED1*-like genes, sequences were available from throughout land plants. In other less-studied gene families sampling was more biased toward model species such as *Arabidopsis*, maize, rice, and tomato. In all cases, we sampled broadly within grasses to explore patterns of gene duplication within the major diversification of the family [the BEP and PACCAD clades (Grass Phylogeny Working Group, 2001)]. In several cases, we were able to include early diverging grasses (*Streptochaeta* and *Pharus*) and near grass relatives (such as *Joinvillea*, Joinvilleaceae) to explore patterns of gene duplication near the origin of grasses. *Streptochaeta* and *Anomochloa* are the sole members of the earliest diverging grass lineage (subfamily Anomochlooideae) and lack true spikelets whereas *Pharus* is a member of the earliest diverging lineage with a true spikelet (Grass Phylogeny Working Group, 2001). Inclusion of these taxa is thus essential to understand how particular genes, gene duplications, and functional diversification might have been involved in the origin of the spikelet.

B. DATA, ALIGNMENT, AND POLYMERASE CHAIN REACTION PRIMER DESIGN

Gene sequences of *A. thaliana TERMINAL FLOWER1* (*AtTFL1*), *Oryza sativa LEAFY HULL STERILE1* (*OsLHS1*), and *Zea mays BARREN STALK1* (*ZmBA1*), *FASCIATED EAR2* (*ZmFEA2*), *INDETERMINATE*

SPIKELET1 (*ZmIDS1*), and *KNOTTED1* (*ZmKN1*) were retrieved from GenBank. Similar sequences were identified using BLAST, BLASTX, BLASTP, and tBLASTX (Altschul *et al.*, 1997) searches at NCBI (http://www.ncbi.nlm.nih.gov/BLAST/) and PlantGDB (http://www.plantgdb.org/PlantGDB-cgi/blast/PlantGDBblast). Nucleotide sequences were aligned, based on the conceptual amino acid translation, using CLUSTALX (Jeanmougin *et al.*, 1998) and MacClade 4 (Maddison and Maddison, 2003). Only sequences or regions within sequences that could be aligned reliably as determined by visual inspection of the aligned matrix of protein sequences were included in subsequent analyses. All aligned matrices were submitted to TREEBASE (www.treebase.org). To identify relationships among grass genes for designing polymerase chain reaction (PCR) primers, preliminary maximum parsimony trees were generated using heuristic searches with TBR and MULPARS on and 100 random addition sequence additions within PAUP*4.0 (Swofford, 2000). Gene-specific PCR primers (Table I) were designed based on available grass sequences (usually only rice and maize) using either OLIGO 4.0 (Molecular Biology Insights, Inc., Cascade, USA) or Primaclade (Gadberry *et al.*, 2005).

C. DNA ISOLATION, PCR, SUBCLONING, AND SEQUENCING

Plants were grown under standard greenhouse conditions at Missouri Botanical Garden or at the University of Missouri-Saint Louis from USDA seed stocks. Total DNA was isolated using either CTAB (Hillis *et al.*, 1996) or SDS-based (Dellaporta, 1994) extraction protocols. Total RNA was extracted from young inflorescences using RNAwiz solution (Ambion, Austin, TX), according to the manufacturer's instructions. cDNA was generated from the extracted total RNA using Superscript III (Invitrogen, Carlsbad, CA) following the manufacturer's instructions using a polyT with adaptor primer (5′-CCGGATCCTCTAGAGCGGCCGCTTTTTTTTTTTTTTTT-3′).

Double-stranded PCR products were amplified in 50-μl reactions containing 2 units of *Taq* polymerase (Promega Corp., Madison, WI), 5 μl of 10× reaction buffer, 5 μl of 25 mM $MgCl_2$, 2 μl dNTP (2.5 mM stock solution), 20 μM each of the gene-specific forward and reverse primers (Table I), 5% DMSO (by volume), and approximately 200 ng of genomic DNA or 4 μl of cDNA. PCR reactions were performed using a hot-start PCR profile with a gene-specific annealing temperature (Table I). PCR fragments were purified and subcloned as described by Malcomber and Kellogg (2004). Dideoxy sequencing was conducted using the Big Dye 3.1 terminator cycle sequencing protocol (Applied Biosystems, Foster City, CA) using the plasmid primers T7 and SP6. Sequencing reactions were analyzed on either an ABI 377 or ABI 3100 automated DNA sequencer (Applied Biosystems). Base calling

TABLE I

Polymerase Chain Reaction (PCR) Primer Sequences for New Gene Sequences, Amplification Conditions and Proportion of the Coding Region Amplified

Gene	PCR primer name	Primer sequence	Annealing temperature	Amplicon size (cDNA or gDNA)	Percent coverage of rice (*Os*) or maize (*Zm*) gene coding region
FASCIATED EAR2 (*FEA2*)	FEA2-228F	5′-CTC CCC GCG TCT CTC CTT-3′	55 °C	~500 bp, cDNA	27% (*ZmFEA2*)
	FEA2-579R	5′-GGT TAC CAG CCA GTG ATA-3′			
INDETERMINATE SPIKELET 1 (*IDS1*)	IDS1-338F	5′-ATG GTG CTG GAT CTC AAT GT-3′	57 °C	~ 650 bp (gDNA)	50% (*ZmIDS1*)
	IDS1-AP2R	5′′-GGT SAC GCC CCT GTA CTG CGA GCT-3′′			
KNOTTED1 (*KN1*)	KN1-1F	5′′-ATG GAG GAG ATC DSC CAM CAC TT-3′′	57 °C	~1000 bp (cDNA)	96% (*ZmKN1*)
	KN1-1041R	5′′-AGC CCG CYG TCG TTG AYG AAG TG-3′′			
RICE CENTRORADIALIS (*RCN*)	RCN1-16F	5′-GAG CCT CTT RTT GTD GGK CGB GTS ATY GG-3′	59 °C	~1000 bp (gDNA)	71% (*OsRCN1*)
	RCN1-393R	5′-GCG YCT CCT GGC WGC AGT CTC YCT CTG-3′			
SEPALLATA-LEAFY HULL STERILE1 (*LHS1*)	MADS-1F	5′′-ATG GGT MGS GGS AAG GTG GAG CTG AAG CGG-3′	57 °C	705 bp (cDNA)	100% (*OsLHS1*)
	LHS1-633R	5′′-TAT CCA KCC RGA TSG RMY RTG YTC ATT SG-3′′			
SEPALLATA3 (*SEP3*)	MADS-1F	See above	57 °C	600 bp, cDNA	75% (*OsMADS8*)
	SEP3-658R	5′- GAT CTG CAR GGT KGG CTC KGC KGC-3′			

within the chromatograms was checked and confidence scores assigned using PHRED (Ewing *et al.*, 1998). Only nucleotide sequences with PHRED scores >20 were used in subsequent analyses. Contiguous alignments were edited using Seqman II (DNASTAR Inc., Madison, WI), and all sequences were submitted to GenBank (DQ317417–DQ317439).

D. ALIGNMENT AND PHYLOGENETIC ANALYSIS

Nucleotide sequences were aligned, based on the conceptual amino acid translation, using RevTrans (Wernersson and Pedersen, 2003) or a combination of CLUSTALX and MacClade 4 (Maddison and Maddison, 2003), before being adjusted manually using MacClade 4. Optimal parameters for the Bayesian analyses were determined using MrModeltest2.0 (Nylander, 2004). Bayesian phylogenetic estimates were produced using MrBayes 3.1 (Huelsenbeck and Ronquist, 2001) on the parallel-processing cluster at the University of Missouri-St. Louis using 10 million generations and default flat priors. Trees were sampled every 500 generations and burn-in was determined empirically by plotting likelihood score against generation number. After burn-in trees had been removed, clade credibility (Bayesian posterior probability) was estimated using MrBayes (Huelsenbeck and Ronquist, 2001). The best topology visited by the Markov chain was retrieved by sorting the MrBayes probability (.p) output file. This topology is expected to be qualitatively identical to the maximum likelihood topology because flat priors were specified prior to conducting the analysis (Larget and Simon, 1999). Only branches supported by 0.95 or greater posterior probability were considered well supported, based on published simulation studies (Alfaro *et al.*, 2003), although other studies have noted that Bayesian phylogenetic analyses can overestimate clade credibility values (Simmons *et al.*, 2004).

III. MORPHOLOGICAL VARIATION AND MOLECULAR EVOLUTION OF GENES IN GRASS INFLORESCENCES

A. FORMATION OF LATERAL STRUCTURES AND NONCORRELATION OF MERISTEM FATES

1. Morphological variation

An inflorescence meristem forms lateral structures that themselves become meristems, either branches or spikelets. Although a few grasses have a single terminal spikelet (e.g., *Lygeum*), in the overwhelming majority the inflorescence meristem produces lateral branches that either end immediately in

spikelets (e.g., *Lolium, Brachyelytrum*), or that go on to branch again. In addition, the fate of the inflorescence meristem is often different from that of lateral meristems (see examples in a later section).

Inflorescences develop both from terminal meristems (terminal inflorescences) and from meristems in the axils of vegetative leaves (axillary inflorescences). In most grasses, the morphology of terminal and axillary inflorescences is the same. On the other hand, in the grass tribe Andropogoneae (which includes maize, sorghum, and sugarcane), axillary inflorescences often have fewer branches than the terminal one (Kellogg, E. A., unpublished observations). This is most obvious in *Z. mays* ssp. *mays*, in which axillary branches terminate in an inflorescence without long branches (the ear), whereas the terminal inflorescence (the tassel) bears multiple long branches. In *Z. mays*, the position of the inflorescence also correlates with its sex expression, but this is unusual among grasses. *Tasselseed6* (*ts6*) and *tasselseed4* mutants of maize have more branches and more orders of branching than wild type (Irish, 1997), but the phenotype is much less severe in the ear (axillary inflorescence) than in the tassel (terminal inflorescence), also suggesting that an unknown factor suppresses branching in axillary inflorescences.

2. KNOTTED1-*like genes*

The ability to form and maintain meristems is affected by a number of genes, the best studied of which is *Z. mays KNOTTED1* (*ZmKN1*). *ZmKN1* is best known for its effects when over expressed; leaf morphology is altered, and ectopic meristems often form. However, the function of *ZmKN1* is apparently meristem formation and/or maintenance. Importantly for this review, *Zmkn1* mutants in permissive inbred backgrounds may become reproductive, but show defects in production of lateral inflorescence branches; if ears form they only produce a few kernels and tassels have fewer branches (Hake *et al.*, 2004).

ZmKN1 was the first plant protein found to contain a homeodomain (Vollbrecht *et al.*, 1991) and is named after the dominant leaf phenotype that has modified "knotted" leaf veins (Hake *et al.*, 2004). The *KNOTTED1*-like homeobox (*KNOX*) genes form a small gene family that has diversified extensively in flowering plants and falls into two monophyletic classes—class I and class II. Members of each class share intron positions, expression patterns, and characteristic residues within the homeodomain (Kerstetter *et al.*, 1994). Class I *KNOX* genes are usually expressed in the shoot apical meristem (SAM) whereas expression of Class II genes is more widespread (Hake *et al.*, 2004). Both classes contain representatives from flowering

plants, gymnosperms, ferns, and mosses, suggesting the two clades probably resulted from a gene duplication event near the origin of land plants.

ZmKN1 is a class I *KNOX* gene. Our phylogenetic analysis of 65 class I *KNOX* genes (Fig. 3) is largely congruent with analyses by Harrison *et al.* (2005). Differences between ours and other analyses, such as Hake *et al.* (2004) and Reiser *et al.* (2000), probably reflect differences in sampling. Our phylogeny finds two major clades, *BP/KN1* and *KNAT2/LG3*, each of which must have originated early in angiosperm evolution. Within each of these clades are multiple duplications within monocots, leading to at least six different class I *KNOX* gene lineages. Of these, *KN1* forms a well-supported clade with *GN1/RS1, KNOX3*, and *KNOX8* sister to a clade of eudicots containing *Arabidopsis BREVIPEDICELLUS* (*AtBP*). The *KN1* clade contains diverse grass sequences including *Hordeum vulgare KNOX3* (*HvKNOX3*), *O. sativa HOMEOBOX1* (*OsH1*), and *ZmKN1*. The *KNOX3* clade comprises *Z. mays KNOX3* (*ZmKNOX3*) and *O. sativa HOMEOBOX3* (*OsH3*), the *KNOX8* clade, *Z. mays KNOX8* (*ZmKNOX8*) and *O. sativa HOMEOBOX43* (*OsH43*), and the *GN1/RS1* clade, *Z. mays GNARLY1* (*ZmGN1*), *Z. mays ROUGHSHEATH1* (*ZmRS1*), *O. sativa HOMEOBOX15* (*OsH15*), and *H. vulgare* "*HvKn1*." Sister to the *BP/KN1* clade is the *KNAT2/LG3* clade. *Z. mays LIGULELESS4* (*ZmLG4*) and *O. sativa HOMEOBOX71* (*OsH71*) are sisters as are *Z. mays LIGULELESS3* (*ZmLG3*) and *O. sativa HOMEOBOX6* (*OsH6*). The *LG3* and *LG4* clades themselves are well supported as sisters and together are sister to a clade of asterids containing the tomato *SeTKN4* gene.

Limited sampling within most of the grass clades makes the exact placement of gene duplication events unclear. However, all clades contain orthologous rice and maize sequences, and the *KN1* clade contains representatives of most grass subfamilies including the early diverging *Pharus* (Pharoideae), indicating that all lineages are at least as old as the grass family, and may be older. The *Arabidopsis* class I *KNOX* gene *BREVIPEDICELLUS* (*BP*) is the closest eudicot relative of *ZmKN1* and is co-orthologous to the monocot *KN1, KNOX3, KNOX8*, and *GN1/RS1* clade.

The *KN1, KNOX3, KNOX8*, and *GN1/RS1* clades have different expression patterns within vegetative tissues but overlapping patterns of expression in the inflorescence and floral meristems (Foster *et al.*, 1999; Hake *et al.*, 2004; Jackson *et al.*, 1994; Sato *et al.*, 1998; Sentoku *et al.*, 1999). *KN1* and *KNOX8* (which are not sister genes) are both expressed throughout maize and rice shoot meristems, but are downregulated as leaves are initiated. Likewise, maize and rice *GN1/RS1* and *KNOX3* gene expression is similar, even though the genes do not form a clade; expression is restricted to meristem and stem regions between successive leaf primordia (Foster *et al.*, 1999; Jackson *et al.*, 1994; Sato *et al.*, 1998; Sentoku *et al.*, 1999).

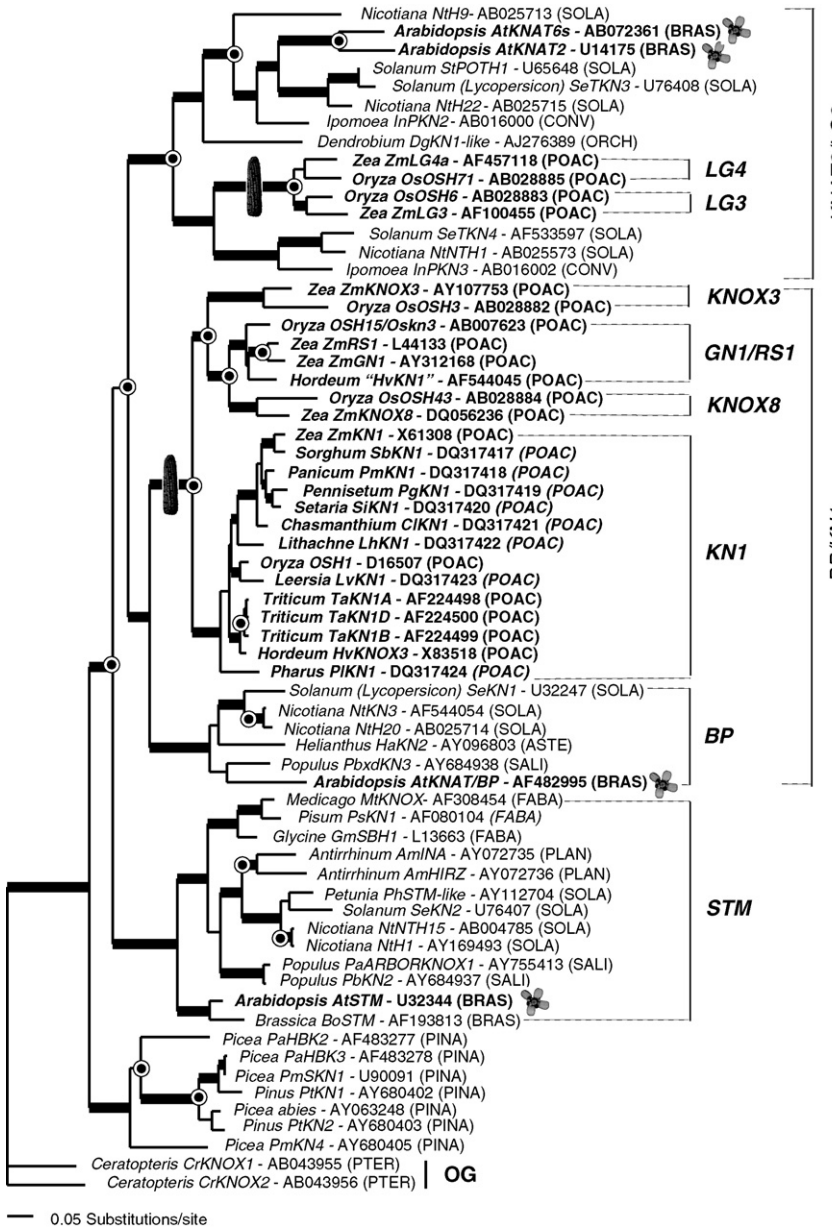

Fig. 3. Maximum likelihood phylogram of the 65 *KNOTTED1-LIKE* HOMEOBOX (*KNOX*) gene data set, comprising 555 bp, visited by Bayesian Markov Chain Monte Carlo search using the General Time Reversible model with some invariant sites and gamma distributed rates (GTR + I + G). –ln = 13505.68. Bold branches are supported by posterior probabilities >0.95. Boldface, grass and *Arabidopsis* sequences. ⊙, Inferred duplication event; *Arabidopsis* flower, *Arabidopsis* sequence; corn

Like the expression studies, genetic studies find similar *KNOX* gene function among distantly related genes. Strong loss-of-function *Zmkn1* mutants have similar phenotypes to *Arabidopsis shoot meristemless* (*stm*) mutants, which produce cotyledons but no further elements of the shoot system (Kerstetter *et al.*, 1997; Long *et al.*, 1996). These phenotypes suggest a role in shoot apical meristem maintenance and/or initiation (Hake *et al.*, 2004). Distantly related maize and *Arabidopsis* genes also co-ordinate cell differentiation and internode growth (Hake *et al.*, 2004). Knockout mutants of the rice *GN1/RS1* gene *Ososh15* are shorter than wild-type plants due to reduced lower internode length (Sato *et al.*, 1998). This dwarf phenotype is similar to *Arabidopsis brevipedicellus* (*bp*) mutants that have shortened internodes and shorter down-pointing pedicels (Venglat *et al.*, 2002).

Similar loss-of-function phenotypes in the distantly related pairs *Zmkn1*/*Atstm* and *Atbp*/*Ososh15* indicate convergence in gene function in at least one of the pairs. If the *KN1/STM*1 function is ancestral, then the *GN1/RS1* and *BP* functions are convergent. Alternatively, if the *OSH15/BP* role is ancestral then the *KN1/STM* function originated independently. It is also possible, although considerably less parsimonious, that the ancestral protein was multifunctional, and that particular subfunctions have been lost repeatedly in evolutionary time. Functional information on other Class I *KNOX* genes would help to assess the ancestral pattern.

In summary, class I *KNOX* genes are generally involved in meristem maintenance and internode growth in grasses and eudicots. However, duplicate genes have diverged in function, and distantly related genes have converged.

3. LAX PANICLE1/BARREN STALK1

LAX PANICLE1 (*LAX*) in rice and *BARREN STALK1* (*BA1*) in maize affect lateral branching in the inflorescence, although neither affects growth of the apical meristem (Gallavotti *et al.*, 2004; Komatsu *et al.*, 2003a). *BA1* is required for initiation of axillary meristems throughout the aerial parts of the plant, both vegetative and reproductive. *LAX*, on the other hand, affects only the inflorescence. Plants carrying strong mutant *lax* alleles produce no spikelets and few panicle branches; those with weaker alleles produce primary branches, but these are almost devoid of spikelets except at their termini. The gene *SMALL PANICLE* (*SPA*) has a similar developmental role as *LAX*,

ear, grass clade. Plant families abbreviated to first four letters of name. *BP, BREVIPEDICELLUS; GN1, GNARLY1; KN1, KNOTTED1; KNAT2, KNOTTED1-LIKE HOMEOBOX PROTEIN FROM Arabidopsis thaliana 2; KNOX3, KNOTTED1-LIKE HOMEOBOX 3; KNOX8, KNOTTED1-LIKE HOMEOBOX 8; LG3, LIGULELESS3; LG4, LIGULELESS4*; OG, Outgroup; *RS1, ROUGHSHEATH1; STM, SHOOT MERISTEMLESS.*

and double mutants have more severe phenotypes than either alone; in the double mutants, all axillary meristems—both vegetative and reproductive—are affected. *SPA* is not yet cloned; Komatsu *et al.* (2003a) speculate that it might be another bHLH protein that dimerizes with *LAX*, or might belong to another class of transcription factor that binds to *LAX*. The developmental role of *BA1* is thus broader than that of *LAX*, and may imply that maize lacks a functional equivalent of *SPA*. Unlike *KN1, BA1* has no effect on the apical meristem during either the vegetative or reproductive phases, or plant growth, but strong mutant alleles have no axillary meristems.

Both *LAX* and *BA1* encode atypical bHLH transcription factors that map to syntenic positions in the rice and maize genomes, respectively. Both are expressed in a slender line of cells just above the point of attachment of lateral meristems (i.e., in the axil of the branches). In normal maize tassel development, *BA1* is expressed just above (adaxial to) initiating long- and short-branch meristems, a pattern analogous to that of *LAX* in rice. Gallavotti *et al.* (2004) hypothesize that *BA1* may be necessary to specify a set of cells that can become an axillary meristem.

Plants contain a very large number of distinct bHLH proteins, with 118 identified from *Arabidopsis* and at least 131 in rice (Buck and Atchley, 2003). The majority of the plant proteins, including 104 of those from *Arabidopsis*, are group B bHLH proteins (Groups A, C, D, and E are not found in plants.). Fourteen of the *Arabidopsis* proteins and 6 from *Oryza* fall into 2 distinct classes, which Buck and Atchley (2003) call PbHLH5 and 6. These proteins have distinctive and uncharacterized DNA-binding domains.

In our phylogenetic analyses including a sample of 43 plant bHLH genes, *LAX, BA1*, and a sequence from *Sorghum* are nested within a well-supported clade of PbHLH6 sequences (Fig. 4), and thus provide the first functional data for members of this clade. Komatsu *et al.* (2003) infer that *LAX* (and by inference, *BA1*, which was not cloned at the time of their paper) represents a type of bHLH protein unique to the grasses because all related proteins in *Arabidopsis* are quite dissimilar outside of the conserved DNA-binding domain. The phylogenetic analysis supports this hypothesis, although sampling of other angiosperm families would provide a more stringent test.

4. MONOCULM1

In contrast to *KN1* and *BA1*, the effect of *MOC1* on the inflorescence has not been extensively characterized. The gene was cloned from rice, in which mutants produce no tillers and very few primary branches in the inflorescence. The gene is expressed in axillary buds throughout their development. *OsOSH1* and *O. sativa TEOSINTE BRANCHED1* (also involved in tiller

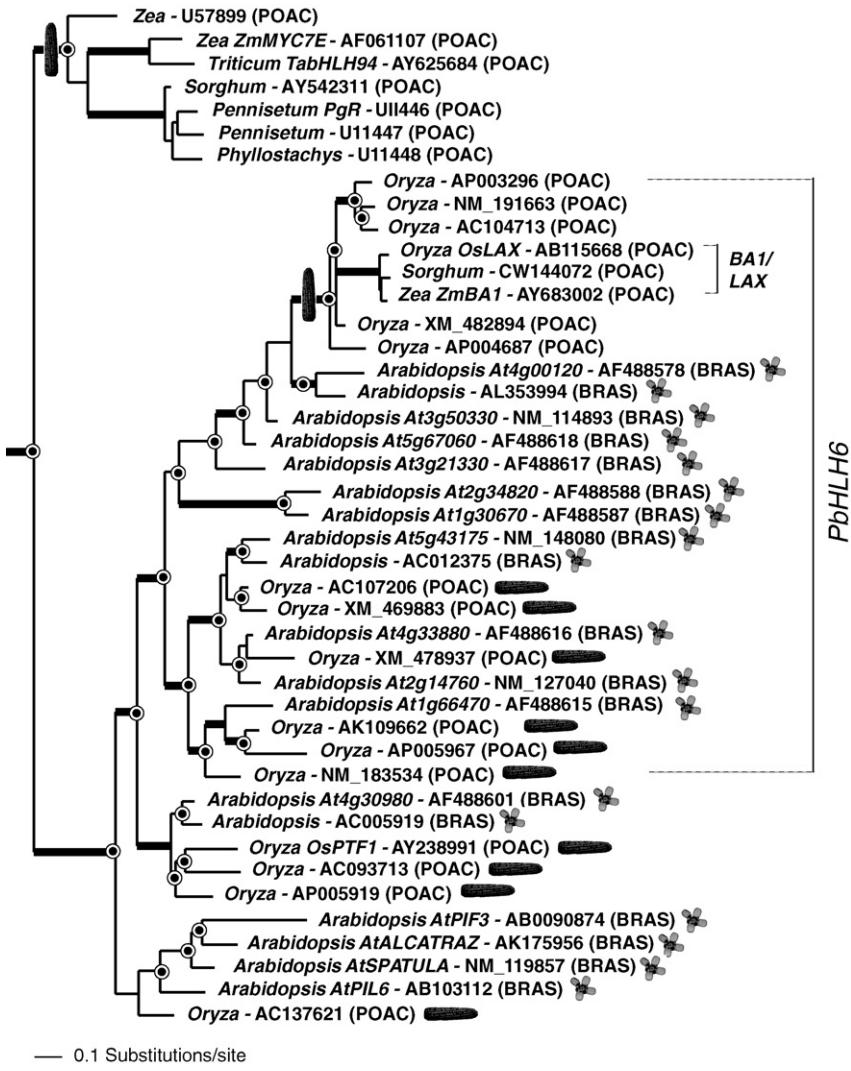

Fig. 4. Maximum likelihood phylogram of the 45 *BARREN STALK1*/*LAX PANICLE1*-like basic helix-loop-helix (bHLH) gene data set, comprising 156 bp, visited by Bayesian Markov Chain Monte Carlo search using the General Time Reversible model with some invariant sites and gamma distributed rates (GTR + I + G). Tree rooted using cat (NM_001009866) and rat (NM_022210) MAX sequences (not shown). –ln = 4803.03. Bold branches are supported by posterior probabilities >0.95. Boldface, grass and *Arabidopsis* sequences. ⊙, Inferred duplication event; *Arabidopsis* flower, *Arabidopsis* sequence; corn ear, grass clade. Plant families abbreviated to first four letters of name. *PbHLH6*, plant basic helix-loop-helix family 6, following the classification of Buck and Atchley (2003).

elongation) were not expressed in axillary buds of *moc1* mutants, indicating that *MOC1* lies upstream in the developmental pathway of bud formation.

MOC1 encodes a GRAS family protein with some similarity to the *LATERAL SUPPRESSOR* proteins of tomato and *Arabidopsis* (*LS* and *LAS*, respectively). A phylogenetic analysis of GRAS proteins (Tian *et al.*, 2004), identified a second rice gene, which they called *OsGRAS-7*, that is sister to *MOC1*; these two are sister to the clade of *LS* and *LAS*. Except for OsGRAS-7, whose function is unknown, members of this clade appear to be similar in regulating formation of axillary meristems throughout the plant.

5. BARREN INFLORESCENCE2

Maize *barren inflorescence2* (*bif2*) mutants are characterized by fewer lateral inflorescence branches, spikelets and ear shoots than wild type plants suggesting that the *ZmBIF2* gene functions in inflorescence axillary meristem initiation and maintenance (McSteen and Hake, 2001). Although *Zmbif2* mutants sometimes have fasciated apical meristems (McSteen and Hake, 2001), it is notable that *ZmBIF2* affects apical and lateral meristems differently. *Zmbif2* mutants appear superficially similar to *pinoid* (*pid*) mutants in *Arabidopsis*. *AtPID* is a regulator of polar auxin transport (Benjamins *et al.*, 2001; Christensen *et al.*, 2000), and it will be of interest to know whether *ZmBIF2* has a similar function.

To summarize this section, three of the four proteins described here affect lateral meristems differently from the apical inflorescence meristem. This sets up a prepattern that allows the fate of lateral meristems to differ from that of the apical one. Such a prepattern permits (but probably does not cause) diversification of inflorescence form, and fits with the combinatorial model of inflorescence development (Kellogg, 2000b).

B. NUMBER OF ORDERS OF BRANCHING

Grass inflorescences vary in the number of times each branch itself branches (orders of branching). The literature is inconsistent in how these orders of branching are numbered, whether the inflorescence meristem is a first-order branch and it produces second-order branches (Doust and Kellogg, 2002; Vollbrecht *et al.*, 2005), or if the numbering should begin by calling the branches produced by the inflorescence meristem primary branches, implicitly calling the inflorescence meristem branch 0 (e.g., most of the rice literature). We will use the latter convention in this chapter and count the spikelets themselves as an order of branching.

In some grasses, such as *Lolium* or *Brachyelytrum*, the inflorescence meristem produces spikelets immediately (i.e., terminating the primary branches);

in some species with this architecture a pedicel elongates beneath the spikelet late in development. In other genera, the primary branches themselves branch, producing two orders of branching beyond the inflorescence axis itself. In *Eleusine* (finger millet), the primary branches each produce multiple short secondary axes, and the secondaries terminate in spikelets. *Oryza* has three orders of branching beyond the main inflorescence axis; the lower lateral meristems (primary branches) each produce several more lateral meristems (secondary branches) before ending in a spikelet; and the secondaries produce spikelets (third-order branches) laterally before themselves terminating in spikelets (Ikeda *et al.*, 2004). Four or more orders of branching are common in many grasses, including maize; and such densely branched species as *Pennisetum* (pearl millet) and *Setaria* (foxtail millet) may have six orders of branching or more (Doust and Kellogg, 2002).

The mechanisms controlling the number of orders of branching are largely unknown. The maize mutants *tasselseed4* and *tasselseed6* (Irish, 1997) both increase the numbers of orders of branching, with the strongest effect produced by *ts4*. Unfortunately, neither is cloned yet, so cross-species comparisons are not possible. Doust *et al.* (2005) identified two major QTL for this character in a mapping population of *Setaria*, which together explained nearly half of the variance in the character. The genes underlying the QTL are unknown.

C. NUMBER OF BRANCHES AT EACH ORDER

1. Morphological variation

The number of branches at each order of branching is often fixed or varies within a narrow range for a given species. For example, *Hordeum* always produces a single primary branch at a node, and the primary branch always produces two secondary branches before itself ending in a spikelet. The secondaries also terminate in spikelets, giving the characteristic of three spikelets per node. Similarly, in the tribe Andropogoneae, the final round of branching produces a pair of spikelets, one sessile and one pedicellate, from a common meristem (the spikelet pair meristem), which can be interpreted as a short branch meristem that itself produces a single branch. The two spikelets thus each terminate branches that represent different orders of branching, evidence for which comes from developmental studies and gene expression data. In *Tripsacum*, as in *Zea*, the sessile spikelet appears to develop slightly later and lateral to the pedicellate one (Orr *et al.*, 2001), as would be expected if it were a higher order branch. In maize, *BA1* is expressed in a narrow zone between the two spikelet meristems (Section III.A.2); because *BA1* is required for lateral meristem initiation, the expression pattern suggests that one of the spikelets is lateral to the other.

The number of primary branches produced by the inflorescence meristem varies independently of the number of secondary branches produced by the primaries. For example, the inflorescence meristems of *Digitaria sanguinalis* (crab grass) and *Cynodon dactylon* (Bermuda grass) each produce fewer than 15 primary branches, but each primary branch produces numerous secondaries; the secondaries may (*Digitaria*) or may not (*Cynodon*) produce a single tertiary and then terminate in a spikelet (Barkworth *et al.*, 2003). The inflorescences of these two species thus appear sparse and spreading (digitate, or more or less antenna-like). In contrast, the inflorescence meristem of *Setaria italica* (foxtail millet) produces numerous primary branches, but few secondaries per primary, and one at each of seven or more orders of branching (Doust and Kellogg, 2002). The result is an inflorescence that is tall, narrow, and very dense.

2. RICE CENTRORADIALIS

Overexpression of *RICE CENTRORADIALIS* (*RCN*) in rice resulted in more secondary branches per primary and more tertiary branches (spikelets, in rice) per secondary (Nakagawa *et al.*, 2002). Although it is difficult to infer wild-type role based on overexpression studies, *RCN1* might regulate the number of branches produced by each higher order meristem by delaying conversion of the branch meristem to a spikelet.

RCN1 and *RCN2* in rice are distantly related to *TERMINAL FLOWER1* (*TFL1*) in *Arabidopsis*, and its ancient paralog *FLOWERING LOCUS T* (*FT*) (Fig. 5), which are key regulators of flowering time and plant architecture in *Arabidopsis*. The two genes have opposite effects, with *TFL1* repressing and *FT* activating flowering. The difference in biochemical function is due to a single amino acid residue; replacing HIS88 in *TFL1* with TYR converts *TFL1* into an activator, and replacing TYR85 in *FT* with HIS converts *FT* into a repressor (Hanzawa *et al.*, 2005).

Our phylogenetic analysis of 43 *TFL1/FT* genes estimates a complex pattern of gene duplication (Fig. 5). Given the distribution of taxa within the respective clades, the *TFL1/FT* duplication occurred before separation of the eudicot and monocot lineages (Hanzawa *et al.*, 2005). Multiple duplication events occurred within the *TFL* lineage itself. Duplications within *Populus* and *Brassica* are likely associated with polyploidization events in Salicaceae and Brassicaceae, respectively. The placement of the *Joinvillea* sequence (*JaRCN1*) as sister to grass *RCN1* genes indicates that the *Oryza RCN1/RCN2* duplication occurred within monocots and before the origin of grasses. However, determining the exact timing of the event requires sequences from other Poales and more distantly related monocots. Yet another duplication occurred during the evolution of the grasses, apparently near the

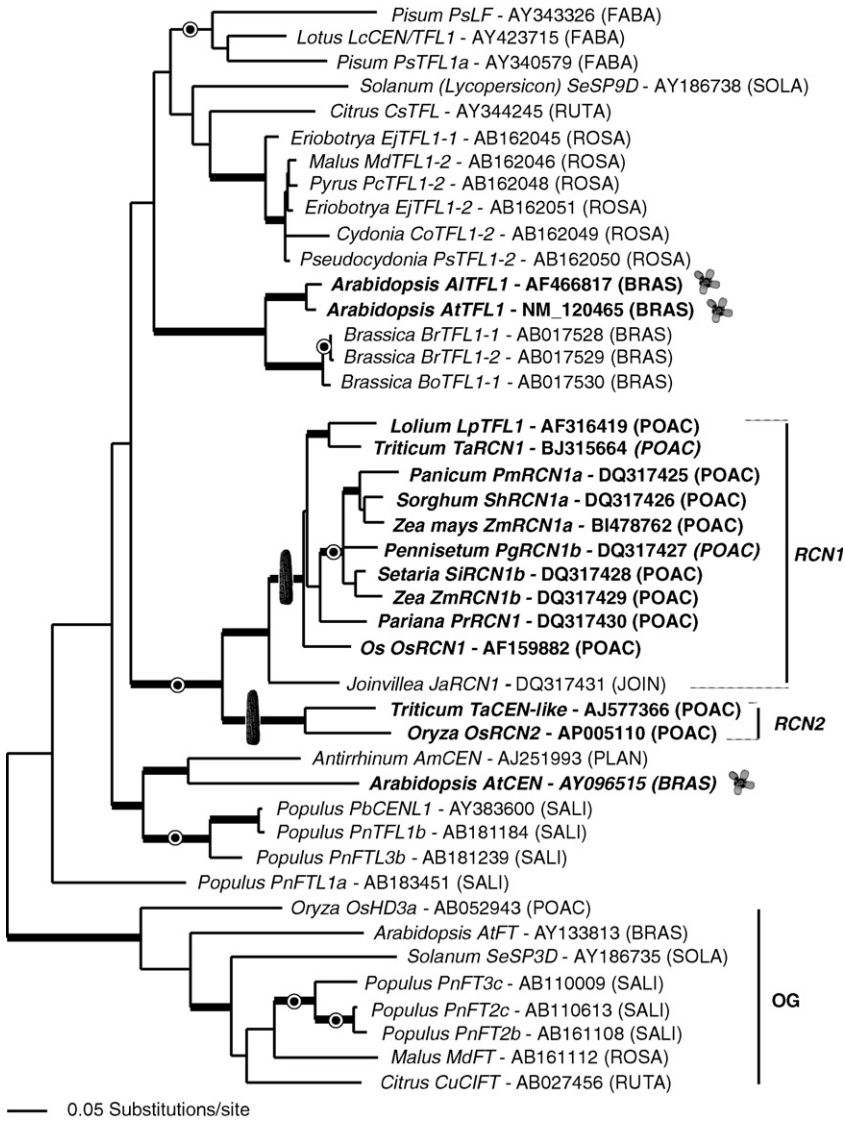

Fig. 5. Maximum likelihood phylogram of the 43 *CENTRORADIALIS, FLOWERING LOCUS T, RICE CENTRORADIALIS*, and *TERMINAL FLOWER1*-like gene data set, comprising 537 bp, visited by Bayesian Markov Chain Monte Carlo search using the General Time Reversible model with some invariant sites and gamma distributed rates (GTR + I + G). Tree rooted using *FLOWERING LOCUS T* sequences. $-\ln = 9556.78$. Bold branches are supported by posterior probabilities >0.95. Boldface = grass and *Arabidopsis* sequences. ⊙, Inferred duplication event; *Arabidopsis* flower, *Arabidopsis* sequence; corn ear, grass clade. Plant families abbreviated to first four letters of name. OG, Outgroup; *RCN1, RICE CENTRORADIALIS1; RCN2, RICE CENTRORADIALIS2.*

base of panicoids. Because of the duplications, none of the grass *TFL*-like genes can be considered directly equivalent to *Arabidopsis TFL1*.

All known *TFL1/CEN/RCN* proteins have HIS rather than TYR at a position equivalent to amino acid 88 in the *Arabidopsis* TFL1 sequence, consistent with the functional analyses of Hanzawa *et al.* (2005), with the exception of the three *Brassica* TFL1-like proteins. In *Bo*TFL1 and *Br*TFL1, HIS is replaced by ARG, resulting from a G to A substitution at the second codon position. Both ARG and HIS are positively charged, with ARG being more strongly charged than HIS, whereas TYR is polar but neutral. Overexpression or knockouts of *BrTFL1* and *BoTFL1* genes will determine whether this amino acid substitution has any functional significance.

Functional analyses of *TFL*-like genes in grasses are available for rice (*O. sativa*) and ryegrass (*Lolium perenne*) (Jensen *et al.*, 2001; Nakagawa *et al.*, 2002). Overexpression of *OsRCN1, OsRCN2*, and *LpTFL1* in *Arabidopsis* caused delayed flowering and extensive branching (Jensen *et al.*, 2001; Nakagawa *et al.*, 2002), similar to overexpressed *AtTFL1* phenotypes (Ratcliffe *et al.*, 1998). Thus, the rice and ryegrass proteins might maintain biochemical functions similar to those of their eudicot co-ortholog, *Arabidopsis* TFL1.

3. RAMOSA1

In the maize tassel, branching is somewhat more complex than in rice. The first primary branches produced by the inflorescence meristem, long branches, themselves produce a large number of secondaries; in the maize literature the long branches are called indeterminate because the precise number of secondary branches is not fixed (Fig. 1). Later, the branches produced by the inflorescence meristem produce only one secondary branch, which produces one tertiary, together forming a spikelet pair (called determinate in the maize literature). The long branches also go on to produce spikelet pairs, which represent tertiary and quaternary branches. The developmental determination of long branches with many secondaries, vs short branches with only one, is controlled by *RAMOSA1* (Vollbrecht *et al.*, 2005). When *RA1* is mutated, more branches in the tassel develop as long branches, with large numbers of secondaries. Above these are long branches that produce a mix of branched and unbranched secondaries, such that some spikelets are paired and others are unpaired. These mixed branches (“spikelet multimers”) are shorter toward the apex of the tassel, until conventional short branches are produced at the tip, giving a distinctive Christmas tree architecture. In maize, short branch (spikelet pair) production on the main inflorescence axis and on the long branches coincides with the onset of *RA1* expression, and in *ra1* mutants, more long branches are produced at the expense of short branches.

RA1 appears to have a similar function in other Andropogoneae. The genus *Miscanthus* produces long primary branches, each of which produces many short branches roughly simultaneously. Consistent with this morphology, *RA1* expression appears somewhat later in development than in maize, being delayed until all primary branches have initiated. The gene is then expressed for a very short time, corresponding to near-simultaneous short branch initiation. Unlike *Zea* and *Miscanthus, Sorghum* produces tertiary and quaternary branches before producing the final short branch meristem. Short branch production occurs on distal branches while proximal branches are themselves producing higher order branches. Consistent with this extended period of short-branch production, *RA1* in sorghum is expressed over a longer portion of inflorescence development than in the other species.

RA1 encodes an EPF-class zinc finger protein, with a DNA-binding domain that differs at one residue from the many other EPF class proteins in rice and *Arabidopsis*. The phylogeny of *RA1* and other EPF-class genes has yet to be explored. The short branches that produce spikelet pairs are only known in the panicoid grasses and are most common in Andropogoneae (Kellogg, 2000c). This observation, and the lack of an obvious *RA1*-like gene in rice, suggests that *RA1* could have originated from a duplication near the base of the Panicoideae.

4. ABERRANT PANICLE ORGANIZATION

ABERRANT PANICLE ORGANIZATION (*APO1*) in rice (Ikeda *et al.*, 2005) controls the conversion of inflorescence and branch meristems to spikelets. In *apo1* mutants, the number of primary branches is reduced, as are numbers of lateral structures (secondary branches or spikelets) on primary branches. The inflorescence meristem, which normally aborts in rice, is converted to a spikelet in many mutant plants. *APO1* has other effects on phyllotaxis (see Section III.D.3). The gene is not yet cloned, so comparative data are unavailable.

5. GRAIN NUMBER 1A/CKX2

A QTL for yield, called *grain number 1A* (*Gn1A*), has been cloned in rice (Ashikari *et al.*, 2005). Differing alleles at this locus affect the numbers of grains produced by controlling numbers of spikelets. One *indica* variety, Habataki, produces more spikelets (and hence has higher yield), but is shorter than a japonica variety Koshihikari; a mapping population derived from a cross between the two identified several loci controlling this phenotype, including *Gn1A*. *Gn1A* was identified as a cytokinin oxidase/dehydrogenase (*OsCKX2*) that is expressed preferentially in inflorescences. Higher levels of gene expression were observed in Koshihikari, and correlated with lower levels of cytokinin, whereas the opposite was true in Habataki.

The authors thus hypothesize that, by regulating the level of cytokinin in the inflorescence meristem, OsCKX2 controls the number of spikelets and hence the number of grains.

In summary, the number of branches produced at each order is fixed or varies within a narrow range for many species of grasses. The plant controls the number of branches in part by mechanisms that determine if and when the apical meristem of the inflorescence or inflorescence branch is converted to a spikelet; the timing of this conversion is affected by proteins such as RCN1, RCN2, APO1, and CKX2. The sharp distinction between long and short branches (spikelet pairs) may be synapomorphic (uniquely derived) for all or part of the Panicoideae; RA1 appears to be a candidate for control of this aspect of morphology.

D. PHYLLOTAXIS

1. Morphological variation

As the shoot apical meristem converts from a vegetative to an inflorescence meristem it elongates and begins producing primary lateral branches. In many grasses, the apical meristem also changes from distichous vegetative phyllotaxis to spiral inflorescence phyllotaxis. The shift from distichous to spiral is known to occur in Ehrhartoideae [*Oryza* (Ikeda *et al.*, 2005; Itoh *et al.*, 2005), *Zizania* (Weir and Dale, 1960)], Panicoideae [*Zea* (Bonnett, 1948; Sundberg and Orr, 1996), *Sorghum* (L. G. Le Roux and E. A. Kellogg, unpublished), *Ixophorus* (Kellogg *et al.*, 2004), *Panicum* (Bess *et al.*, 2005; Reinheimer *et al.*, 2005), *Pennisetum, Setaria* (Doust and Kellogg, 2002)], and Chloridoideae [*Eragrostis, Eleusine* (Moncur, 1981)]. Spiral inflorescence phyllotaxis is also clear in *Streptochaeta* (Anomochlooideae), which is sister to all other grasses, and in the grass outgroups *Joinvillea* (Joinvilleaceae; Malcomber and Kellogg, unpublished) and *Ecdeiocolea* (Ecdeiocoleaceae; Rudall *et al.*, 2005). Portions of the inflorescence of Restionaceae are also spirally arranged (Ronse Decraene *et al.*, 2002), but the early development of the inflorescence as a whole has not been studied in detail. Spiral inflorescence phyllotaxis is thus very probably ancestral in the grasses.

Some grasses do not undergo phyllotactic change. In all Pooideae [e.g., *Hordeum* (Klaus, 1966; Bossinger, 1990), *Avena* (Moncur, 1981), *Stipa* (Kellogg, E. A., unpublished), *Phaenosperma* (Kellogg, E. A., unpublished), and multiple other species (Evans, 1940)], the inflorescence is distichous, like the leaves. The failure to shift phyllotaxis also occurs in some Panicoideae [e.g. *Urochloa* (Reinheimer *et al.*, 2005); *Heteropogon* (Le Roux and Kellogg, 1999)], and in the woody bamboo *Fargesia* (Kellogg, E. A., unpublished). In all cases, producing distichous primary branches appears to be evolutionarily derived.

2. CLAVATA-*like*

The *CLAVATA*-like proteins [FASCIATED EAR2 and THICK TASSEL DWARF1 in maize (Bommert *et al.*, 2005a; Taguchi-Shiobara *et al.*, 2001) and FLORAL ORGAN NUMBER1 in rice (Suzaki *et al.*, 2004)] affect phyllotaxis by regulating the size of the inflorescence meristem. Maize *fasciated ear2* (*Zmfea2*) mutants are characterized by enlarged inflorescence and floral meristems. *ZmFEA2* encodes a membrane-localized leucine-rich repeat (LRR) receptor-like protein that is very similar to *Arabidopsis* CLAVATA2 (AtCLV2) (Taguchi-Shiobara *et al.*, 2001). Our phylogenetic analysis of 28 *FEA2*-like LRR genes indicates that *ZmFEA2* and *AtCLV2* are the sole remaining co-orthologs (Fig. 6). The *FEA2/CLV2* duplicates produced by the grass (Paterson *et al.*, 2004; Yu *et al.*, 2005), eudicot, and Brassicaceae (Bowers *et al.*, 2003) duplication events therefore must have been lost during evolutionary time. Although *FEA2/CLV2* genes have not been isolated from other eudicots, the clade appears to date back to at least the separation of the eudicot and monocot lineages.

ZmFEA2 and *AtCLV2* have similar broad-level expression patterns but nonidentical functional roles. Both genes are expressed in developing leaves and inflorescence tissues, but are downregulated in mature leaves (Jeong *et al.*, 1999; Taguchi-Shiobara *et al.*, 2001). *ZmFEA2* is also not expressed in roots (Taguchi-Shiobara *et al.*, 2001). Although both *Zmfea2* and *Atclv2* mutants have enlarged inflorescence and floral meristems, and in at least some flowers, longer pedicels and more stamens, only *Atclv2* mutants have enlarged vegetative meristems (Jeong *et al.*, 1999; Kayes and Clark, 1998; Taguchi-Shiobara *et al.*, 2001). Thus, both genes appear to limit meristem size, although the role of *ZmFEA2* seems to be restricted to inflorescence tissues.

The three *Arabidopsis* genes *AtCLV1, AtCLV2*, and *AtCLV3* form a single receptor–ligand complex with *AtCLV2* interacting with *AtCLV1*, and the *AtCLV1* + *AtCLV2* heterodimer acting as a receptor for the secreted *AtCLV3* signaling protein (Sharma *et al.*, 2003). This *CLV* signaling pathway functions interdependently with *WÜSCHEL* (*WUS*), with *WUS* promoting meristem fate and *CLV* restricting meristem size (Sharma *et al.*, 2003).

The maize *THICK TASSEL DWARF1* (*TD1*) and rice *FLORAL ORGAN NUMBER1* (*FON1*) genes form a clade that is sister to a eudicot clade containing *AtCLV1* (Bommert *et al.*, 2005a). All three mutants have enlarged meristems, although which meristems are enlarged depends upon the species. As in *Atclv2, Atclv1* mutants have enlarged vegetative, inflorescence, and floral meristems (Clark *et al.*, 1997). In contrast, only the inflorescence, spikelet and floral meristems are enlarged in *Zmtd1* mutants, and only floral meristems are affected in *Osfon1* mutants (Bommert *et al.*, 2005; Suzaki *et al.*, 2004). Expression patterns also vary among the three species.

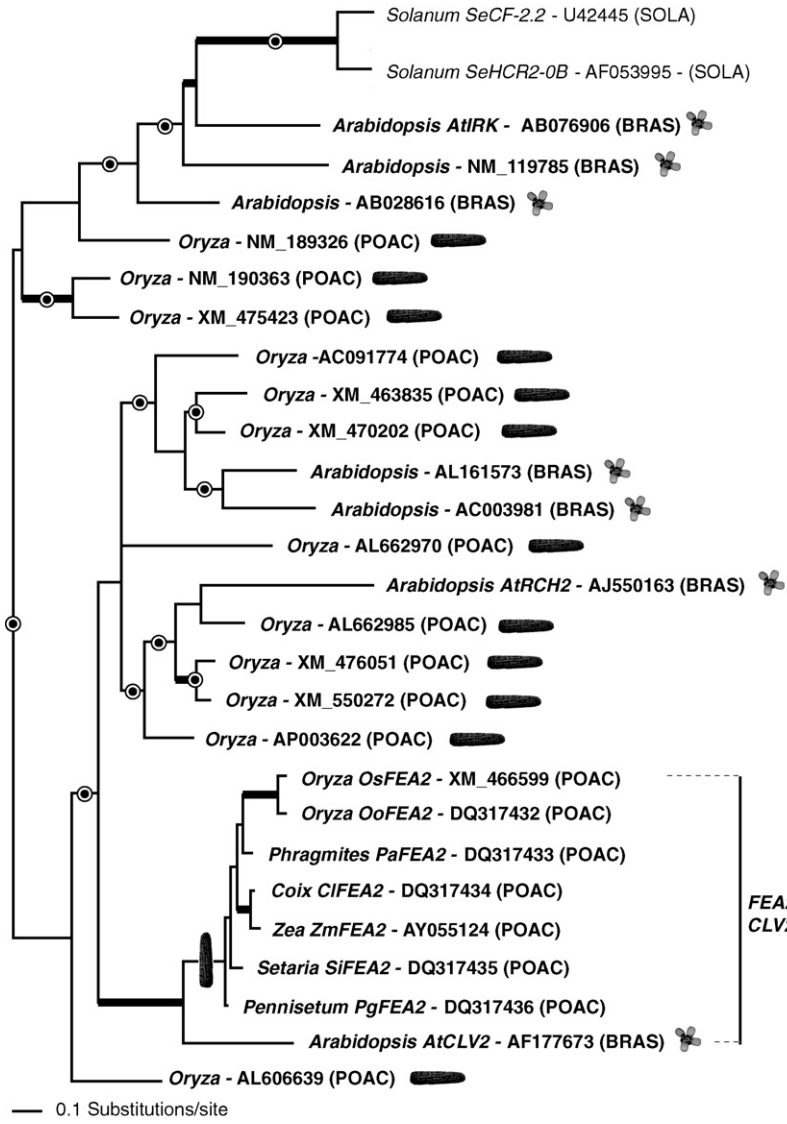

Fig. 6. Maximum likelihood phylogram of the 28 *FASCIATED EAR2* and *CLAVATA2*-like gene data set, comprising 375 bp, visited by Bayesian Markov Chain Monte Carlo search using the General Time Reversible model with some invariant sites and gamma distributed rates (GTR + I + G). Tree rooted at midpoint. $-\ln = 7072.01$. Bold branches are supported by posterior probabilities >0.95. Boldface = grass and *Arabidopsis* sequences. ◉, Inferred duplication event; *Arabidopsis* flower, *Arabidopsis* sequence; corn ear, grass clade. Plant families abbreviated to first four letters of name. *CLV2, CLAVATA2; FEA2, FASCIATED EAR2.*

AtCLV1 is restricted to inflorescence and floral meristems (Clark *et al.*, 1997), whereas *ZmTD1* and *OsFON1* are expressed broadly in all above ground tissues, including floral organs such as glumes, lemmas, and stamens (Bommert *et al.*, 2005a; Suzaki *et al.*, 2004). This shows that the different types of meristems in grasses are distinct in their development and regulation, a necessary condition for diversification of form.

The phenotypic similarity of *Zmtd1* and *Zmfea2* mutants suggests that the genes might belong to a single signaling pathway, like *AtCLV1* and *AtCLV2* (Bommert *et al.*, 2005). However, *Zmtd1/Zmfea2* double mutants exhibit more severe aberrations than either of the single mutants, with twice as many kernel rows and fewer leaves. If the two were simply components of a single signaling pathway, then the double mutant should be similar to either single mutant. Thus, unlike *Arabidopsis*, additional factors must be involved in the *TD1/FEA2* pathway. Despite broad similarity in *CLV*-like regulation of meristem size in grasses and *Arabidopsis*, there are significant differences in how such regulation is accomplished in the different species (see also Lunde and Hake, 2006).

3. ABPHYL1

Mutations in *ABPHYL1* in maize create plants with decussate phyllotaxis throughout the plant, including the first long branches of the tassel (Jackson and Hake, 1999). *Abphyl1* mutants have larger meristems than normal plants, which provide sufficient space for formation of two opposite leaves, rather than one, as is normal. The gene is a cytokinin-inducible response regulator related to the family of *Arabidopsis* response regulators (ARR, To *et al.*, 2004). A phylogenetic analysis of *ARR*-like genes in *Arabidopsis*, maize and rice indicates that *ABPHYL1* (labeled as *ZmRR3*) is sister to the rice gene BAC15873 within the type-A response regulator clade (To *et al.*, 2004). The grass *ABPHYL1* clade, in turn, is sister to a clade containing the maize genes *ZmRR1* and *ZmRR2*, indicating that the lineage is likely restricted to monocots and has duplicated at least once since the monocot/eudicot divergence. Distichous vegetative phyllotaxis is a synapomorphy for all graminoid Poales (Stevens, 2001 onwards), and conceivably *ABPHYL1* was involved in imposing this pattern, which then became fixed throughout the clade.

No grasses are known to have opposite leaves or inflorescence branches, although in some species with very dense inflorescences decussate phyllotaxis might be difficult to detect if it were present. No grasses have a phenotype similar to the *abphyl1* mutant inflorescence suggesting that this gene has a conserved function throughout the family. However, it remains possible that subtle changes in gene structure or regulation could have modest effects on meristem size and inflorescence phyllotaxis without creating truly opposite leaves.

4. ABERRANT PANICLE ORGANIZATION

Mutants of *APO1* in rice (Ikeda *et al.*, 2005; see also Section III.C) exhibit unusual phyllotaxis of the inflorescence axis. While this axis should be spiral in rice, in *apo1* mutants the phyllotaxis is distichous or biased distichous (two-ranked, but with a divergence angle other than 180°). Aberrant phyllotaxis correlates with modification of meristem shape, which is taller and narrower in the mutants. Some mutant alleles have a similar arrangement of branches as some pooids, raising the possibility that variation in the structure or regulation of this gene correlates with diversification in inflorescence form in the grasses. It will be of considerable interest to investigate this gene in other species once it is cloned.

Whatever the phyllotaxis of the inflorescence meristem, most higher order meristems in most species produce lateral structures (whether branch or spikelet meristems) distichously. Some exceptions to this occur, however. The phyllotaxis of primary branches in rice is biased distichous, with a divergence angle of 144° (Ikeda *et al.*, 2004). In some species such as *Urochloa* (Reinheimer *et al.*, 2005) or *Paspalum* (Kellogg, E. A. and LeRoux, L. G., unpublished), two ranks of secondary branches are produced on the abaxial side of the primary branch. And in genera such as *Zea* or *Tripsacum*, secondary branches are produced distichously, but the tertiaries are produced on the same side of the primary branch, creating a clear dorsiventral structure (e.g., Orr and Sundberg, 2004; Orr *et al.*, 2001). The genetic basis of these patterns is unknown.

E. BRACTS AND LEAVES SUBTENDING BRANCHES

1. Morphological variation

In the grasses as in most other angiosperms, inflorescence branches are always subtended by bracts, which are more or less prominent early in development as ridges forming just before and just below the inflorescence branches. In some species (e.g., *Stipa*), the bracts persist as small flaps, but more often are undetectable in mature inflorescences.

A number of species, particularly in Bambuseae (Bambusoideae) and in Andropogoneae (Panicoideae), also develop spathes subtending parts of the inflorescence (e.g., *Fargesia* in Bambuseae, *Heteropogon, Coelorachis, Hyparrhenia* in Andropogoneae). It is not clear whether these should be interpreted as cauline leaves with secondary inflorescences (paracladia) in their axils, or as subtending bracts that have not been developmentally suppressed, or if the distinction between the two is even meaningful.

2. LEAFY

The control of bract development is not well understood in any angiosperm. The best-studied gene is the *Arabidopsis* gene *LEAFY* (*LFY*) (Weigel *et al.*, 1992).

In *lfy* mutants of *Arabidopsis*, a floral bract develops beneath each flower. Maize has two *LFY*-like genes, *ZFL1* and *ZFL2*, which are largely redundant; single mutants have no obvious phenotype (Bomblies *et al.*, 2003). The double mutant, however, disrupts inflorescence development. Tassel architecture and sex determination are greatly altered; internodes subtending the terminal inflorescence are shortened and deformed, the pattern of branching is complex, and reduction in *ZFL* copy number correlates with reduction in number of kernel rows in the ear.

The lower branches of the tassel in *zfl1 zfl2* double mutants are replaced by peculiar "tassel ears" that are subtended by leaves. The total number of leaves per plant is higher in the mutants than in wild-type, which suggests that the leaves could correspond to derepressed inflorescence bracts. The tassel ears themselves terminate leafy branches whose leaves may correspond to additional bracts, although the structure is so abnormal that direct comparison to a normal inflorescence is difficult.

LFY-like genes have also been cloned from rice (*RFL*; Kyozuka *et al.*, 1998) and from *Lolium* (*LtLFY*; Pooideae; Gocal *et al.*, 2001), and are apparently single-copy in both species. Comparison with *ZFL* is unfortunately not possible because expression has not been studied at comparable stages in the different species and *rfl* and *Ltlfy* mutants have not been reported, so gene function is unknown. *RFL* is expressed very early in the development of young inflorescence axes, but is excluded from the primordia of primary and secondary branches. The comparable stage in maize would be during formation of the long branches of the tassel, but *ZFL* expression is not reported for that stage, so it is unclear if the *RFL* pattern is novel. Because *ZFL* mutations do affect long-branch morphology, it is conceivable that the mutations in maize correlate with the expression pattern in rice, and the role of the gene in early inflorescence development is conserved. The inflorescence of *Lolium* is unbranched, so is completely uninformative about branch production in general.

The two *ZFL* loci are the product of a gene duplication event that precedes the divergence of *Tripsacum* and maize, but follows the diversification of most Andropogoneae (Bomblies and Doebley, 2005). The genes are under purifying selection throughout the Andropogoneae, and within *Z. mays* itself. Thus sequence analysis suggests that gene function is conserved throughout the tribe.

In the grasses, *LFY* has no apparent function in either short branch (spikelet pair) formation or spikelet formation. Mutations of *ZFL1* and *ZFL2* in maize do not affect formation of short branches, and glumes, lemmas, and paleas are likewise normal. *ZFL* is expressed only weakly, if at all, in glumes, and *RFL* (rice) is not expressed in spikelets. Thus *LFY*-like genes are not involved in specification of developed bracts in spikelets.

The *Arabidopsis* gene *JAGGED* is also involved in development of inflorescence bracts, and may ultimately provide some insight (Dinneny *et al.*, 2004; Ohno *et al.*, 2004). *JAGGED*-like genes have not yet been cloned from grasses, however.

F. PRESENCE OF A TERMINAL SPIKELET

Grass species vary considerably in whether the inflorescence meristem terminates in a spikelet or not. For example, in rice (Ikeda *et al.*, 2004), maize (Bonnett, 1948), foxtail millet (Doust and Kellogg, 2002), and *Urochloa* (Reinheimer *et al.*, 2005), the apical meristem of the inflorescence never produces a spikelet and may remain clearly visible until quite late in development. Conversely, in ryegrass [*Lolium*; (Gocal *et al.*, 2001)], oats (*Avena*; Hiser, K. M., and Kellogg, E. A., unpublished), and proso millet [*Panicum miliaceum*, (Bess *et al.*, 2005)], among many others, the inflorescence meristem apparently is converted to a spikelet, although it is often hard to rule out completely the possibility of a tiny residual inflorescence meristem; in this case the ''terminal'' spikelet would actually be lateral.

Arabidopsis LEAFY is sufficient to convey floral identity (Weigel and Nilsson, 1995), and might be involved in formation of terminal flowers in grasses; however, expression of *LFY*-like genes in grasses does not correlate with presence of a terminal flower. *LFY*-like genes have been cloned from *Oryza* (Kyozuka *et al.*, 1998), *Lolium* (Gocal *et al.*, 2001), and maize (Bomblies *et al.*, 2003). *LtLFY, ZFL1*, and *ZFL2* are not expressed in the inflorescence meristems of *Lolium* or maize, respectively, even though the former has a terminal spikelet and the latter does not. *RFL* is expressed in the early inflorescence meristem of rice, but expression disappears by the time the secondary branches start to form on the primaries.

Mutants of *TFL1* in *Arabidopsis* convert the inflorescence meristem to a flower (Shannon and Meeks-Wagner, 1991), and when overexpressed delay the transition to flowering (Ratcliffe *et al.*, 1998). As noted in Section III.C, *RCN1* and *RCN2* in rice, which are related to *TFL1*, exhibit similar overexpression phenotypes and are thus good candidates for involvement in this aspect of diversification. Loss-of-function mutations in the rice genes are not known, however.

Grasses also vary in whether lateral branches end in spikelets or not. Most species do have terminal spikelets on secondary and higher order branches, but in genera such as *Paspalidium* (Panicoideae) and *Dactyloctenium* (Chloridoideae), secondary branches end in a stiff point. Candidate genes controlling this aspect of plant architecture are unknown.

G. ELONGATION OF INFLORESCENCE INTERNODES

Much of the obvious variation in grass inflorescences is created by differential elongation of internodes. For example, the spreading inflorescence of *Sorghum halepense* looks quite different from the contracted one of *Sorghum bicolor*. As with patterns of branching, elongation of the main inflorescence axis is independent of that of primary branches, and secondary branches may elongate independently from primaries. Thus, inflorescences that look very similar early in development can look very different later. In addition, elongation may not occur equally in all internodes even along the same axis. Reinheimer *et al.* (2005) have shown for *Panicum maximum* (= *Megathyrsus maximus*) that the most basal internodes on the main inflorescence axis fail to elongate at maturity, producing a pseudowhorl of primary branches. Such pseudowhorls are evident in other species as well (e.g., *Panicum mertensii, Setaria verticillata*), although their development has not been studied.

In most species, elongation occurs very late in development, after all branching has occurred and spikelets are largely formed (Bess *et al.*, 2005; Doust and Kellogg, 2002; Ikeda *et al.*, 2004). It is thus developmentally separable from specification of branch and spikelet identity.

Perhaps because it is a quantitative rather than qualitative character, few genes are reported to modify internode length. *PANICLE PHYTOMER1* in rice is one such gene, although it also increases the numbers of branches (Takahashi *et al.*, 1998). Several QTL have been identified that affect primary branch density in *Setaria* (Doust *et al.*, 2005). Also, treatment of plants with gibberellins and with GA inhibitors affects the length of panicle branches in *Setaria* (Bess, Doust, Preston, and Kellogg, unpublished). We thus expect that considerable information on internode length lies buried in the literature. Of most interest for comparative morphology would be loci that affect length of internodes in inflorescences but not in vegetative parts of the plant.

H. GLUMES AND SPIKELETS

1. Morphological variation

The ultimate unit of the grass inflorescence is the spikelet, literally a little spike. Although not actually synapomorphic for the family, the spikelet characterizes all grasses except the earliest diverging lineage, Anomochlooideae, that includes *Anomochloa* and *Streptochaeta* (Grass Phylogeny Working Group, 2001). The spikelet meristem produces lateral organs in a distichous arrangement, beginning with two bracts known as glumes. Above the glumes are one or more flowers, produced laterally on a short axis, the rachilla.

Floral development is discussed in detail by Whipple and Schmidt, Chapter 10 (this volume) so we will consider here only formation of glumes and the architecture of the spikelet itself.

The spikelet first appears as a lateral or terminal meristem on an inflorescence or inflorescence branch. The first product of the spikelet meristem is a leaf-like structure, the first glume. This is formed on the abaxial side of the spikelet meristem (relative to the axis on which the spikelet is borne) in many panicoid grasses but is lateral in most other species. The second glume forms opposite and slightly above the first, in a distichous arrangement.

Glumes are generally interpreted as bracts (Clifford, 1987), which are extrafloral structures. In a conventional monocot, flowers and inflorescence branches are subtended by bracts. Each inflorescence axis, and in a few taxa each floral axis, then bears an adaxial prophyll. To extend this conventional interpretation to the glumes, then, would require interpreting the first glume as an inflorescence bract and the second as a prophyll. Arguing against this interpretation is the position of the glumes and their morphology. Both glumes are produced by the spikelet meristem and are borne on the spikelet axis, suggesting that neither is an inflorescence bract and the first one (rather than the second) might be considered a prophyll. However, monocot prophylls are usually two-keeled structures and are adaxial, whereas the first glume is rarely adaxial, and even more rarely two keeled. The morphology and position of the first glume thus suggests that it may not be homologous to the prophyll in other monocots.

2. FRUITFULL 1/2/3

The MIKC MADS-box *FRUITFULL*-like (*FUL*-like) genes (*FRUITFULL* [*FUL*], *APETALA1* [*AP1*], and *CAULIFLOWER* [*CAL*]) play important roles in *Arabidopsis* floral development (Ferrandiz *et al.*, 2000) but act only after the genes determining the transition to flowering time such as *TFL1, FT, SUPPRESSOR OF OVEREXPRESSION OF CONSTANS1* (*SOC1*) and *CONSTANS* (*CO*), have initiated the floral pathway (Blazquez, 2000). The three genes have redundant roles in specifying floral meristems (Ferrandiz *et al.*, 2000), but *AtAP1* has a nonredundant role in specifying the sepal and petal whorls of the flower (Pelaz *et al.*, 2001), and *AtFUL* has a nonoverlapping role during fruit development, preventing ectopic activity of *INDEHISCENT* to maintain proper elongation of the silique (Liljegren *et al.*, 2004).

The three *Arabidopsis FRUITFULL*-like genes fall into two clades: the core eudicot *AP1*-like clade (*AtAP1* and *AtCAL*), and the core eudicot *FUL* clade (*AtFUL*) (Litt and Irish, 2003). Litt and Irish (2003) also showed that monocot genes of the *AP1-FUL* group are much more similar to *FUL* than

to *AP1*. Their extensive *AP1-FUL* phylogeny demonstrated that orthologs of *Arabidopsis AP1* (*AtAP1*) and *CAULIFLOWER* (*AtCAL*) genes are restricted to Brassicaceae, and *AP1/CAL* genes are restricted to core eudicots. *FUL* genes form a grade, but magnoliid, monocot, noncore eudicot, and two core eudicot clades of *FUL* genes were recognized (Litt and Irish, 2003). Following their results, and contrary to much recent literature on grasses, we refer to the grass genes as *FUL*-like, rather than *AP1*-like.

Sampling within our phylogenetic analysis of 35 *FUL*-like genes concentrates on the grass *FUL* genes (Fig. 7). These fall into three clades: *FUL1* (containing *Triticum monococcum AP1, H. vulgare M5, L. temulentum MADS1, O. sativa MADS14, Z. mays MADS4, ZmMADS15*), *FUL2* (*H. vulgare M8, L. perenne MADS2, O. sativa MADS15, Z. mays AP1* and *ZmMADS3*), and *FUL3* (containing *H. vulgare M3, L. perenne MADS3, O. sativa MADS18* and *Z. mays MADS28*). *FUL1* and *FUL2* are sisters, and the *FUL1/FUL2* clade is sister to *FUL3*. All three clades result from gene duplications apparently near the base of grasses, although the exact placement of the duplication events requires additional sequences from related members of Poales.

All grass *FUL*-like genes appear to be expressed broadly throughout the plant, although most information is available for the *FUL1* and *FUL2* clades. Within spikelets, barley, ryegrass, rice, and maize *FUL1* and *2* genes are expressed in glumes, palea, and lemma of all species, but expression in other organs varies from species to species (J. Preston, unpublished data). *FUL1* and *2* thus may be involved in specifying the glumes, lemmas, and paleas, but orthologous genes in different species may have different roles in the inner three floral whorls (Pelucchi *et al.*, 2002; Schmitz *et al.*, 2000). All three *FUL* genes are expressed in floral meristems, consistent with a plesiomorphic (ancestral) role of the gene family in specifying floral meristem identity throughout angiosperms (Pelucchi *et al.*, 2002; Schmitz *et al.*, 2000). RNAi silencing of *OsMADS18* (a *FUL3* gene) did not produce an obvious phenotype (Fornara *et al.*, 2004), suggesting that *FUL3* may be at least partially redundant with *FUL1* and *FUL2* in rice.

Grass FUL proteins play a role in determining flowering time, unlike the *Arabidopsis* FUL-like proteins. Grass FUL1 proteins in barley, ryegrass, oats, and wheat have all been implicated in vernalization-induced competence to flower (Danyluk *et al.*, 2003; Fu *et al.*, 2005; Jensen *et al.*, 2005; Trevaskis *et al.*, 2003; Yan *et al.*, 2003). *H. vulgare M5* (*HvM5*) and *T. monococcum AP1* (*TmAP1*) show limited expression in winter cultivars that have not been vernalized but are strongly expressed following vernalization treatment (Trevaskis *et al.*, 2003; Yan *et al.*, 2003). Upregulation of *LpMADS2* (*FUL2*) and *LpMADS3* (*FUL3*) during floral transition in perennial ryegrass also points to a general role for the other grass *FUL*

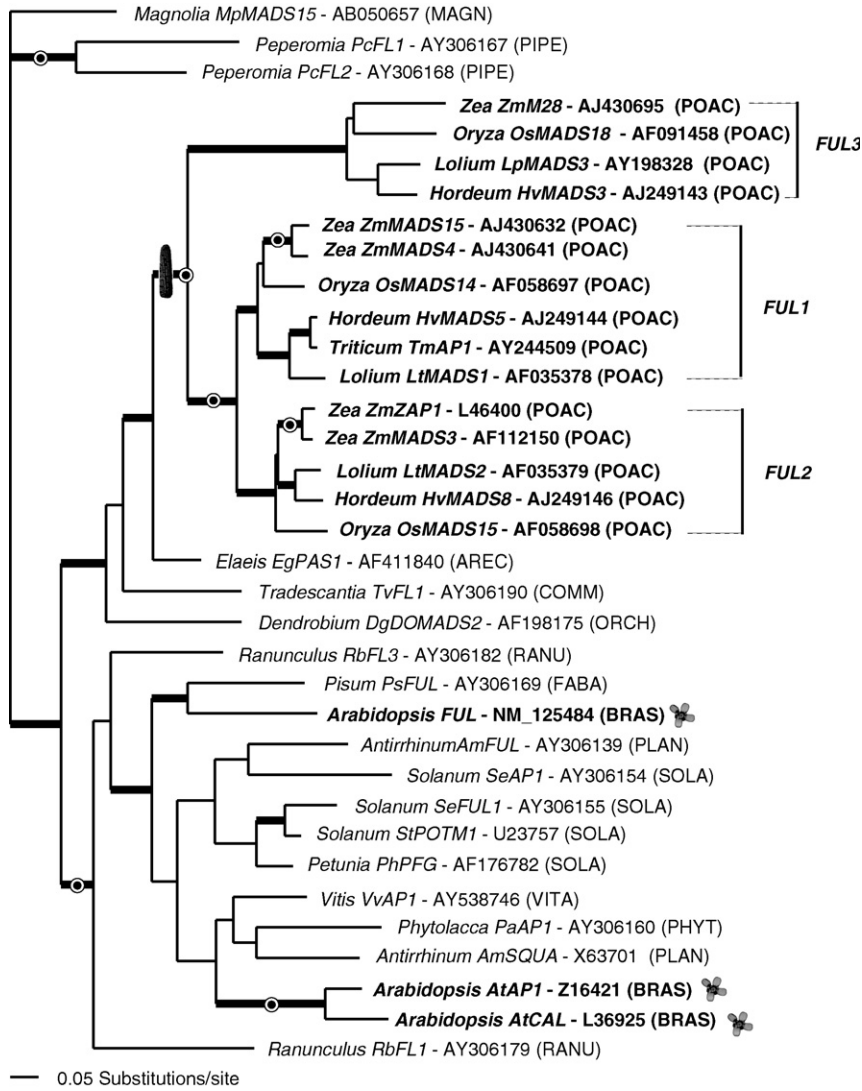

Fig. 7. Maximum likelihood phylogram of the 35 *FRUITFULL/APETALA1*-like gene data set, comprising 657 bp, visited by Bayesian Markov Chain Monte Carlo search using the General Time Reversible model with some invariant sites and gamma distributed rates (GTR + I + G). Tree rooted using *Magnolia MpMADS15* sequence. $-\ln = 11057.87$. Bold branches are supported by posterior probabilities >0.95. Boldface, grass and *Arabidopsis* sequences. ⊙, Inferred duplication event; *Arabidopsis* flower, *Arabidopsis* sequence; corn ear, grass clade. Plant families abbreviated to first four letters of name. *FUL1, FRUITFULL1; FUL2, FRUITFULL2; FUL3, FRUITFULL3.*

genes in flowering time (Andersen *et al.*, 2004). All species investigated to date are in subfamily Pooideae, so it will be interesting to know if the same proteins are involved in vernalization of species from other subfamilies.

These data suggest general roles for at least *FUL1* and *FUL2* genes, and conceivably *FUL3*, in specifying glumes, palea, lemma, and floral meristems, and also novel roles in the flowering time pathway in certain species following gene duplications near the origin of the grass family.

AP1-like genes in eudicots appear to be regulated by *LFY*, but this regulatory interaction is either reversed or lost in the grasses (see also Section III.E.1). The rice gene *RFL* is expressed in very young spikelets, but this expression is soon lost (Kyozuka *et al.*, 1998). *LtLFY* from *Lolium* and both *ZFL* genes are expressed in spikelet meristems, but expression in glumes is transitory (Bomblies *et al.*, 2003; Gocal *et al.*, 2001). *LtLFY* (in *Lolium*) is expressed notably later than *FUL*-like genes. In *zfl1 zfl2* double mutants, glumes, lemmas, and paleas—organs with conserved *FUL* expression—are unaffected. Together, these data imply that grass *LFY* orthologs are not involved in specifying spikelet meristem, glume, lemma, or palea identity. In addition, they do not regulate the *FUL*-like genes, because they are expressed after the *FUL*-like genes and are not expressed in the same organs; the lack of mutant phenotype in maize glumes also indicates that *LFY* does not function there.

3. *BRANCHED SILKLESS1/FRIZZY PANICLE1*

In most grasses, no meristem forms in the axil of either glume. Extensively branched bracteate structures do occur, however, in some woody bamboos (Judziewicz *et al.*, 1999). Because it is difficult to determine if highly branched units are at all homologous to conventional spikelets, they are generally called pseudospikelets; it is not known whether the bracts in the pseudospikelet are homologous to glumes. Occasional terata are also reported in which an extra spikelet forms in the axil of the glume (Sharman, 1947).

The BRANCHED SILKLESS1 (BD1) protein in maize, and its ortholog FRIZZY PANICLE1 (FZP) in rice control the outgrowth of the rachilla and production of flowers (Chuck *et al.*, 2002; Komatsu *et al.*, 2003). When the gene is mutated in maize or rice, glume-like structures are produced, but the "spikelet" meristem continues to produce bracts and to branch.

Both genes belong to the ethylene response element-binding factor (ERF) class of transcription factors that are involved in diverse developmental processes including ethylene-mediated responses to pathogens, cold and abiotic factors (Chuck *et al.*, 2002). Our phylogenetic analysis of 36 *BD1/FZP* genes using only the conserved ERF domain did not have enough

nucleotide variation to provide a well-supported estimate of relationships (not shown). However, preliminary results from these analyses indicate that the *Arabidopsis LEAFY PETIOLE* gene is likely the closest relative of *BD1/FZP*.

BD1/FZP genes from seven diverse grasses have conserved amino acid sequences in the ERF domain and are 45–75% identical in other regions of the protein (Chuck *et al.*, 2002). Based on this sequence similarity and the similar mutant phenotypes in rice and maize, *BD1/FZP* genes are hypothesized to have similar roles in most BEP and PACCAD grasses. The exception appears to be the maize *BD1* duplicate that is conserved in the ERF domain, but divergent elsewhere, pointing to either a different or loss of gene function (Chuck *et al.*, 2002). Given the similarity between maize *bd1* mutant spikelets and bamboo pseudospikelets, it would be of considerable interest to isolate *BD1/FZP* genes from a bamboo.

I. GLUME VS LEMMA IDENTITY

1. Morphological variation

After it produces two glumes, the spikelet meristem continues to produce lateral structures. These may be morphologically similar to or distinct from the glumes and may bear an axillary floral meristem or not. Conventionally, any lateral bract-like structure above the second glume is considered to be a lemma, which is generally interpreted as a floral bract. The lemma is then described as sterile (if it has no axillary meristem) or fertile (if it does have an axillary meristem). For many grasses (e.g., *Ehrharta, Chasmanthium*), the sterile lemma is morphologically similar to the glumes, implying that there is a developmental gradient from the glumes through the sterile lemmas to the fertile ones, whereby floret identity is acquired gradually during development, rather than abruptly.

Evidence of such a gradient comes from overexpression studies of the *SEPALLATA* gene *LEAFY HULL STERILE1* in rice (Prasad *et al.*, 2001). In rice, the glumes are reduced to tiny flaps, called rudimentary glumes in the literature (Bommert *et al.*, 2005b). Above them are two larger subulate structures in the position of sterile lemmas, and morphologically distinct from either the true glumes or the fertile lemma; these are generally called "empty glumes." When *LHS1* is overexpressed, the sterile lemmas enlarge and become morphologically similar to the fertile lemma, whereas the true glumes are unaffected (Prasad *et al.*, 2001). These results, along with supporting data from *OsFZP* (Komatsu *et al.*, 2003b), support the hypothesis first proposed by Stapf (1917) that rice has a three-flowered spikelet with one fertile upper floret and two lower florets that are reduced to sterile lemmas.

2. SEPALLATA *genes*

SEPALLATA (*SEP*) genes play fundamental roles in the development of all floral whorls in *Arabidopsis* where they are hypothesized to act as cofactors with A-, B-, C-, and D-class MADS-box floral homeotic genes (Ditta *et al.*, 2004; Pelaz *et al.*, 2000). *Arabidopsis* has four *SEP* genes, *AtSEP1, AtSEP2, AtSEP3*, and *AtSEP4*. Expression of *AtSEP1, AtSEP2*, and *AtSEP3* is restricted to the flower, whereas *AtSEP4* is expressed in all above-ground organs (Huang *et al.*, 1995; Ma *et al.*, 1991; Savidge *et al.*, 1995). The four *Arabidopsis SEP* genes are developmentally redundant, with discernible phenotypes only when several genes have been removed. For example, *Arabidopsis sep1 sep2 sep3* (*sepallata*) triple mutant flowers are composed entirely of sepal-like structures, and *sep1 sep2 sep3 sep4* quadruple mutants have flowers composed entirely of leaf-like structures (Ditta *et al.*, 2004; Pelaz *et al.*, 2000).

The phylogenetic analysis of *SEPALLATA* genes by Zahn *et al.* (2005) includes 113 sequences that span flowering plant diversity. An early gene duplication event occurred before the origin of extant angiosperms and produced the *SEP3* and *LOFSEP* clades, containing *Arabidopsis AtSEP3*, and rice *LEAFY HULL STERILE1* (*LHS1*), *O. sativa MADS5* (*OsMADS5*) and *OsMADS34* genes, petunia *FLORAL BINDING PROTEIN9* (*PhFBP9*) and *PhFBP23* genes, and *Arabidopsis AtSEP1, AtSEP2,* and *AtSEP4* genes, respectively (Malcomber and Kellogg, 2005). Within the *LOFSEP* clade, additional duplications near the base of core eudicots produced the *SEP1/2, FBP9/23*, and *SEP4* clades, and conceivably the Brassicaceae duplication (Bowers *et al.*, 2003) produced the *AtSEP1* and *AtSEP2* genes. Additional duplications occurred near the origin of Solanaceae within the *SEP1/2* and *FBP9/23* clades. Two duplications near the base of grasses within the *LOFSEP* clade produced the *LHS1, OsMADS5*, and *OsMADS34* subclades. Polyploidy has also produced additional duplicates in monocots and eudicots throughout the phylogeny. These duplications imply that genes similar to *AtSEP3* are found throughout flowering plants, whereas *AtSEP1* and *AtSEP2* orthologs are restricted to Brassicaceae and *AtSEP4* orthologs to core eudicots (Malcomber and Kellogg, 2005).

Our phylogenetic analysis of 119 *SEP* genes comprises the Zahn *et al.* (2005) 113 *SEP* gene data set plus six additional sequences from early diverging grasses (*Streptochaeta* and *Pharus*) and additional monocots to localize putative grass-specific duplication events within the *LOFSEP* and *SEP3* clades (Fig. 8). This analysis is congruent with Zahn *et al.*'s (2005) analysis, with the exception of the placement of *Helianthus HaM137, Dendrathema DgCDM77*, and *Gerbera GhGRCD1* (ASTERACEAE SEP3, Malcomber and Kellogg, 2005). In our analysis, this clade appears to have

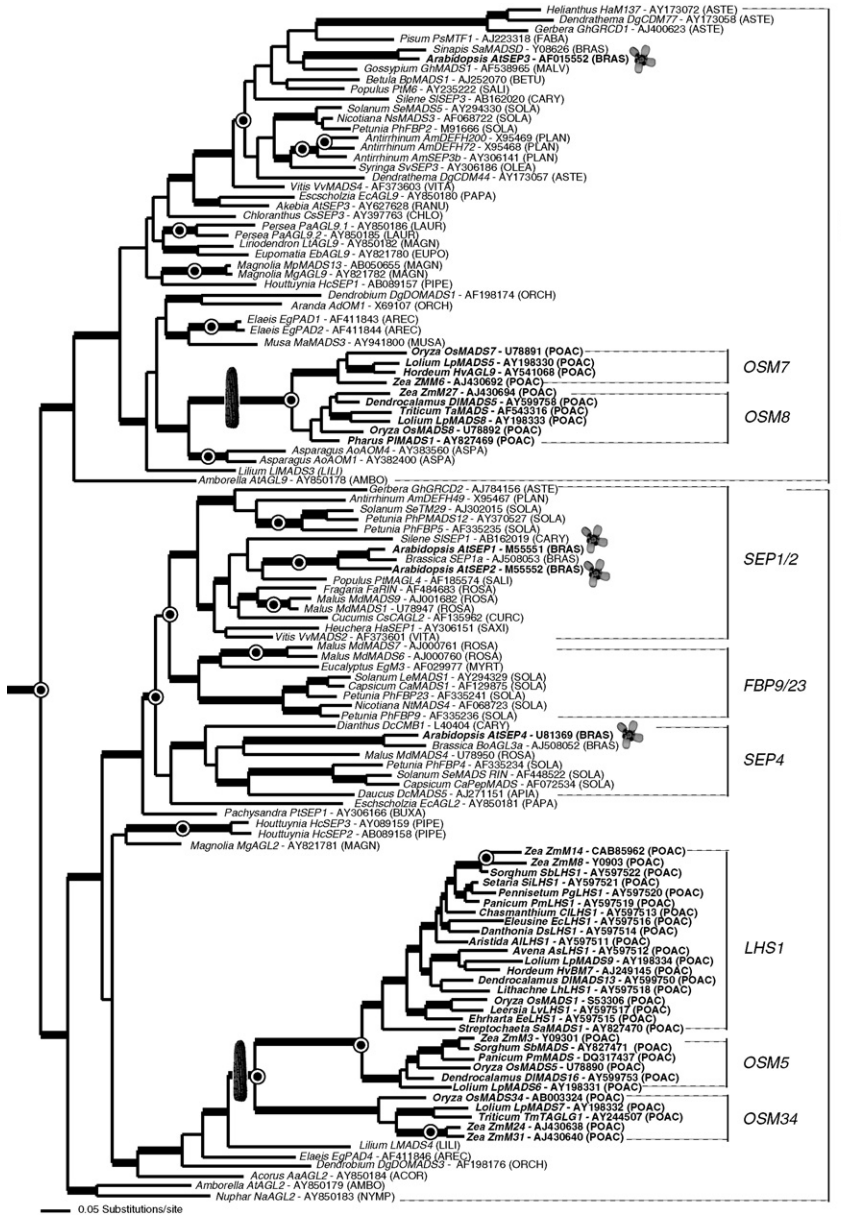

Fig. 8. Maximum likelihood phylogram of the 119 *SEPALLATA* gene data set, comprising 1200 bp, visited by Bayesian Markov Chain Monte Carlo search using the General Time Reversible model with some invariant sites and gamma distributed rates (GTR + I + G). Tree rooted using four *AP1/FUL*-like gene sequences (not shown). Analysis also included 26 *AGL6*-like gene sequences (not shown). −ln = 56818.93. Bold branches are supported by posterior probabilities >0.95. Boldface, grass and *Arabidopsis* sequences. ◉, Inferred duplication event; *Arabidopsis* flower, *Arabidopsis* sequence; corn ear, grass clade. Plant families abbreviated to first four letters of

been produced from a duplication near the base of eudicots, whereas in Zahn *et al.* (2005) the clade diverges within basal angiosperms.

Streptochaeta SaMADS1 falls within the *LHS1* clade of grass sequences, and shares C-terminal motifs with all other *LHS1* genes (Malcomber and Kellogg, 2005; Vandenbussche *et al.*, 2003). Placement of *SaMADS1* within the *LHS1* clade suggests the *LHS1/OsMADS5* duplication occurred prior to the origin of extant grasses, although the exact placement of this duplication requires additional sampling from related members of the Poales. The *Pharus PlMADS1* gene is sister to other *OsMADS8* genes, although its placement is not well supported. This position, or a placement sister to other *OsMADS7* and *OsMADS8*, is consistent with the grass *SEP3* duplication occurring near the origin of grasses. As in the *LHS1/OsMADS5* duplication, the exact placement of this duplication will require additional sequences from related members of the Poales.

SEP mRNA expression patterns are heterogeneous within flowering plants, and *SEP* gene function varies from redundant, as in *Arabidopsis*, to nonredundant with roles in fruit maturation, floral whorl identity and plant architecture (Malcomber and Kellogg, 2005). Expression patterns within the grass *LHS1* clade are particularly diverse (Malcomber and Kellogg, 2004). *LHS1* orthologs are always expressed in the palea and lemma, but expression in other organs of the floret and regions of the spikelet varies among species (Malcomber and Kellogg, 2004).

The semidominant negative rice *lhs1* mutant has leafy palea, lemma, and lodicules, fewer stamens, and occasionally an extra pistil or floret (Jeon *et al.*, 2000). Ectopic expression of this gene in rice produces plants with short panicles and irregularly positioned branches; the sterile lemmas of the two lower flowers are similar to the palea and lemma of the fertile upper flower (Prasad *et al.*, 2001). As discussed in an earlier section, these functional analyses indicate that *OsLHS1* influences palea and lemma morphology, but not glumes.

J. FLORET NUMBER

1. Morphological variation

Grasses vary considerably in the number of florets produced by the spikelet meristem. Multiflowered spikelets originated relatively late in grass evolution, just before the divergence of the Puelioideae, and reversals to single

---

name. *FBP9/23, FLORAL BINDING PROTEIN9/23; LOFSEP: LEAFY HULL STERILE1; OsMADS5, OsMADS34; FLORAL BINDING PROTEIN9/23 SEPALLATA1/2, SEPALLATA4* clade; *LHS1, LEAFY HULL STERILE1; OSM5, OsMADS5; OSM7, OsMADS7; OSM8, OsMADS8; OSM34, OsMADS34; SEP1/2, SEPALLATA1/2; SEP3, SEPALLATA3; SEP4, SEPALLATA4.*

flowered spikelets have occurred frequently (Grass Phylogeny Working Group, 2001). In some species, including all the panicoid grasses, the number of florets is fixed both within and between species, whereas in other species or clades the number varies within a range.

Spikelets may mature from the top-down (basipetal maturation) or the bottom-up (acropetal maturation). Spikelets with a fixed number of florets exhibit basipetal maturation, whereas in those with a variable number maturation is acropetal. The difference in maturation patterns also correlates with differences in expression domain of *LHS1*; in basipetal species, *LHS1* is expressed only in the uppermost floret, whereas it is expressed in multiple florets in acropetal species (Malcomber and Kellogg, 2004). The pattern in basipetal grass species supports the hypothesis that *LHS1* orthologs specify the terminal flower of the spikelet (the "selector gene" hypothesis; Cacharrón *et al.*, 1999), whereas the gene may have different or additional developmental roles in species with acropetal maturation of florets (Malcomber and Kellogg, 2004).

2. INDETERMINATE SPIKELET1

Mutations in the maize gene *INDETERMINATE SPIKELET1* (*IDS1*) lead to production of more than two flowers in the spikelet. It thus seems possible that modulation of *ids1* might result in some of the variation in floret number observed in the grasses.

IDS1 is an *APETALA2* (*AP2*)-like gene cloned and characterized in maize by Chuck *et al.* (1998). *ZmIDS1* is expressed broadly in both vegetative and floral tissues, but only appears to function in inflorescence development where it regulates the number of florets a spikelet produces (Chuck *et al.*, 1998).

Our phylogenetic analysis of 21 *IDS1*-like AP2 genes estimates a well-supported grass clade containing the *ZmIDS1* clade and two rice genes (Fig. 9). The position of these two rice genes suggests that *IDS1* and an *IDS1*-like rice sequence result from a duplication near the base of grasses, although additional *IDS1*-like sequences would help test this hypothesis. The wheat *IDS1* gene (*T. monococcum IDS1*) is the "*Q* gene" that confers the square-headed phenotype and free-threshing character of domesticated bread and durum wheat, and also affects the presence of keels on glumes, rachis toughness, spike length, spike type, and culm height (Faris *et al.*, 2003; Simons *et al.*, 2006). Molecular analyses reveal that the domesticated (*Q*) and wild type (*q*) gene products differ by a single amino acid at position 329; the *Q* gene has an isoleucine whereas *q* has a valine (Simons *et al.*, 2006). This sequence change increases the stability of *Q* gene homodimers relative

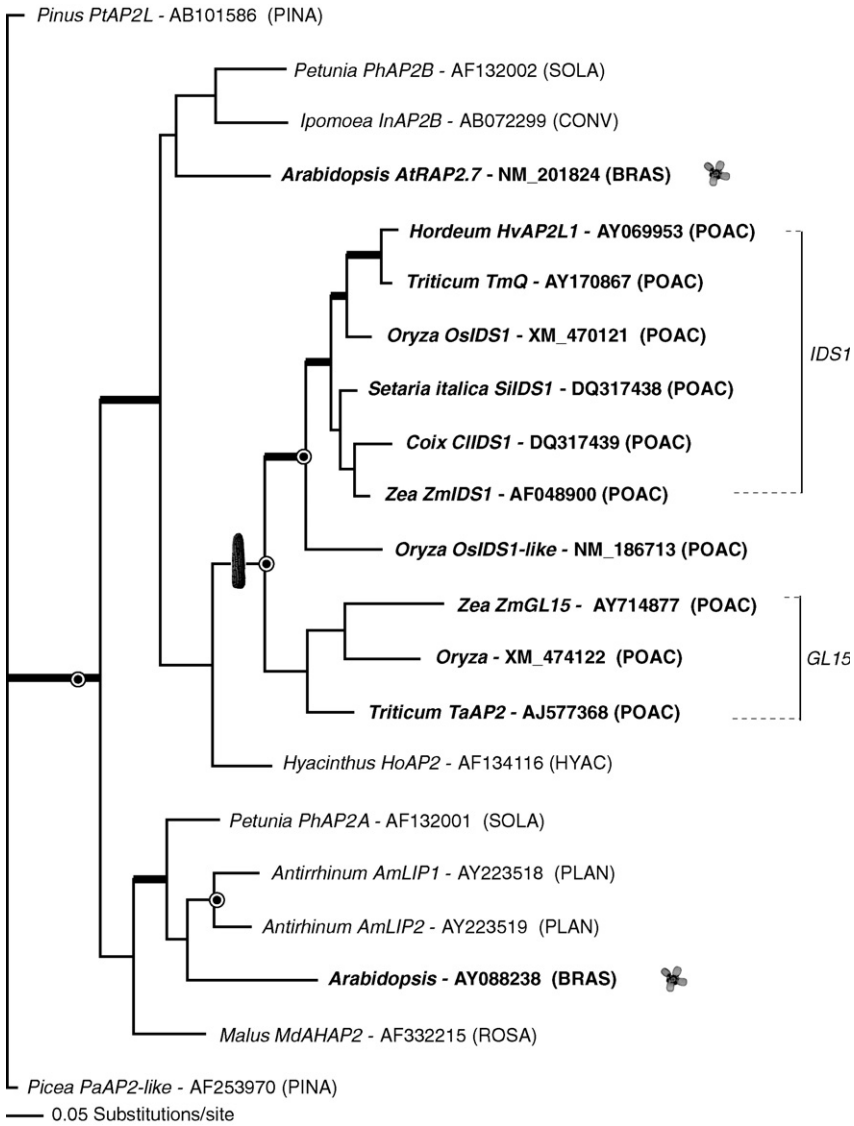

Fig. 9. Maximum likelihood phylogram of the 21 *INDETERMINATE SPIKELET1*-like gene data set, comprising 486 bp, visited by Bayesian Markov Chain Monte Carlo search using the General Time Reversible model with some invariant sites and gamma distributed rates (GTR + I + G). Tree rooted using *Pinus PtAP2L* sequence. $-\ln = 4601.91$. Bold branches are supported by posterior probabilities >0.95. Boldface, grass and *Arabidopsis* sequences. ◉, Inferred duplication event; *Arabidopsis* flower, *Arabidopsis* sequence; corn ear, grass clade. Plant families abbreviated to first four letters of name. *GL15, GLOSSY15; IDS1, INDETERMINATE SPIKELET1*.

to the wild type gene, suggesting that the domesticated inflorescence phenotype is dosage related and caused by increased amounts of the *Q* gene transcript (Simons *et al.*, 2006). Sister to the *IDS1* + *IDS1*-like clade is a clade of rice, wheat, and maize sequences containing the maize *GLOSSY15* (*ZmGL15*) gene, which regulates leaf epidermal development (Lauter *et al.*, 2005; Moose and Sisco, 1996).

The closest eudicot relatives of *ZmIDS1* are *Arabidopsis AtRAP2.7, Petunia PhAP2B*, and *Ipomoea ImAP2B*. Information in addition to the nucleotide sequence is only available for *PhAP2B*, which is expressed strongly in the outer cells of young inflorescence bracts, the epidermis of sepals and the ovary, and seed endosperm (Maes *et al.*, 2001).

Available data for *IDS1* and related sequences suggest the *IDS1* clade was produced from one of several duplications within monocots. The *ZmGL15* clade seems to have maintained the expression and possible function of related eudicot genes, whereas all available information on the *IDS1* clade suggests a novel role restricted to spikelet development resulting from a duplication near the origin of grasses.

Transcription factors containing *AP2* domains play roles in regulating root, leaf, flower, seed, and ovule development, often via miRNA posttranscriptional regulation (Lauter *et al.*, 2005; Riechmann and Meyerowitz, 1998). Regulation by miRNAs has been demonstrated for *GLOSSY15* (Lauter *et al.*, 2005). We expect that this mode of regulation should apply to *IDS1* as well.

IV. CONCLUSIONS

A goal of evolutionary developmental genetics is to explain the diversity of life in terms of modification of underlying genes. This requires linking phenotype to genotype not just in a couple of model organisms, but in entire groups of species. In this chapter, we have described some of the phenotypic diversity among grass inflorescences and shown that multiple genes have been cloned that are good candidates for regulating that diversity. In drawing connections between genes and phenotypes, it has been essential to describe the inflorescence in terms of meristems and primoridia and to focus on events in inflorescence development. The classic typology for describing inflorescences—spike, raceme, panicle—is not precise enough for this purpose.

The majority of the phenotypes described here do not occur in *Arabidopsis*. In addition, many of the genes we have analyzed have duplicated extensively in grasses or monocots. We suggest that the two observations are related—grass-specific genes have acquired new functions to create novel

grass-specific morphology. Multiple model systems are going to be necessary to understand the genetic basis of evolutionary diversification, and this study is just one of many illustrations of that point (Gewin, 2005). Valuable insights can come from comparison to eudicots, and in a few cases close comparisons can be made (e.g., *FEA2* with *CLV2*) but more often the grasses appear to be a unique system, which can only be studied by direct analysis of genes in the grass models, rice and maize.

Despite tremendous progress in cloning relevant genes, only a few have so far been studied broadly enough to document diversification within the grasses themselves. *LHS1, FUL1/2/3*, and *RA1* all exhibit variation suggesting that they may be involved in diversification within the grasses. Most of the other genes described above are good candidates, however. Once orthologs have been cloned from a variety of species, the next step is clearly to investigate differences in expression pattern, which is the first step toward determining function.

Gene duplication is one source of novel genes. Extensive large-scale, possibly whole-genome, duplications have occurred frequently throughout the tree of life. Flowering plants are no exception, with large-scale duplications estimated near the base of Brassicaceae, grasses, core eudicots, and papilionoid legumes (Bowers *et al.*, 2003; Paterson *et al.*, 2004). The genes described here vary widely in how much evidence they retain of these duplication events. At one extreme, phylogenetic analyses including *ZmFEA2* (Fig. 6) and *ZmTD1* (Bommert *et al.*, 2005a) do not identify any duplication events in grasses, Brassicaceae, or elsewhere; thus any duplicate genes have suffered the fate of most gene copies and have been removed from the genome, likely via selection (Lynch and Force, 2000). Phylogenetic analyses of other developmental genes identify duplication events near the base of grasses and/or Brassicaceae, but the number of estimated events varies, as does the timing of the duplication. In the *KN1*/*BP* clade of *KNOX* genes, three duplications occurred before divergence of the grasses, but no duplicates are retained in Brassicaceae (Fig. 3). In the *LG3*/*KNAT2* clade, both the grass duplication and the Brassicaceae duplication are evident. Among *TFL*-like genes, the *RCN1*/*RCN2* duplication appears to have preceded the grass duplication (based on the position of *Joinvillea*), but the Brassicaceae duplication left only one descendant gene in *Arabidopsis* (Fig. 5). The *AP1*/*CAL*/*FUL* phylogeny reveals descendants of three monocot duplications, at least one of which may correspond to the grass duplication, as well as one at the base of eudicots and one in Brassicaceae (Fig. 7). One duplication event is also estimated to have occurred within the *SEP3* lineage, although only a single *Arabidopsis SEP3* gene has been identified (Fig. 8). As in the grass *FUL*-like clade, two duplication events are estimated near the base of grasses

in the *SEPALLATA LOFSEP* clade, and in the *IDS1*-like gene family. The persistence of gene products following these gene duplication events has resulted in most of these grass genes being "novel."

Interpreting how and whether inflorescence genes have diversified functionally is limited by available data. However, the sparse data suggest a complex pattern. *ZmFEA2* and *AtCLV2*, along with *ZmTD1, OsFON1*, and the related *AtCLV1* have overlapping roles in defining meristem size, but their expression patterns and the pathways by which the genes regulate meristem size differ between grasses and *Arabidopsis*. In *Arabidopsis, clv1* and *clv2* mutants and *clv1 clv2*-double mutants have identical phenotypes, suggesting a simple pathway, whereas *Zmtd1 Zmfea2* double mutants have more pronounced phenotypes than either single mutant, indicating a more elaborate pathway in maize. Thus protein function has apparently diverged despite a lack of retained duplicates.

In contrast, the *TFL1* clade appears to be a small gene family in which successive gene duplications have produced little or no diversification in gene function. Despite different expression patterns and multiple duplications within flowering plants, the *Arabidopsis AtTFL1* and *AtCEN*, and grass *OsRCN1* and *OsRCN2* genes all regulate inflorescence branching and flowering time. Loss-of-function mutants in grasses, however, may uncover more and different gene functions.

Among class I *KNOX* genes, functional convergence has followed gene duplication. *ZmKN1* and *AtSTM* genes both play roles in meristem maintenance and *OsOSH15* and *AtBP* both regulate internodes (Long *et al.*, 1996; Sato *et al.*, 1998; Venglat *et al.*, 2002; Vollbrecht *et al.*, 2000). The phylogeny shows, however, that *ZmKN1* is actually more closely related to *AtBP*, whereas *AtSTM* is more closely related to *OsOSH15*. Thus, at least one pair of genes must have arrived at their similarity independently.

Gene duplications have also produced novel roles. Examples of this can be seen in the *SEPALLATA, FRUITFULL, IDS1*, and presumably *RA1* clades. The SEPALLATA protein LHS1 determines palea and lemma size and texture, whereas all *Arabidopsis* SEP proteins have redundant developmental roles (Ditta *et al.*, 2004; Pelaz *et al.*, 2000). In the case of FRUITFULL proteins, some aspects of function are conserved among family members (e.g., specifying meristem identity) whereas others appear to be novel roles that have evolved within grasses (vernalization-induced flowering). *Arabidopsis AP1/FUL* genes are restricted to floral tissue whereas grass *FUL*-like genes are expressed broadly throughout the plant. In the case of IDS1-like proteins, duplication has resulted in a diversification of functional roles, with *GLOSSY15* maintaining the apparent plesiomorphic condition of specifying

epidermal morphology, whereas the role of IDS1 proteins is restricted to the spikelet. The function of the *IDS1*-like gene from rice that is sister to *IDS1* in our analysis might expand the functional domain of this small gene family even further. *RA1* also appears to be a novel gene, and it will be of interest to reconstruct its evolutionary history.

Morphological novelty might arise in many ways. Here we have focused on gene duplication as one important phenomenon that provides the raw material on which selection can act. The phenotypic result of gene duplication depends, however, on changes within the genes themselves or the sequences that regulate them. Changes in gene expression are likely to underlie much evolutionary change, and these changes may arise easily through modification of *cis*-regulatory sequences (Doebley and Lukens, 1998; Moore and Purugganan, 2005). Partial loss of function can change developmental pathways by leading to accumulation of intermediate gene products, and then selecting for enhanced processing of the intermediates (Grotewold, 2005).

Gene level changes can then lead to novel morphological outcomes by changing the timing (heterochrony) or location (heterotopy) of genetic programs (Kellogg, 2000). Vollbrecht *et al.* (2005) have suggested that the proteins that control branching in the grasses, such as RA1, BA1, and others, may modulate the timing of development. Thus by extending or narrowing the developmental window for primary branch production, more or fewer are produced, leading to a marked change in morphology. They suggest that this may occur by genetic regulation of a mobile signal, or possibly the gene products themselves are the signal.

In conclusion, the rich diversity of the grasses, coupled with the availability of several model systems, makes the family an ideal group for the study of morphological evolution. Much of the structure of the grass inflorescence has no homolog in *Arabidopsis* or Brassicaceae, and the terminal inflorescence unit, the spikelet, is unique in the angiosperms. The grass family thus has provided and will continue to provide new information on the origins of novel structures and mechanisms of morphological diversification.

ACKNOWLEDGMENTS

We thank Doug Soltis, Pam Soltis, and Jim Leebens-Mack for the opportunity to contribute to this volume and Volker Brendel, George Chuck, Sarah Hake, Dave Jackson, Rob Martienssen, Torbert Rocheford, Bob Schmidt, and Erik Vollbrecht for discussions, information, and for commenting on the ideas presented here. We also thank two anonymous reviewers

for helpful suggestions, Lilla Bartkó for providing some of the grass *KNOTTED1* sequences, and Missouri Botanical Garden and the USDA for access to plant materials. This work was supported by National Science Foundation Grant DBI-0110189 (to E. A. K.).

REFERENCES

Alfaro, M. E., Zoller, S. and Lutzoni, F. (2003). Bayes or bootstrap? A simulation study comparing the performance of Bayesian Markov chain Monte Carlo sampling and bootstrapping in assessing phylogenetic confidence. *Molecular Biology and Evolution* **20**, 255–266.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Research* **25**, 3389–3402.

Andersen, C. H., Jensen, C. S. and Petersen, K. (2004). Similar genetic switch systems might integrate the floral inductive pathways in dicots and monocots. *Trends in Plant Science* **9**, 105–107.

Arabidopsis Genome Initiative. (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**, 796–815.

Ashikari, M., Sakakibara, H., Lin, S., Yamamoto, T., Takashi, T., Nishimura, A., Angeles, E. R., Qian, Q., Kitano, H. and Matsuoka, M. (2005). Cytokinin oxidase regulates rice grain production. *Science* **309**, 741–745.

Barkworth, M. E., Capels, M., Long, S. and Piep, M. B. (eds.) (2003). "Flora of North America: North of Mexico Volume 25, Magnoliophyta, Commelinidae (in part): Poaceae, part 2." Flora of North America. Oxford University Press, Oxford.

Baum, D. A. (2002). Identifying the genetic causes of phenotypic evolution: A review of experimental strategies. *In* "Developmental Genetics and Plant Evolution." (Q. E. B. Cronk, R. M. Bateman and J. Hawkins, eds.), pp. 493–507. Taylor and Francis, London.

Benjamins, R., Quint, A., Weijers, D., Hooykaas, P. and Offringa, R. (2001). The *PINOID* protein kinase regulates organ development in *Arabidopsis* by enhancing polar auxin transport. *Development* **128**, 4057–4067.

Bess, E. C., Doust, A. N. and Kellogg, E. A. (2005). A naked grass in the bristle clade: A phylogenetic and developmental study of *Panicum* section *Bulbosa* (Paniceae: Poaceae). *International Journal of Plant Sciences* **166**, 371–381.

Blazquez, M. (2000). Flower development pathways. *Journal of Cell Science* **113**, 3547–3548.

Bomblies, K. and Doebley, J. F. (2005). Molecular evolution of *FLORICAULA/LEAFY* orthologs in the Andropogoneae (Poaceae). *Molecular Biology and Evolution* **22**, 1082–1094.

Bomblies, K., Wang, R. L., Ambrose, B. A., Schmidt, R. J., Meeley, R. B. and Doebley, J. (2003). Duplicate *FLORICAULA/LEAFY* homologs *zfl1* and *zfl2* control inflorescence architecture and flower patterning in maize. *Development* **130**, 2385–2395.

Bommert, P., Lunde, C., Nardmann, J., Vollbrecht, E., Running, M., Jackson, D., Hake, S. and Werr, W. (2005a). *thick tassel dwarf1* encodes a putative maize ortholog of the *Arabidopsis CLAVATA1* leucine-rich repeat receptor-like kinase. *Development* **132**, 1235–1245.

Bommert, P., Satoh-Nagasawa, N., Jackson, D. and Hirano, H. Y. (2005b). Genetics and evolution of inflorescence and flower development in grasses. *Plant Cell and Physiology* **46**, 69–78.

Bonnett, O. T. (1948). Ear and tassel development in maize. *Annals of the Missouri Botanical Garden* **35**, 269–287.

Bossinger, G. (1990). *Klassifizierung von entwicklungsmutanten der gerste anhand einer interpretation des pflanzanaufbaus de Poaceae aus phytomeren*. Dissertation, Rheinischen Friedrich-Wilhelms-Universität zu Bonn, Bonn, Germany.

Bowers, J. E., Chapman, B. A., Rong, J. and Paterson, A. H. (2003). Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature* **422**, 433–438.

Buck, M. and Atchley, W. (2003). Phylogenetic analysis of plant basic helix-loop-helix proteins. *Journal of Molecular Evolution* **56**, 742–750.

Cacharrón, J., Saedler, H. and Theissen, G. (1999). Expression of MADS box genes *ZMM8* and *ZMM14* during inflorescence development of *Zea mays* discriminates between the upper and the lower floret of each spikelet. *Developmental Genes and Evolution* **209**, 411–420.

Campbell, C. S. and Kellogg, E. A. (1987). Sister group relationships of the Poaceae. *In* "Grass Systematics and Evolution" (T. R. Soderstrom, K. W. Hilu, C. S. Campbell and M. E. Barkworth, eds.), pp. 217–224. Smithsonian Press, Washington, DC.

Chantret, N., Salse, J., Sabot, F., Rahman, S., Bellec, A., Laubin, B., Dubois, I., Dossat, C., Sourdille, P., Joudrier, P., Gautier, M. F., Cattolico, L., *et al.* (2005). Molecular basis of evolutionary events that shaped the hardness locus in diploid and polyploid wheat species (*Triticum* and *Aegilops*). *Plant Cell* **17**, 1033–1045.

Christensen, S. K., Dagenais, N., Chory, J. and Weigel, D. (2000). Regulation of auxin response by the protein kinase PINOID. *Cell* **100**, 469–478.

Chuck, G., Meeley, R. B. and Hake, S. (1998). The control of maize spikelet meristem fate by the *APETALA2-like* gene *indeterminate spikelet1*. *Genes and Development* **12**, 1145–1154.

Chuck, G., Muszynski, M., Kellogg, E., Hake, S. and Schmidt, R. J. (2002). The control of spikelet meristem identity by the *branched silkless1* gene in maize. *Science* **298**, 1238–1241.

Clark, L. G. and Judziewicz, E. J. (1996). The grass subfamilies Anomochlooideae and Pharoideae (Poaceae). *Taxon* **45**, 641–645.

Clark, L. G. and Pohl, R. W. (1996). "Agnes Chase's First Book of Grasses." Smithsonian Institution Press, Washington, DC.

Clark, S. E., Williams, R. W. and Meyerowitz, E. M. (1997). The *CLAVATA1* gene encodes a putative receptor kinase that controls shoot and floral meristem size in *Arabidopsis*. *Cell* **89**, 575–585.

Clayton, D. K. and Renvoize, S. A. (1986). "Genera Graminum: Grasses of the World." Her Majesty's Stationery Office, London.

Clifford, H. T. (1987). Spikelet and floral morphology. *In* "Grass Systematics and Evolution" (T. R. Soderstrom, K. W. Hilu, C. S. Campbell and M. E. Barkworth, eds.), pp. 21–30. Smithsonian Institute, Washington, DC.

Danyluk, J., Kane, N. A., Breton, G., Limin, A. E., Fowler, D. B. and Sarhan, F. (2003). *TaVRT-1*, a putative transcription factor associated with vegetative to reproductive transition in cereals. *Plant Physiology* **132**, 1849–1860.

Dellaporta, S. L. (1994). Plant DNA miniprep and microprep: Versions 2.1–2.3. *In* "The Maize Handbook" (M. Freeling and V. Walbot, eds.), pp. 522–525. Springer-Verlag, New York.

Dinneny, J. R., Yadegari, R., Fischer, R. L., Yanofsky, M. F. and Weigel, D. (2004). The role of *JAGGED* in shaping lateral organs. *Development* **131**, 1101–1110.

Ditta, G., Pinyopich, A., Robles, P., Pelaz, S. and Yanofsky, M. F. (2004). The *SEP4* gene of *Arabidopsis thaliana* functions in floral organ and meristem identity. *Current Biology* **14**, 1935–1940.

Doebley, J. and Lukens, L. (1998). Transcriptional regulators and the evolution of plant form. *Plant Cell* **10**, 1075–1082.

Doust, A. N. and Kellogg, E. A. (2002). Inflorescence diversification in the panicoid bristle grass clade (Paniceae, Poaceae): Evidence from molecular phylogenies and developmental morphology. *American Journal of Botany* **89**, 1203–1222.

Doust, A. N., Devos, K. M., Gadberry, M. D., Gale, M. D. and Kellogg, E. A. (2005). The genetic basis for inflorescence variation between foxtail and green millet (Poaceae). *Genetics* **169**, 1659–1672.

Evans, M. W. (1940). Developmental morphology of the growing point of the shoot and the inflorescence in grasses. *Journal of Agricultural Research* **61**, 481–520.

Ewing, B., Hillier, L., Wendl, M. C. and Green, P. (1998). Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Research* **8**, 175–185.

Faris, J. D., Fellers, J. P., Brooks, S. A. and Gill, B. S. (2003). A bacterial artificial chromosome contig spanning the major domestication locus Q in wheat and identification of a candidate gene. *Genetics* **164**, 311–321.

Ferrandiz, C., Gu, Q., Martienssen, R. and Yanofsky, M. F. (2000). Redundant regulation of meristem identity and plant architecture by *FRUITFULL, APETALA1* and *CAULIFLOWER*. *Development* **127**, 725–734.

Fitch, W. M. (1970). Distinguishing homologous from analogous proteins. *Systematic Zoology* **19**, 99–113.

Force, A., Lynch, M., Pickett, F. B., Amores, A., Yan, Y. L. and Postlethwait, J. (1999). Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* **151**, 1531–1545.

Force, A., Cresko, W. A., Pickett, F. B., Proulx, S. R., Amemiya, C. and Lynch, M. (2005). The origin of subfunctions and modular gene regulation. *Genetics* **170**, 433–446.

Fornara, F., Parenicova, L., Falasca, G., Pelucchi, N., Masiero, S., Ciannamea, S., Lopez-Dee, Z., Altamura, M. M., Colombo, L. and Kater, M. M. (2004). Functional characterization of *OsMADS18*, a member of the *AP1/SQUA* subfamily of MADS box genes. *Plant Physiology* **135**, 2207–2219.

Foster, T., Yamaguchi, J., Wong, B. C., Veit, B. and Hake, S. (1999). *Gnarley1* is a dominant mutation in the *knox4* homeobox gene affecting cell shape and identity. *Plant Cell* **11**, 1239–1252.

Fu, D., Szucs, P., Yan, L., Helguera, M., Skinner, J. S., von Zitzewitz, J., Hayes, P. M. and Dubcovsky, J. (2005). Large deletions within the first intron in *VRN-1* are associated with spring growth habit in barley and wheat. *Molecular Genetics and Genomics* **273**, 54–65.

Gadberry, M. D., Malcomber, S. T., Doust, A. N. and Kellogg, E. A. (2005). Primaclade—a flexible tool to find conserved PCR primers across multiple species. *Bioinformatics* **21**, 1263–1264.

Gallavotti, A., Zhao, Q., Kyozuka, J., Meeley, R. B., Ritter, M. K., Doebley, J. F., Pe, M. E. and Schmidt, R. J. (2004). The role of *barren stalk1* in the architecture of maize. *Nature* **432**, 630–635.

Gewin, V. (2005). Functional genomics thickens the biological plot. *PLoS Biology* **3**, e219.

Gocal, G. F., King, R. W., Blundell, C. A., Schwartz, O. M., Andersen, C. H. and Weigel, D. (2001). Evolution of floral meristem identity genes. Analysis of *Lolium temulentum* genes related to *APETALA1* and *LEAFY* of *Arabidopsis*. *Plant Physiology* **125**, 1788–1801.

Goff, S. A., Ricke, D., Lan, T. H., Presting, G., Wang, R., Dunn, M., Glazebrook, J., Sessions, A., Oeller, P., Varma, H., Hadley, D. Hutchison, D., *et al*. (2002). A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science* **296**, 92–100.

Grass Phylogeny Working Group. (2001). Phylogeny and subfamilial classification of the grasses (Poaceae). *Annals of the Missouri Botanical Garden* **88**, 373–457.

Grotewold, E. (2005). Plant metabolic diversity: A regulatory perspective. *Trends in Plant Science* **10**, 57–62.

Hake, S., Smith, H. M., Holtan, H., Magnani, E., Mele, G. and Ramirez, J. (2004). The role of *knox* genes in plant development. *Annual Review of Cell and Developmental Biology* **20**, 125–151.

Haldane, J. B. S. (1932). "The Causes of Evolution." Longmans, Green and Co., London.

Haldane, J. B. S. (1933). The part played by recurrent mutation in evolution. *American Naturalist* **67**, 5–19.

Hanzawa, Y., Money, T. and Bradley, D. (2005). A single amino acid converts a repressor to an activator of flowering. *Proceedings of the National Academy of Science of the United States of America* **102**, 7748–7753.

Harrison, C. J., Corley, S. B., Moylan, E. C., Alexander, D. L., Scotland, R. W. and Langdale, J. A. (2005). Independent recruitment of a conserved developmental mechanism during leaf evolution. *Nature* **434**, 509–514.

He, X. and Zhang, J. (2005). Rapid subfunctionalization accompanied by prolonged and substantial neofunctionalization in duplicate gene evolution. *Genetics* **169**, 1157–1164.

Hillis, D. M., Moritz, C. and Mable, B. K. (1996). "Molecular Systematics." Sinauer Associates, Sunderland, MA.

Huang, H., Tudor, M., Weiss, C. A., Hu, Y. and Ma, H. (1995). The *Arabidopsis* MADS-box gene *AGL3* is widely expressed and encodes a sequence-specific DNA-binding protein. *Plant Molecular Biology* **28**, 549–567.

Huelsenbeck, J. P. and Ronquist, F. (2001). MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**, 754–755.

Hughes, A. L. (1999). "Adaptive Evolution of Genes and Genomes." Oxford University Press, New York.

Ikeda, K., Sunohara, H. and Nagato, Y. (2004). Developmental course of inflorescence and spikelet in rice. *Breeding Science* **54**, 147–156.

Ikeda, K., Nagasawa, N. and Nagato, Y. (2005). *ABERRANT PANICLE ORGANIZATION 1* temporally regulates meristem identity in rice. *Developmental Biology* **282**, 349–360.

Irish, E. (1997). Class II tassel seed mutations provide evidence for multiple types of inflorescence meristems in maize (Poaceae). *American Journal of Botany* **84**, 1502–1515.

Itoh, J., Nonomura, K., Ikeda, K., Yamaki, S., Inukai, Y., Yamagishi, H., Kitano, H. and Nagato, Y. (2005). Rice plant development: From zygote to spikelet. *Plant Cell and Physiology* **46**, 23–47.

Jackson, D. and Hake, S. (1999). Control of phyllotaxy in maize by the *abphyl1* gene. *Development* **126**, 315–323.

Jackson, D., Veit, B. and Hake, S. (1994). Expression of maize *KNOTTED1* related homeobox genes in the shoot apical meristem predicts patterns of morphogenesis in the vegetative shoot. *Development* **120**, 405–413.

Jeanmougin, F., Thompson, J. D., Gouy, M., Higgins, D. G. and Gibson, T. J. (1998). Multiple sequence alignment with Clustal X. *Trends in Biochemical Sciences* **23**, 403–405.

Jensen, C. S., Salchert, K. and Nielsen, K. K. (2001). A *TERMINAL FLOWER1-Like* gene from perennial ryegrass involved in floral transition and axillary meristem identity. *Plant Physiology* **125**, 1517–1528.

Jensen, L. B., Andersen, J. R., Frei, U., Xing, Y., Taylor, C., Holm, P. B. and Lubberstedt, T. (2005). QTL mapping of vernalization response in perennial ryegrass (*Lolium perenne* L.) reveals co-location with an orthologue of wheat *VRN1*. *Theoretical and Applied Genetics* **110**, 527–536.

Jeon, J. S., Jang, S., Lee, S., Nam, J., Kim, C., Lee, S. H., Chung, Y. Y., Kim, S. R., Lee, Y. H., Cho, Y. G. and An, G. (2000). *leafy hull sterile1* is a homeotic mutation in a rice MADS box gene affecting rice flower development. *Plant Cell* **12**, 871–884.

Jeong, S., Trotochaud, A. E. and Clark, S. E. (1999). The *Arabidopsis CLAVATA2* gene encodes a receptor-like protein required for the stability of the *CLAVATA1* receptor-like kinase. *Plant Cell* **11**, 1925–1934.

Judziewicz, E. J., Clark, L. G., Londoño, X. and Stern, M. J. (1999). "American Bamboos." Smithsonian Institution Press, Washington, DC.

Kayes, J. M. and Clark, S. E. (1998). *CLAVATA2*, a regulator of meristem and organ development in *Arabidopsis*. *Development* **125**, 3843–3851.

Kellogg, E. A. (1996). Integrating genetics, phylogenetics, and developmental biology. *In* "The Impact of Plant Molecular Genetics." (B. W. S. Sobral, ed.), pp. 159–172. Birkhäuser, Cambridge, MA.

Kellogg, E. A. (2000a). The grasses: A case study in macroevolution. *Annual Review of Ecology and Systematics* **31**, 217–238.

Kellogg, E. A. (2000b). A model of inflorescence development. *In* "Monocots: Systematics and Evolution" (K. L. Wilson and D. A. Morrison, eds.), pp. 84–88. CSIRO, Melbourne.

Kellogg, E. A. (2000c). Molecular and morphological evolution in Andropogoneae. *In* "Grasses: Systematics and Evolution" (S. W. L. Jacobs and J. E. Everett, eds.). CSIRO, Melbourne, Australia.

Kellogg, E. A. (2003). Genome evolution: It's all relative. *Nature* **422**, 383–384.

Kellogg, E. A., Hiser, K. M. and Doust, A. N. (2004). Taxonomy, phylogeny, and inflorescence development of the genus *Ixophorus* (Panicoideae: Poaceae). *International Journal of Plant Sciences* **165**, 1089–1105.

Kellogg, E. A. and Shaffer, H. B. (1993). Model organisms in evolutionary studies. *Systematic Biology* **42**, 409–414.

Kerstetter, R., Vollbrecht, E., Lowe, B., Veit, B., Yamaguchi, J. and Hake, S. (1994). Sequence analysis and expression patterns divide the maize *knotted1-like* homeobox genes into two classes. *Plant Cell* **6**, 1877–1887.

Kerstetter, R. A., Laudencia-Chingcuanco, D., Smith, L. G. and Hake, S. (1997). Loss-of-function mutations in the maize homeobox gene, *knotted1*, are defective in shoot meristem maintenance. *Development* **124**, 3045–3054.

Klaus, H. (1966). Ontogenetische und histogentische Ungersuchungen an der Gerste (*Hordeum distichon* L.). *Botanisches Jahrbucher* **85**, 45–79.

Komatsu, K., Maekawa, M., Ujiie, S., Satake, Y., Furutani, I., Okamoto, H., Shimamoto, K. and Kyozuka, J. (2003a). *LAX* and *SPA*: Major regulators of shoot branching in rice. *Proceedings of the National Academy of Science of the United States of America* **100**, 11765–11770.

Komatsu, M., Chujo, A., Nagato, Y., Shimamoto, K. and Kyozuka, J. (2003b). *FRIZZY PANICLE* is required to prevent the formation of axillary meristems and to establish floral meristem identity in rice spikelets. *Development* **130**, 3841–3850.

Kyozuka, J., Konishi, S., Nemoto, K., Izawa, T. and Shimamoto, K. (1998). Down-regulation of *RFL*, the *FLO/LFY* homolog of rice, accompanied with panicle branch initiation. *Proceedings of the National Academy of Science of the United States of America* **95**, 1979–1982.

Larget, B. and Simon, D. L. (1999). Markov Chain Monte Carlo algorithms for the Bayesian analysis of phylogenetic trees. *Molecular Biology and Evolution* **16**, 750–759.

Lauter, N., Kampani, A., Carlson, S., Goebel, M. and Moose, S. P. (2005). *microRNA172* down-regulates *glossy15* to promote vegetative phase change in maize. *Proceedings of the National Academy of Science of the United States of America* **102**, 9412–9417.

Lawrence, C. J., Seigfried, T. E. and Brendel, V. (2005). The maize genetics and genomics database. The community resource for access to diverse maize data. *Plant Physiology* **138**, 55–58.

Le Roux, L. G. and Kellogg, E. A. (1999). Floral development and the formation of unisexual spikelets in the Andropogoneae (Poaceae). *American Journal of Botany* **86**, 354–366.

Liljegren, S. J., Roeder, A. H., Kempin, S. A., Gremski, K., Ostergaard, L., Guimil, S., Reyes, D. K. and Yanofsky, M. F. (2004). Control of fruit patterning in *Arabidopsis* by *INDEHISCENT*. *Cell* **116**, 843–853.

Litt, A. and Irish, V. F. (2003). Duplication and diversification in the *APETALA1/FRUITFULL* floral homeotic gene lineage: Implications for the evolution of floral development. *Genetics* **165**, 821–833.

Long, J. and Barton, M. K. (2000). Initiation of axillary and floral meristems in *Arabidopsis*. *Developmental Biology* **218**, 341–353.

Long, J. A., Moan, E. I., Medford, J. I. and Barton, M. K. (1996). A member of the *KNOTTED* class of homeodomain proteins encoded by the *STM* gene of *Arabidopsis*. *Nature* **379**, 66–69.

Lunde, C. and Hake, S. (2006). Florets and rosettes: Meristem genes in maize and *Arabidopsis*. *Maydica* **50**, 451–458.

Lynch, M. and Force, A. (2000). The probability of duplicate gene preservation by subfunctionalization. *Genetics* **154**, 459–473.

Ma, H., Yanofsky, M. F. and Meyerowitz, E. M. (1991). *AGL1-AGL6*, an *Arabidopsis* gene family with similarity to floral homeotic and transcription factor genes. *Genes and Development* **5**, 484–495.

Maddison, D. R. and Maddison, W. P. (2003). "MacClade: Analysis of Phylogeny and Character Evolution." Sinauer Associates, Sunderland, MA.

Maes, T., Van de Steene, N., Zethof, J., Karimi, M., D'Hauw, M., Mares, G., Van Montagu, M. and Gerats, T. (2001). *Petunia Ap2-like* genes and their role in flower and seed development. *Plant Cell* **13**, 229–244.

Malcomber, S. T. and Kellogg, E. A. (2004). Heterogeneous expression patterns and separate roles of the *SEPALLATA* gene *LEAFY HULL STERILE1* in grasses. *Plant Cell* **16**, 1692–1706.

Malcomber, S. T. and Kellogg, E. A. (2005). *SEPALLATA* gene diversification: Brave new whorls. *Trends in Plant Science* **10**, 427–435.

McSteen, P. and Hake, S. (2001). *barren inflorescence2* regulates axillary meristem development in the maize inflorescence. *Development* **128**, 2881–2891.

Moncur, M. W. (1981). "Floral Intiation in Field Crops." CSIRO, Melbourne.

Moore, R. C. and Purugganan, M. D. (2003). The early stages of duplicate gene evolution. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 15682–15687.

Moore, R. C. and Purugganan, M. D. (2005). The evolutionary dynamics of plant duplicate genes. *Current Opinion in Plant Biology* **8**, 122–128.

Moose, S. P. and Sisco, P. H. (1996). *Glossy15*, an *APETALA2-like* gene from maize that regulates leaf epidermal cell identity. *Genes and Development* **10**, 3018–3027.

Nakagawa, M., Shimamoto, K. and Kyozuka, J. (2002). Overexpression of *RCN1* and *RCN2*, rice *TERMINAL FLOWER 1/CENTRORADIALIS* homologs, confers delay of phase transition and altered panicle morphology in rice. *Plant Journal* **29**, 743–750.

Nylander, J. A. A. (2004). *MrModeltest, Version 2*, Program distributed by the author. Evolutionary Biology Centre, Uppsala University.

Ohno, S. (1970). "Evolution by Gene Duplication." Springer-Verlag, London.

Orr, A. R. and Sundberg, M. D. (2004). Inflorescence development in a new teosinte: *Zea nicaraguensis* (Poaceae). *American Journal of Botany* **91**, 165–173.

Orr, A. R., Kaparthi, R., Dewald, C. L. and Sundberg, M. D. (2001). Analysis of inflorescence organogenesis in Eastern gamagrass, *Tripsacum dactyloides* (Poaceae): The wild type and the gynomonoecious *GSF1* mutant. *American Journal of Botany* **88**, 363–381.

Ohno, C. K., Reddy, G. V., Heisler, M. G. and Meyerowitz, E. M. (2004). The *Arabidopsis JAGGED* gene encodes a zinc finger protein that promotes leaf tissue development. *Development* **131**, 1111–1122.

Paterson, A. H., Bowers, J. E. and Chapman, B. A. (2004). Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. *Proceeding of the National Academy of Sciences of the United States of America* **101**, 9903–9908.

Pelaz, S., Ditta, G. S., Baumann, E., Wisman, E. and Yanofsky, M. F. (2000). B and C floral organ identity functions require *SEPALLATA* MADS-box genes. *Nature* **405**, 200–203.

Pelaz, S., Gustafson-Brown, C., Kohalmi, S. E., Crosby, W. L. and Yanofsky, M. F. (2001). *APETALA1* and *SEPALLATA3* interact to promote flower development. *Plant Journal* **26**, 385–394.

Pelucchi, N., Fornara, F., Favalli, C., Masiero, S., Lago, C., Pe, M. E., Colombo, L. and Kater, M. M. (2002). Comparative analysis of rice MADS-box genes expressed during flower development. *Sexual Plant Reproduction* **15**, 113–122.

Prasad, K., Sriram, P., Kumar, C. S., Kushalappa, K. and Vijayraghavan, U. (2001). Ectopic expression of rice *OsMADS1* reveals a role in specifying the lemma and palea, grass floral organs analogous to sepals. *Developmental Genes and Evolution* **211**, 281–290.

Ratcliffe, O. J., Amaya, I., Vincent, C. A., Rothstein, S., Carpenter, R., Coen, E. S. and Bradley, D. J. (1998). A common mechanism controls the life cycle and architecture of plants. *Development* **125**, 1609–1615.

Reinheimer, R., Pozner, R. and Vegetti, A. C. (2005). Inflorescence, spikelet, and floral development in *Panicum maximum* and *Urochloa plantaginea* (Poaceae). *American Journal of Botany* **92**, 565–575.

Reiser, L., Sanchez-Baracaldo, P. and Hake, S. (2000). Knots in the family tree: Evolutionary relationships and functions of *knox* homeobox genes. *Plant Molecular Biology* **42**, 151–166.

Riechmann, J. L. and Meyerowitz, E. M. (1998). The *AP2/EREBP* family of plant transcription factors. *Journal of Biological Chemistry* **379**, 633–646.

Ronse Decraene, L. P., Linder, H. P. and Smets, E. F. (2002). Ontogeny and evolution of the flowers of South African Restionaceae with special emphasis on the gynoecium. *Plant Systematics and Evolution* **231**, 225–258.

Rudall, P. J., Stuppy, W., Cunniff, J., Kellogg, E. A. and Briggs, B. G. (2005). Evolution of reproductive structures in grasses inferred by sister-group comparisons with their putative closest living relatives, Ecdeiocoleaceae. *American Journal of Botany* **92**, 1432–1443.

Sato, Y., Sentoku, N., Nagato, Y. and Matsuoka, M. (1998). Isolation and characterization of a rice homebox gene, *OSH15*. *Plant Molecular Biology* **38**, 983–997.

Savidge, B., Rounsley, S. D. and Yanofsky, M. F. (1995). Temporal relationship between the transcription of two *Arabidopsis* MADS box genes and the floral organ identity genes. *Plant Cell* **7**, 721–733.

Schmitz, J., Franzen, R., Ngyuen, T. H., Garcia-Maroto, F., Pozzi, C., Salamini, F. and Rohde, W. (2000). Cloning, mapping and expression analysis of barley MADS-box genes. *Plant Molecular Biology* **42**, 899–913.

Sentoku, N., Sato, Y., Kurata, N., Ito, Y., Kitano, H. and Matsuoka, M. (1999). Regional expression of the rice *KN1*-type homeobox gene family during embryo, shoot, and flower development. *Plant Cell* **11**, 1651–1664.

Shannon, S. and Meeks-Wagner, D. R. (1991). A mutation in the *Arabidopsis TFL1* gene affects inflorescence meristem development. *Plant Cell* **3**, 877–892.

Sharma, V. K., Carles, C. and Fletcher, J. C. (2003). Maintenance of stem cell populations in plants. *Proceedings of the National Academy of Science of the United States of America* **100** (Suppl. 1), 11823–11829.

Sharman, B. C. (1947). The biology and developmental morphology of the shoot apex in the Gramineae. *New Phytologist* **46**, 20–34.

Simmons, M. P., Pickett, K. M. and Miya, M. (2004). How meaningful are Bayesian support values? *Molecular Biology and Evolution* **21**, 188–199.

Simons, K. J., Fellers, J. P., Trick, H. N., Zhang, Z., Tai, Y.-S., Gill, B. S. and Faris, J. D. (2006). Molecular characterization of the major wheat domestication gene *Q*. *Genetics* **172**, 547–555.

Smyth, D. R., Bowman, J. L. and Meyerowitz, E. M. (1990). Early flower development in *Arabidopsis*. *Plant Cell* **2**, 755–767.

Soltis, D. E., Soltis, P. S., Endress, P. K. and Chase, M. W. (2005). "Phylogeny and Evolution of Angiosperms." Sinauer Associates, Sunderland, MA.

Sonnhammer, E. L. L. (2002). Orthology, paralogy and proposed classification for paralog subtypes. *Trends in Genetics* **18**, 619–620.

Stapf, O. (ed.) (1917). "Gramineae. Flora of Tropical Africa."London, Lowell Reeve & Co.

Stevens, P. F. (2001 onwards). Angiosperm Phylogeny Website, Version 6, http:/www.mobot.org/MOBOT/research/APweb/.

Sundberg, M. D. and Orr, A. R. (1996). Early inflorescence and floral development in *Zea mays* land race Chapalote (Poaceae). *American Journal of Botany* **83**, 1255–1265.

Suzaki, T., Sato, M., Ashikari, M., Miyoshi, M., Nagato, Y. and Hirano, H. Y. (2004). The gene *FLORAL ORGAN NUMBER1* regulates floral meristem size in rice and encodes a leucine-rich repeat receptor kinase orthologous to *Arabidopsis CLAVATA1*. *Development* **131**, 5649–5657.

Swofford, D. L. (2000). "PAUP*: Phylogenetic Analysis Using Parsimony." Sinauer Associates, Sunderland, MA.

Taguchi-Shiobara, F., Yuan, Z., Hake, S. and Jackson, D. (2001). The *fasciated ear2* gene encodes a leucine-rich repeat receptor-like protein that regulates shoot meristem proliferation in maize. *Genes and Development* **15**, 2755–2766.

Takahashi, M., Nagasawa, N., Kitano, H. and Nagato, Y. (1998). *panicle phytomer1* mutations affect panicle architecture of rice. *Theoretical and Applied Genetics* **96**, 1050–1056.

Tian, C., Wan, P., Sun, S., Li, J. and Chen, M. (2004). Genome-wide analysis of the GRAS gene family in rice and *Arabidopsis*. *Plant Molecular Biology* **54**, 519–532.

To, J. P. C., Haberer, G., Ferreira, F. J., Deruere, J., Mason, M. G., Schaller, G. E., Alonso, J. M., Ecker, J. R. and Kieber, J. J. (2004). Type-A *Arabidopsis* response regulators are partially redundant negative regulators of cytokinin signaling. *Plant Cell* **16**, 658–671.

Trevaskis, B., Bagnall, D. J., Ellis, M. H., Peacock, W. J. and Dennis, E. S. (2003). MADS box genes control vernalization-induced flowering in cereals. *Proceedings of the National Academy of Science of the United States of America* **100**, 13099–13104.

Vandenbussche, M., Theissen, G., Van de Peer, Y. and Gerats, T. (2003). Structural diversification and neo-functionalization during floral MADS-box gene evolution by C-terminal frameshift mutations. *Nucleic Acids Research* **31**, 4401–4409.

Vegetti, A. C. and Weberling, F. (1996). The structure of the paracladial zone in Poaceae. *Taxon* **45**, 453–460.

Venglat, S. P., Dumonceaux, T., Rozwadowski, K., Parnell, L., Babic, V., Keller, W., Martienssen, R., Selvaraj, G. and Datla, R. (2002). The homeobox gene *BREVIPEDICELLUS* is a key regulator of inflorescence architecture in *Arabidopsis*. *Proceedings of the National Academy of Science of the United States of America* **99**, 4730–4735.

Vollbrecht, E., Veit, B., Sinha, N. and Hake, S. (1991). The developmental gene *Knotted-1* is a member of a maize homeobox gene family. *Nature* **350**, 241–243.

Vollbrecht, E., Reiser, L. and Hake, S. (2000). Shoot meristem size is dependent on inbred background and presence of the maize homeobox gene, *knotted1*. *Development* **127**, 3161–3172.

Vollbrecht, E., Springer, P. S., Goh, L., Buckler, E. S., IV and Martienssen, R. (2005). Architecture of floral branch systems in maize and related grasses. *Nature* **436**, 1119–1126.

Walsh, J. B. (1995). How often do duplicated genes evolve new functions? *Genetics* **139**, 421–428.

Ware, D., Jaiswal, P., Ni, J., Pan, X., Chang, K., Clark, K., Teytelman, L., Schmidt, S., Zhao, W., Cartinhour, S., McCouch, S. and Stein, L. (2002). Gramene: A resource for comparative grass genomics. *Nucleic Acids Research* **30**, 103–105.

Watson, L. and Dallwitz, M. J. (1992 onwards). Grass Genera of the World: Descriptions, Illustrations, Identification, and Information Retrieval; including Synonyms, Morphology, Anatomy, Physiology, Phytochemistry, Cytology, Classification, Pathogens, World and Local Distribution, and References. http://biodiversity.uno.edu/delta/.

Weigel, D., Alvarez, J., Smyth, D. R., Yanofsky, M. F. and Meyerowitz, E. M. (1992). *LEAFY* controls floral meristem identity in *Arabidopsis*. *Cell* **69**, 843–859.

Weigel, D. and Nilsson, O. (1995). A developmental switch sufficient for flower initiation in diverse plants. *Nature* **377**, 495–500.

Weir, C. E. and Dale, H. M. (1960). A developmental study of wild rice, *Zizania aquatica* L. *Canadian Journal of Botany* **38**, 719–739.

Wernersson, R. and Pedersen, A. G. (2003). RevTrans: Multiple alignment of coding DNA from aligned amino acid sequences. *Nucleic Acids Research* **31**, 3537–3539.

Yan, L., Loukoianov, A., Tranquilli, G., Helguera, M., Fahima, T. and Dubcovsky, J. (2003). Positional cloning of the wheat vernalization gene VRN1. *Proceedings of the National Academy of Science of the United States of America* **100**, 6263–6268.

Yu, J., Hu, S., Wang, J., Wong, G. K., Li, S., Liu, B., Deng, Y., Dai, L., Zhou, Y., Zhang, X., Cao, M., Liu, J., *et al*. (2002). A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). *Science* **296**, 79–92.

Yu, J., Wang, J., Lin, W., Li, S., Li, H., Zhou, J., Ni, P., Dong, W., Hu, S., Zeng, C., Zhang, J., Zhang, Y., *et al*. (2005). The genomes of *Oryza sativa*: A history of duplications. *PLoS Biology* **3**, e38.

Zahn, L. M., Kong, H., Leebens-Mack, J. H., Kim, S., Soltis, P. S., Landherr, L. L., Soltis, D. E., Depamphilis, C. W. and Ma, H. (2005). The evolution of the *SEPALLATA* subfamily of MADS-box genes: A pre-angiosperm origin with multiple duplications throughout angiosperm history. *Genetics* **169**, 2225–2239.

Zhang, J. Z. (2003). Evolution by gene duplication: An update. *Trends in Ecology & Evolution* **18**, 292–298.