New Phytologist

Research

Evolution of unisexual flowers in grasses (Poaceae) and the putative sex-determination gene, *TASSELSEED2* (*TS2*)

Simon T. Malcomber and Elizabeth A. Kellogg

Research 223, Department of Biology, University of Missouri – St Louis, One University Boulevard, Saint Louis, MO 63121, USA

Author for correspondence:
Elizabeth A. Kellogg
Tel: +1 314 516 6217
Fax: +1 314 516 6233
Email: tkellogg@umsl.edu

Received: *30 November 2005*
Accepted: *7 February 2006*

Summary

• Unisexuality has evolved repeatedly in flowering plants, but its genetic control is not understood in most cases. In maize (*Zea mays*), unisexual flower development is regulated by a short-chain dehydrogenase/reductase protein, TASSELSEED2 (TS2), but its role in other grass lineages is unknown.

• *TS2* was cloned and sequenced from a broad range of grasses and compared to available sequences from other flowering plants using phylogenetic analysis and tests for selection. Gene expression was investigated using reverse transcriptase–polymerase chain reaction (RT-PCR) and *in situ* hybridization.

• *TS2* orthologs appear to be restricted to monocots. The TS2 protein sequence was found to be generally under purifying selection in bisexual and unisexual lineages alike. Only one site, in unisexual herbaceous bamboos, is potentially under positive selection. *TS2* was expressed broadly in all sampled tissues of unisexual and bisexual grasses, and was also expressed in rice flowers in floral organs that do not abort.

• TS2 may have a more general developmental role in most grasses than programmed cell death of the developing gynoecium, but has been co-opted to this role within a subset of Poaceae, probably as a result of alterations in the activity or regulation of other genes in the gynoecial pathway.

Key words: evolution, gibberellin, grass, tasselseed, unisexual flower, *Zea*.

New Phytologist (2006) **170**: 885–899

doi: 10.1111/j.1469-8137.2006.01726.x

Introduction

Flowering plants exhibit a diverse array of sexual systems (reviewed in Barrett, 2002; Tanurdzic & Banks, 2004). Although most species produce only bisexual flowers, with functional staminate (pollen-producing, often called male) and pistillate (ovule-producing, often called female) organs in the same flower, an estimated 30% of species produce at least some unisexual flowers (Richards, 1997). These unisexual species are distributed throughout the angiosperm phylogeny, indicating that dicliny (the presence of unisexual flowers) has evolved repeatedly, likely through diverse mechanisms (Charlesworth, 1985; Mitchell & Diggle, 2005).

Evolution of sex expression in grasses

Transitions between bisexual and unisexual flowers have occurred frequently in the 70 million year history of the Poaceae. The grass outgroups in the graminoid Poales include a mixture of unisexual and bisexual taxa (Stevens, 2001 (onwards)). The earliest diverging member of the graminoids, *Flagellaria*, is bisexual, the next branch leads to Restionaceae and Anarthriaceae, which are both unisexual, and the next branch leads to *Joinvillea*, with bisexual flowers. The sister to Poaceae is now thought to be the western Australian family Ecdeiocoleaceae (Michelangeli *et al.*, 2003), in which flowers are exclusively unisexual (Rudall *et al.*, 2005). The earliest branch of the Poaceae includes *Anomochloa* and *Streptochaeta*, both of which are bisexual, but the next lineage leads to present-day Pharoideae, in which all species have exclusively unisexual flowers. These shifts between different forms of sex expression have continued throughout the history of the family, and estimation of the ancestral condition for the family is highly sensitive to assumptions about branch lengths and transition probabilities (data not shown). Even among closely related species, sex expression is highly variable; at least six transitions in sex expression have

been estimated in the chloridoid genus *Bouteloua* alone (Columbus, 1999; Kinney *et al.*, 2003). Phylogenetic data thus show that changes in sex expression occur frequently, and gain or loss of unisexuality is apparently common.

It is convenient to describe grass taxa as being unisexual or bisexual, but this oversimplifies the variation in sex expression, and indeed almost all possible breeding systems can be found in the grasses (Connor, 1981; Chapman, 1990; Watson & Dallwitz, 1992). Dioecy (staminate and pistillate flowers on separate plants) occurs in subfamilies Chloridoideae, Pooideae, and Danthonioideae; in the latter two subfamilies it is often associated with apomixis. Monoecy also occurs, with the most familiar example being maize (*Zea mays*). Among monoecious species, the staminate and pistillate flowers may be borne in separate inflorescences (e.g. maize), in separate spikelets in the same inflorescence (e.g. Pharoideae, tribe Olyreae in the Bambusoideae, and *Tripsacum*, *Coix*, and *Heteropogon* in Panicoideae), or in separate flowers in the same spikelet (e.g. *Ixophorus unisetus*). Other grasses are andromonoecious, with bisexual and staminate flowers in the same spikelet (e.g. most members of Panicoideae). Gynomonoecy, with bisexual and pistillate flowers in the same spikelet, occurs sporadically among the grasses (Connor, 1981); generally spikelets contain multiple flowers, the lower of which are bisexual and the upper of which often produce only a gynoecium.

Sterile flowers are also common and may co-occur with bisexual or unisexual flowers. In some cases these appear to be simply underdeveloped, perhaps as a result of environmental conditions. In other cases, sterility is clearly genetically and developmentally imposed. Examples of the latter include the lower two flowers of the spikelet in *Oryza sativa*. These are reduced to empty lemmas, and often called 'glumes' in the literature; the true glumes are reduced to tiny flaps, sometimes known as rudimentary glumes.

In all studied species of Poaceae and their relatives, flowers initiate both gynoecia and androecia (Cheng *et al.*, 1983; Le Roux & Kellogg, 1999; Zaitchik *et al.*, 2000; Rudall *et al.*, 2005) even if the mature flower will be unisexual (Fig. 1).

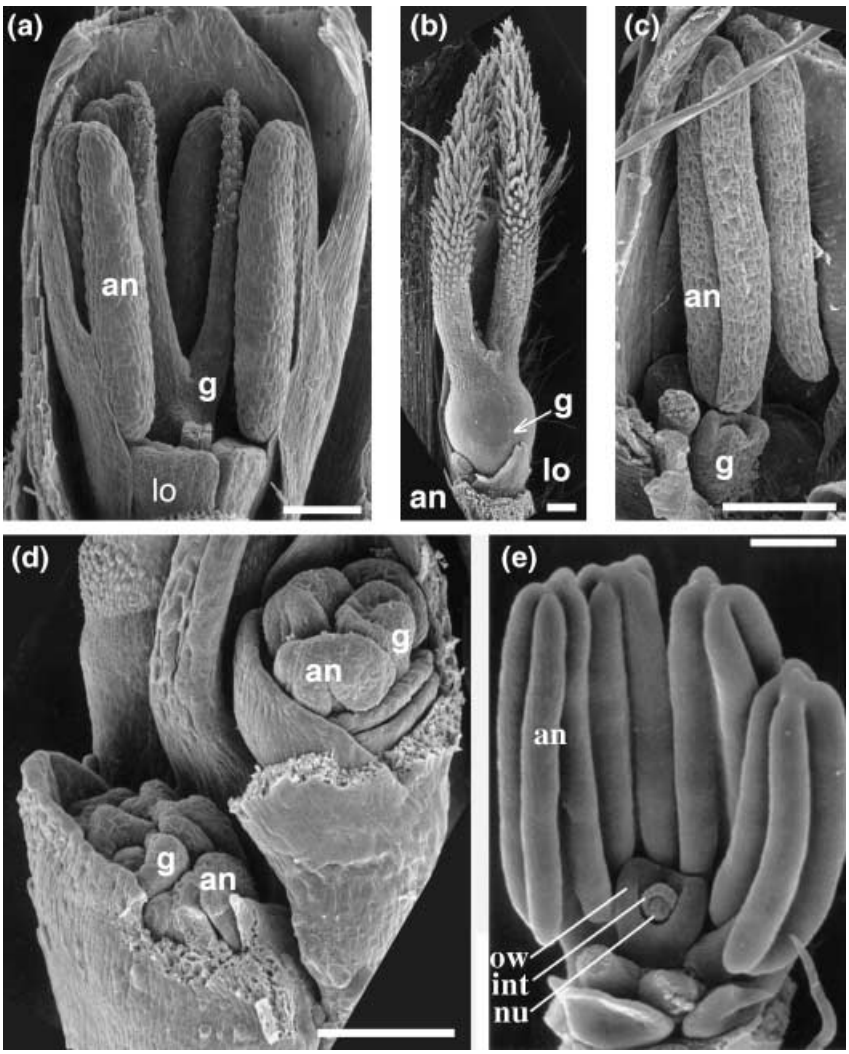

Fig. 1 Scanning electron micrographs of developing flowers in selected grasses. (a) A bisexual flower in *Coelorachis aurita*. Reproduced from Le Roux & Kellogg (1999), with permission from the Botanical Society of America; (b) a pistillate flower, showing undeveloped anthers, in *Heteropogon contortus*; (c) a staminate flower, showing gynoecium arrested at the gynoecial ridge stage, in *Hyparrhenia hirta*; (d) adjacent flowers in the inflorescence of *H. hirta*; the lower flower is in a sessile spikelet and will become bisexual; the upper flower is in a pedicellate spikelet and will become staminate; (e) a staminate flower of *Zizania aquatica*, showing well-developed gynoecium with ovule and integument. Reproduced from Zaitchik *et al.* (2000), with permission from the International Journal of Plant Sciences. Bars, 100 μm. an, anther; g, gynoecium; int, integument; lo, lodicule; nu, nucellus; ow, ovary wall.

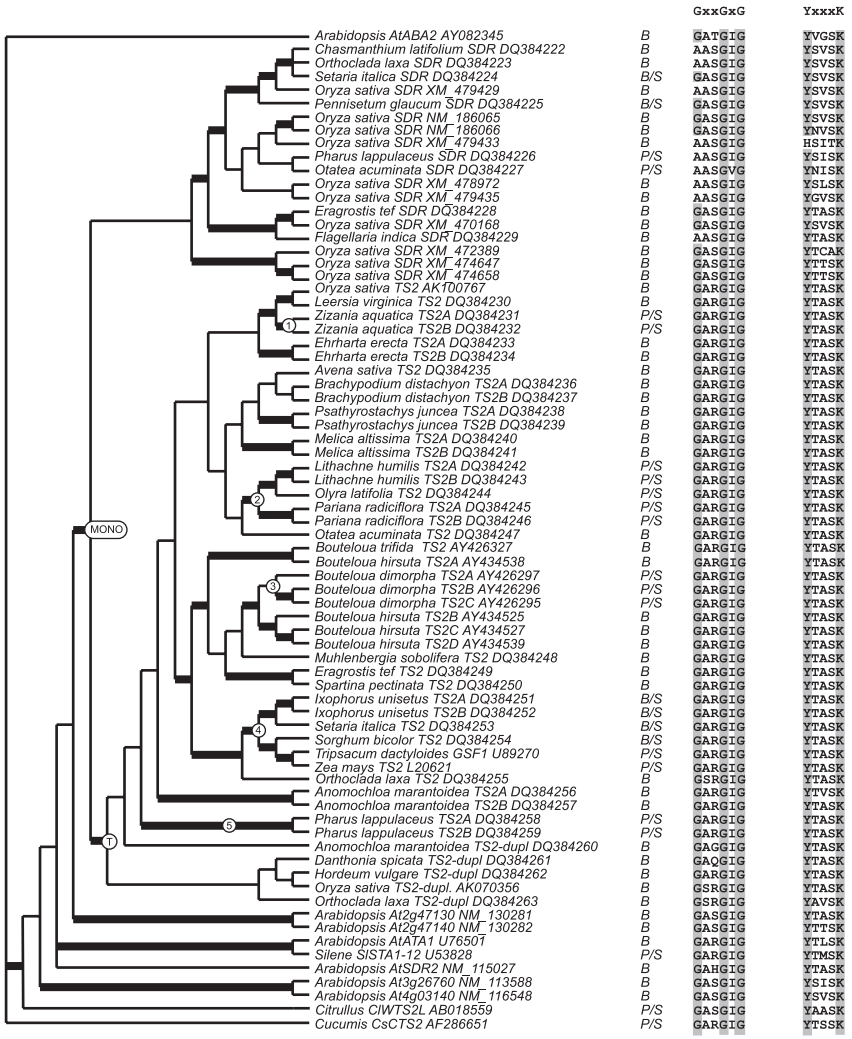

Fig. 2 Best cladogram of relationships of 74 *TASSELSEED2* (*TS2*) and related short-chain dehydrogenase/reductase (SDR) genes from the Bayesian analysis using the General time-reversible (GTR) model, invariant sites and gamma distributed rates (GTR + I + Γ). Bold branches are supported by clade credibility ≥ 0.95. B, bisexual flowers; MONO, monocots; P, pistillate flowers; S, staminate flowers; T, *TS2*/*TS2-duplicate* clade. 1–5, branches where unisexuality is inferred to have evolved; GxGxxG, NAD- or NADP-binding site; YxxxK, catalysis or subunit interaction motif.

Staminate flowers form when the gynoecium fails to develop fully, and pistillate flowers when the androecium fails to mature. This pattern of selective organ abortion is also found in unisexual flowers in many other families, including Actinidiaceae, Asparagaceae, Caryophyllaceae, Curcubitaceae, Fabaceae and Polygonaceae (reviewed in Ainsworth, 2000). Control of sex expression is thus frequently the control of floral organ development, stopping or starting an existing, functional developmental pathway.

We hypothesize that there is no single sex-determining pathway in the grasses, but rather that sex expression has been modified in different ways in different lineages. The few studies that have looked at development of unisexual flowers in the grasses support this hypothesis. In staminate flowers of multiple species of panicoid grasses (most of which have staminate and bisexual flowers in the same spikelet), subepidermal cells in the gynoecium undergo cell death soon after formation of the gynoecial ridge (Fig. 1c), a pattern that may characterize the entire clade of approximately 3300 species (Li *et al.*, 1997; Calderon-Urrea & Dellaporta, 1999; Le Roux & Kellogg, 1999; Zaitchik *et al.*, 2000). In staminate flowers in maize, the cells that die in the gynoecium become vacuolized, and lose free ribosomes and other organelles from the cytoplasm (Cheng *et al.*, 1983). However, in wild rice (*Zizania aquatica*, subfamily Ehrhartoideae, branch 1; Fig. 2), the gynoecium of staminate flowers develops much farther than those in Panicoideae. The stigmatic arms enlarge, the ovule differentiates, and the integuments become visible (Fig. 1e). Growth finally ceases at the stage at which the carpels have nearly closed; developmental arrest of the gynoecium correlates with deposition of dark-staining material in the ovary walls (Zaitchik *et al.*, 2000).

We also hypothesize that development of androecia and gynoecia should be under separate genetic control. Because developmental arrest of the two sorts of organs does not covary in evolutionary time, or even necessarily within the same plant, we conclude that production of a unisexual staminate flower proceeds via a different mechanism from production of a unisexual pistillate flower. They are not, in other words, simply two sides of the same coin.

Genetic basis of dicliny in grasses

The genetic basis of dicliny is best understood in maize, a monoecious plant with staminate flowers borne in apical tassel inflorescences and pistillate flowers in axillary ear inflorescences. In *anther ear1* (*an1*) and *dwarf plant1* (*d1*), *d2*, *d3*, *d5*, and *d8* mutants, stamens within the ear florets fail to abort, resulting in a bisexual upper floret and staminate lower floret. Staminate florets in the tassel are unaffected in these mutants. These mutations all affect gibberellin (GA) biosynthesis, indicating that GAs play a key role in the stamen abortion process in ear florets (Phinney, 1961, 1984).

The mutant phenotype of the ear is a reversion to the ancestral state. The genus *Zea* and its sister *Tripsacum* were derived from ancestors that had bisexual upper florets and staminate lower florets, the latter pattern being synapomorphic for Panicoideae (Giussani *et al.*, 2001). Thus the origin of one-flowered, pistillate florets in *Zea* and *Tripsacum* could have been a response to altered regulation of GA concentrations.

The GA data also confirm that development of different floral organs can be completely decoupled. In the *an1* and *dwarf* mutants, gynoecium development is normal and only anther development is affected.

Multiple loci in maize affect gynoecium development, particularly in the tassel (Coe *et al.*, 1988; Irish *et al.*, 1994). Most of these have pleiotropic effects, affecting aspects of inflorescence branching as well as development of the plant as a whole. However, *tasselseed* (*ts*) *1*, *ts2*, and *Ts5* mutants, designated class I *tasselseed* mutants by Irish *et al.* (1994), affect only gynoecial development in the tassel and lower floret development in the ear, without affecting other aspects of floral, inflorescence or plant development. In *ts1*, *ts2* and *Ts5* mutants, sex expression in the tassel is reversed, with the gynoecium developing and stamens being suppressed. In addition, in *silkless1* (*sk1*) mutants the pistils of the ear florets fail to develop, resulting in plants that cannot produce seeds. Staminate florets in the tassel are unaffected in *sk1* mutants. Because they affect sex determination specifically and with minimal pleiotropy, *TS1*, *TS2*, *TS5*, and *SK1* are good candidates for genes that might have been modified in the evolution of different patterns of sex expression in the grasses.

Double mutant analysis has suggested that *TS2* acts in the same pathway as *TS1* and *TS5* (Irish *et al.*, 1994) and *SK1* (Jones, 1934). In *ts1*; *ts2* and *ts1*; *Ts5* double mutants the mutant phenotype resembles the *ts1* single mutants, suggesting that *TS1* may be upstream of both *TS2* and *TS5* (Irish *et al.*, 1994). Likewise, *ts2*; *Ts5* double mutants resemble *ts2* single mutants, placing *TS2* genetically upstream of *TS5*. In *ts2*; *sk1* double mutants, pistils developed in all ear florets and a mixture of pistillate and staminate florets developed in the tassel. Thus the lack of TS2 gene product suppressed the *sk1* mutant phenotype in the ear and lack of SK1 gene product was partially able to correct the *ts2* mutant phenotype in the tassel (Jones, 1934; Irish *et al.*, 1994).

Of the class I *tasselseed* loci, only *TS2* has been cloned (DeLong *et al.*, 1993). The gene product resembles a short-chain dehydrogenase/reductase (SDR), with significant similarity to hydroxysteroid dehydrogenases, and is hypothesized to play a role in hormone metabolism (DeLong *et al.*, 1993). Expression of *TS2* RNA in would-be staminate flowers correlates with death of cells – presumably the ones in which it is expressed – which lose their cytoplasm; death of these cells presumably prevents continued development of the gynoecium. Calderon-Urrea & Dellaporta (1999) showed that *TS2* RNA is expressed in pistils of both pistillate and staminate flowers, although only the pistils of staminate flowers undergo cell death. They hypothesized that the pistil primordium of the pistillate (ear) inflorescence is protected from *TS2*-mediated cell death by presence of a wild-type *SK1* gene, perhaps by formation of a TS2:SK1 protein complex, by disruption of a downstream step in the cell-death pathway, or by sequestration of TS2 or other cell death factors to an organelle compartment (Calderon-Urrea & Dellaporta, 1999). Alternatively, Veit *et al.* (1993) proposed that absence of TS2 protein in the *ts2* mutant resulted in ectopic expression of *SK1* in the tassel, leading to feminization. Explicit testing of these hypotheses awaits cloning and characterization of the *SK1* gene.

The *TS2* ortholog in gama grass (*Tripsacum dactyloides*, tribe Andropogoneae, subfamily Panicoideae), *GYNOMONOECIOUS SEX FORM1* (*GSF1*), plays a similar role to *TS2* in maize (Li *et al.*, 1997). *TS2*-like SDR genes have also been isolated from *A. thaliana* (*AtATA1*) and *Silene latifolia* (*SlSTA1–12*), where expression is restricted to tapetal cells, suggesting that these *TS2*-like SDR genes are not involved in sex determination (Lebel-Hardenack *et al.*, 1997). However, the expression of these genes in tapetal tissue, which breaks down during pollen formation, suggests a more general role in cell death (Lebel-Hardenack *et al.*, 1997). The precise relationship of the *A. thaliana and Silene* genes to *TS2* is not clear, but could easily be determined with additional sequences similar to those of both the dicot and monocot genes.

If TS2 is a major sex-determining protein, then comparative sequencing of *TS2* orthologs might be expected to show evidence of positive selection in at least some lineages of grasses. Such tests have only been carried out in two closely related species of *Bouteloua*. In the dioecious *Bouteloua dimorpha*, *TS2* appears to be evolving neutrally, whereas in the bisexual species *Bouteloua hirsuta*, *TS2* is under strong purifying selection (Kinney *et al.*, 2003). *Bouteloua* is in the grass subfamily Chloridoideae, and thus is distantly related to subfamily Panicoideae, which includes maize. It is unknown whether gynoecial abortion in *Bouteloua* correlates with death of cells in the subepidermal layers, and the expression pattern of *TS2* is also unknown.

To evaluate the hypothesis that *TS2* is a major player in the evolution of unisexual flowers in grasses, we analyzed its molecular evolution and expression in a variety of grasses.

Materials and Methods

Generation and analysis of sequences

Seventy-four *TS2* and *TS2*-like short-chain dehydrogenase/reductase (SDR) genes from 34 eudicots and monocots were examined, including the putative *TS2* homologs in *Silene latifolia* Poir. (*SlSTA1–12*) and *Arabidopsis thaliana* (L.) Heynh. (*AtATA1* and *AtSDR2*), and diverse unisexual and bisexual grasses. Forty-one genes were isolated for this study and 33 were identified by BLAST (Altschul *et al.*, 1997) searches at National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/). Degenerate PCR primers to amplify diverse *TS2-like* SDR genes were designed using PRIMACLADE (Gadberry *et al.*, 2005) based on the *A. thaliana ATA1*, *S. latifolia* Poir. SlSTA1–12, *Tripsacum dactyloides* (L.) L. *TdGSF1* and *Zea mays* L. *ZmTS2* gene sequences.

Total DNA was isolated using an sodium dodecyl sulphate (SDS) protocol (Dellaporta, 1994), and double-stranded *TS2-like* SDR PCR products were amplified using a standard reaction mix [2 U *Taq* polymerase (Promega Corp., Madison, WI, USA), 5 μl of 10 × reaction buffer, 5 μl of 25 mM $MgCl_2$, and 2 μl of 2.5 mM dNTP], plus 20 μM each of the primers *TS2-59F*, 5′-AGA GGC TGG AMG GGA AGG TG-3′ and *TS2–854R*, 5′-GTC SAC GAC RAG GTT GTG GC-3′, 10% (by volume) 5 M betaine (Sigma, St Louis, MO, USA), 5% dimethyl sulfoxide (DMSO) (by volume), and 100–200 ng of genomic DNA. PCR reactions used a hot-start, touchdown PCR profile (three cycles at 65°C, three at 63°C, three at 60°C and 25 at 57°C). PCR fragments were purified and subcloned as described by Malcomber & Kellogg (2004), and two to five clones were sequenced per species. Dideoxy sequencing used plasmid primers *T7* and *SP6*, with reactions analyzed on ABI 377, ABI 3100 or ABI 3130XL DNA sequencers (Applied Biosystems, Foster City, CA, USA). Only nucleotide sequences with scores > 20, as determined by PHRED, (Ewing *et al.*, 1998) were used in subsequent analyses. Alignments were edited in Seqman II (DNASTAR Inc., Madison, WI, USA) and all sequences submitted to GenBank (DQ384222–DQ384263).

Nucleotide sequences were aligned based on the conceptual amino acid translation using REVTRANS (Wernersson & Pedersen, 2003), and adjusted manually using MACCLADE 4 (Maddison & Maddison, 2003); nucleotides were used for all analyses. Maximum likelihood (ML) and Bayesian phylogenetic analyses used PAUP* 4.0 (Swofford, 2000) and MRBAYES 3.1 (Huelsenbeck & Ronquist, 2001) on the Beowulf parallel processing cluster at the University of Missouri – St Louis. The ML search used 10 separate heuristic searches with TBR and MULPARS on and 10 random sequence additions. Bayesian analyses used two separate searches of 8 million generations using default flat priors and the General time-reversible (GTR) model of sequence evolution, invariant sites and gamma distributed rates (GTR + I + Γ) (estimated by MODELTEST; Posada & Crandall, 1998). 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 (CC) values and the 95% credible set of trees were estimated using MRBAYES (Huelsenbeck & Ronquist, 2001).

Copy number of genes was estimated using PCR and confirmed for some taxa by Southern blots. Approximately 10 μg of total DNA was digested with *Bam*HI, *Eco*RI, or *Hin*dIII, separated on a 1.2% agarose gel, blotted onto a nylon membrane, and hybridized with a full-length ^{32}P-dCTP-labeled *TS2* probe for 16 h at 65°C following Laurie *et al.* (1993). After hybridization, blots were washed at 65°C twice in 2 × saline sodium citrate (SSC) (1 × SSC is 0.15 M NaCl and 0.015 M sodium citrate)/0.5% sodium dodecyl sulfate (SDS) for 20 min each time, and twice in 0.1 × SSC/0.1% SDS.

Molecular evolution

Tests for selection on *TS2* and related SDR genes used the best ML topology and the CODEML program within PAML (Yang, 1997) version 3.14 on the Beowulf parallel processing cluster at University of Missouri – St Louis. Evidence of positive selection at particular codons was tested using the nested codon models M0 and M3, M1a and M2a, and M7 and M8 (Yang & Nielsen, 2002; Wong *et al.*, 2004; Yang *et al.*, 2005); the standard likelihood ratio test (LRT) statistic was applied against a χ^2 distribution with two degrees of freedom. Model M0 is the simplest codon model with a single ω parameter (ratio of nonsynonymous to synonymous sites, or d*N*/d*S*) for all sites and all branches of the phylogeny. M1a is the 'nearly neutral' model with two site classes, $0 < \omega_0 < 1$ and $\omega = 1$ (Wong *et al.*, 2004). M2a is the selection model and is an extension of M1a, and has the additional class ω_2, which can take any value. M3, the discrete model, is an extension of M0 and has *k* site classes, each with a separate ω ratio. M3 with two sites classes is the null model for the model B branch-site model described in the following paragraph. M7 and M8 use a discrete β distribution to approximate among-site ω variation. M8 differs from M7 in allowing an additional site class with $\omega > 1$.

We also tested for sites potentially under positive selection on the five branches within the *TS2* clade leading to grasses with unisexual flowers (Fig. 2) using the modified branch-site models A and B (Wong *et al.*, 2004; Yang *et al.*, 2005). Model A was compared with model M1a (NearlyNeutral) and model B was compared with M3 (discrete) with two site classes in a LRT against a χ^2 distribution with two degrees of freedom (Wong *et al.*, 2004; Yang *et al.*, 2005).

TS2 protein structure modeling

The tertiary structure of the TS2 protein was estimated via homology modeling using the ESyPred3D server (Lambert

et al., 2002). The proximity of sites that were identified as being potentially under positive selection to the NAD/NADP binding site (GxxGxG) and catalysis or subunit interaction site (YxxxK) was investigated using PROTEIN EXPLORER 2.75 (Martz, 2002).

Gene expression

Total RNA was extracted from developing rice inflorescences, culms, young leaves and roots using RNAwiz solution (Ambion, Austin, TX, USA) according to the manufacturer's instructions. *TS2* expression profiles were inferred using the Superscript One-step RT-PCR kit (Invitrogen, Carlsbad, CA, USA) following the manufacturer's instructions, except that 4% [volume/volume (v/v)] DMSO and 10% (v/v) 5 M betaine (Sigma) were included to facilitate strand separation. *TS2*-specific PCR primers were designed using PRIMACLADE (Gadberry *et al.*, 2005). *Oryza sativa* L. *TS2* (*OsTS2*) gene fragments were amplified with a 55°C annealing temperature using the primers *OsTS2-828F*, 5′-GAA GAT GGA GGA GGT GGT CA-3′, and *OsTS2-1110R*, 5′-AGT CCA ATT AAC ACA TTG AAT CAA GA-3′. *Zea mays TS2* (*ZmTS2*) and *Sorghum bicolor* (L.) Moench *TS2* (*SbTS2*) gene fragments were both amplified with a 57°C annealing temperature using the primers *ZmTS2-915F*, 5′-GAG GTG GAG AAG ATG GAG GAG-3′ and *ZmTS2-1196R*, 5′-ATG AAT CAA TCA ACC AAA TGA AAA-3′ (for maize), and *SbTS2-43F*, 5′-ATC GTG MGK CTG TTC GTG A-3′, and *TS2-854R* (for sorghum). All RT-PCR products were verified by subcloning and sequencing as previously described. The amplified *TS2* fragment lacks introns, so that the cDNA and gDNA PCR products are expected to be the same size. We therefore included triose phosphate isomerase (*TPI*), which has introns, as a positive control to verify that none of the RNA extractions was contaminated by gDNA. Exons four and five of *TPI* were PCR-amplified using the same conditions using degenerate primers *TPIX4F*, 5′-AAG GTC ATT GCA TGT GTT GG-3′, and *TPIX6R*, 5′-TTT ACC AGT TCC AAT AGC CA-3′ (Strand *et al.*, 1997), which span a 500–800-bp intron. If RNA were contaminated with DNA we would see two bands in the RT-PCR reaction.

RNA *in situ* hybridization was conducted on developing rice inflorescences using 3′ untranslated region (UTR) probes derived from RT-PCR products using the *OsTS2-828F* and *OsTS2-1110R* primers, as described in Malcomber & Kellogg (2004).

Results

TS2 is single-copy in the grasses, but has an ancient duplicate

PCR fragments ranged from 710 bp (*Flagellaria indica* L.) to 804 bp (*Hordeum vulgare* L.) and represent ~79% of the *Z. mays TS2* coding region. Alignment of the predicted amino acid sequences was largely unambiguous, except for a hypervariable region between amino acids 225 and 242 that was removed from subsequent analyses (Fig. 3). Interestingly, this region represents an apparent insertion in members of the Bambusoideae, Ehrhartoideae, and Pooideae, which together form a clade (the BEP clade) in some phylogenetic analyses (Grass Phylogeny Working Group, 2001).

The best ML tree for the 75 SDR gene dataset (–ln 17592.537) was identical to the tree with the highest posterior probability from the Bayesian analysis, and estimated a well-supported (≥ 0.95 clade credibility) clade of monocot SDR genes, and several well-supported subclades (Fig. 2). The 95% credible set of topologies included 15969 distinct tree topologies, indicating that the SDR dataset has only limited resolving power.

The *Z. mays TS2* (*ZmTS2*) sequence falls within a well-supported clade that we will refer to as the *TS2*/*TS2-duplicate* clade (T; Fig. 2). The *TS2*/*TS2-duplicate* clade contains a broad sample of grasses, including a representative of the earliest diverging lineage, *Anomochloa*, indicating that the duplication event producing the two lineages occurred before the origin of extant grasses. Sister to the *TS2*/*TS2-duplicate* clade is a well-supported clade of SDR genes from several grasses and *F. indica*, a member of the Poales. The presence of 10 distinct rice sequences within this clade indicates that multiple duplication events occurred during diversification of this lineage, although current sampling limits our ability to infer where within monocots the duplication events occurred. The *A. thaliana* genes *At2g47140* and *At2g47130* are inferred to be sister to the clade of monocot SDR genes, although this relationship is not well supported. Other closely related *A. thaliana* SDR genes include *At3g51680* (= *AtSDR2*; Cheng *et al.*, 2002) and *AtATA1* (Fig. 2). However, the complex pattern of duplication of SDR genes within monocots means that none of the *A. thaliana* genes is orthologous to *ZmTS2*.

Estimated relationships within the *TS2* clade by the Bayesian and ML analyses are largely congruent with the Grass Phylogeny Working Group (2001) analysis, although the placement of *Pharus* and *Anomochloa* is reversed in the *TS2* phylogeny relative to the inferred organismal phylogeny. However, this placement in the *TS2* topology is not well supported.

We isolated up to two distinct *TS2* sequences per species (A-B; Fig. 2) from sequencing two to five clones. Kinney *et al.* (2003) reported up to 20 *TS2* alleles in their analysis of *Bouteloua dimorpha* J.T. Columbus and *Bouteloua hirsuta* Lag. All *TS2* sequences from the same species coalesce in our analysis except for *B. hirsuta* (Fig. 2), with raw divergence among clones ranging from 0.005 (*Ehrharta erecta* Lam.) to 0.118 (*Brachypodium distachyon* (L.) P. Beauv.). Both sequenced clones in *B. distachyon* have an open reading frame, suggesting that they are functional. *B. distachyon* ($n = 5$) may be diploid, tetraploid, or hexaploid (Index of Plant Chromosome Numbers; http://mobot.mobot.org/W3T/Search/ipcn.html), indicating that

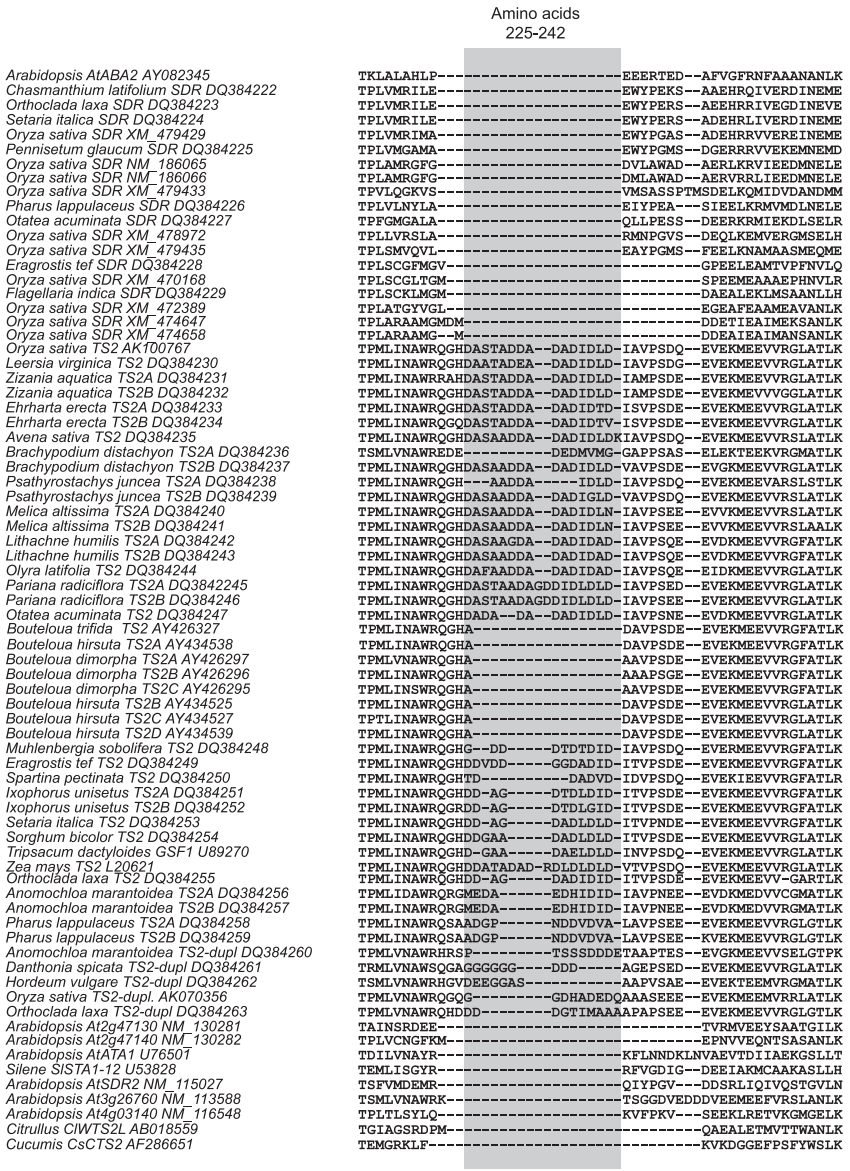

Fig. 3 C-terminal amino acid alignment of 74 TASSELSEED2 (TS2)-like SDR proteins included in the analysis, with the hypervariable region between amino acids 225 and 242 that was removed from subsequent analyses shaded gray.

the high divergence between the sequenced colonies may reflect different *TS2* copies within our polyploid accession. The minimal variation detected among *TS2* clones in the other species is likely caused by PCR error, although some of the variation could be caused by persistence of ancestral polymorphism (lineage sorting). One *B. hirsuta* clone forms a clade with *Bouteloua trifida* Thurb. ex S. Watson which is sister to a clade of *TS2* sequences from diverse chloridoid grasses (*Bouteloua*, *Eragrostis*, *Muhlenbergia*, and *Spartina*), suggesting a duplication near the base of subfamily Chloridoideae or possibly deeper within the PACCAD clade.

We verified the copy number of the *TS2-like* genes in grasses by Southern blots using species-specific *TS2* probes in the diploid species *O. sativa*, *Pharus lappulaceus* Aubl. and *S. bicolor* (Index of Plant Chromosome Numbers; http://mobot.mobot.org/W3T/Search/ipcn.html). Only one gene was detected on the *O. sativa* and *S. bicolor* blots, whereas a faint second copy was visible on the *P. lappulaceus* blot (Fig. 4). Uncorrected sequence divergence between the rice *TS2* and *TS2-duplicate* genes was 0.206, so the *TS2* probe was unlikely to hybridize to the *TS2-duplicate* using our high-stringency hybridization conditions. In contrast, the faint extra band in the *Hin*dIII lane of the *P. lappulaceus* blot could represent the *TS2-duplicate* gene. Sequence divergence between *TS2* and the *TS2-duplicate* gene in *Anomochloa*, the other early diverging grass included within our analysis, was 0.18207, just within the limits of detection on a high-stringency DNA gel blot. *Hin*dIII does not cut within the *P. lappulaceus TS2* gene fragment and the low sequence divergence between the two distinct *P. lappulaceus TS2* clones (0.0078) included in our analysis is more likely to have been a result of *Taq* error than different gene copies.

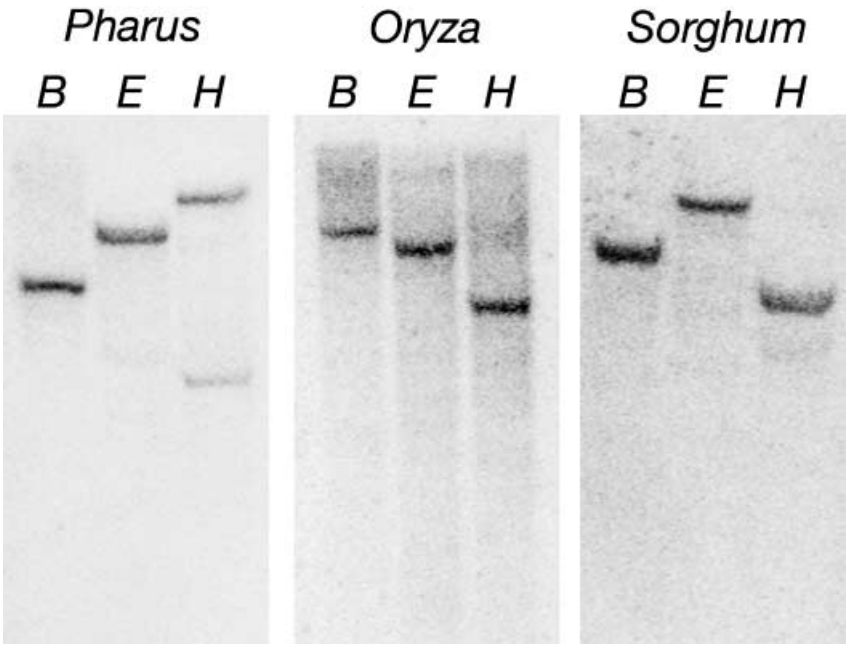

Fig. 4 DNA gel blots of *Oryza sativa*, *Pharus lappulaceus* and *Sorghum bicolor*. Blots were probed with species-specific *TASSELSEED2* (*TS2*) probes. Genomic DNA was digested with *Bam*HI (*B*), *Eco*RI (*E*) and *Hin*dIII (*H*) restriction enzymes.

The sequence of *TS2* is highly conserved and under strong purifying selection

All SDR enzymes have an N-terminal NAD- or NADP coenzyme-binding pattern of GxxGxG and a motif considered to be involved in catalysis or subunit interaction of YxxxK (Jörnvall *et al.*, 1999). The GxxGxG motif is conserved as GARGIG in all *TS2* and *TS2-like* sequences, except that the *O. sativa TS2-duplicate* and *Orthoclada laxa* (Rich.) P. Beauv. *TS2* and *TS2-duplicate* sequences have Ser instead of Ala, the *Danthonia spicata* (L.) P. Beauv. ex Roem. & Schult. *TS2-duplicate* sequence has Gln instead of Arg, and the *Anomochloa marantoidea* Brongn. *TS2-duplicate* sequence has Gly instead of Arg (C to G at position 1). The YxxxK motif is conserved as YTASK in all *TS2* and *TS2-duplicate* sequences except that *A. marantoidea TS2A* and *O. laxa TS2-duplicate* both have Val instead of Ala (a conservative substitution).

Despite widespread conservation of the GxxGxG binding site in all other SDR genes (Jörnvall *et al.*, 1999), several monocot SDR genes have an Ala instead of a Gly at the first position of the NAD/NADP binding domain (Fig. 2). This amino acid change might affect their ability to bind NAD/NADP, although biochemical analyses have not been conducted on these proteins. In monocot SDR genes where the GxxGxG and YxxxK motifs are conserved, they are generally not GARGIG and YTASK, also indicating possible diversification of biochemical function.

All site-specific models indicate that *TS2* sequences were under strong purifying selection ($\omega << 1.0$) with no evidence for positive selection having acted on sites across the whole phylogeny. The one-ratio model (M0) estimated $\omega = 0.0934$. The discrete model (M3) was significantly better than the one-ratio model ($P < 0.001$), indicating considerable variability in ω among sites, but also identified no positively selected sites. The nearly neutral model (model M1a) allowed two ratios, $0 < \omega_0 < 1$ and $\omega_1 = 1$, and estimated 97.1% of sites to be under considerable purifying selection ($\omega_0 = 0.029$). The more complex models also failed to find any evidence of positive, or even relaxed, selection.

Using branch-site models A and B, we tested for evidence of positive section on the branches where unisexuality is inferred to have evolved. Our species sample included five such branches: (1) *Zizania aquatica* (L.), (2) the herbaceous bamboo clade (*Lithachne*, *Olyra* and *Pariana*), (3) *B. dimorpha*, (4) the panicoid clade (*Ixophorus*, *Setaria*, *Sorghum*, *Tripsacum* and *Zea*), and (5) *P. lappulaceus* (1–5; Figs 2 and 5). Testing all five unisexual branches simultaneously, we found ω_2 greater than 1 in model A ($\omega_2 = 1.796$), in which two sites were identified as potentially being under positive selection (Ala^{29} to Leu and Arg^{117} to Gly); however, neither site had a significant posterior probability (0.507 and 0.503, respectively) and the model A likelihood score was not significantly better than that of the null model. Positive selection was also inferred in model B ($\omega_2 = 1.418$), and the likelihood score was significantly better than that of the null model, but no sites were identified as potentially under positive selection.

Tests for positive selection on each branch individually identified the branch subtending the herbaceous bamboo clade (*Lithachne*, *Otatea* and *Pariana*) (branch 2; Fig. 2). Under both models (A and B) the foreground parameter ω_2 was large and likelihood scores were significantly better than that of the null model. However, only in model A was a site potentially

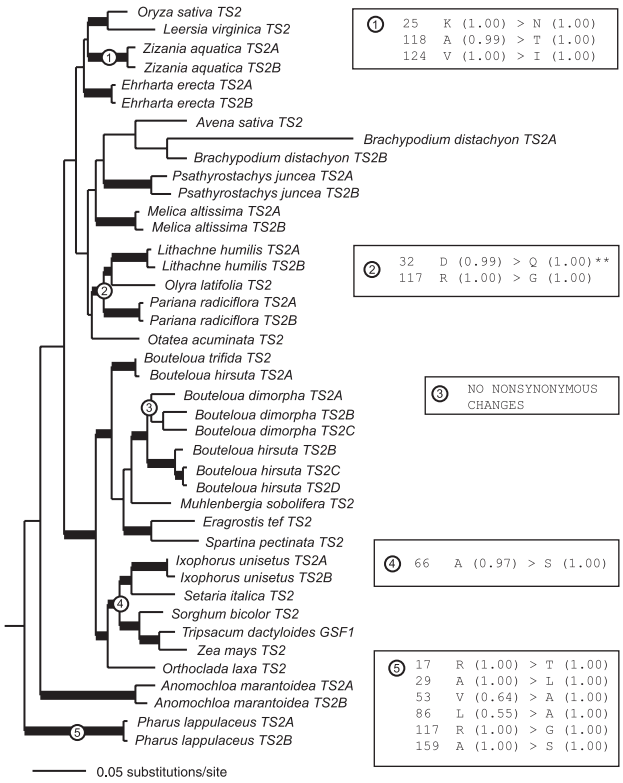

Fig. 5 Maximum likelihood (ML) phylogram of the *TASSELSEED2* (*TS2*) clade with nonsynonymous changes on each of the five branches (1–5) where unisexuality is inferred to have evolved as estimated by PAML (Yang, 1997). The number represents the amino acid position within the TS2 protein followed by the estimated change and the likelihood in parentheses.

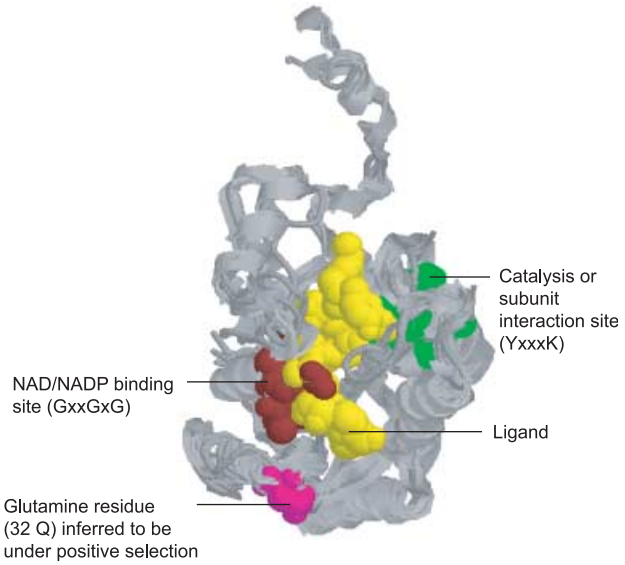

Fig. 6 Tertiary structure of the *Lithachne humilis* TASSELSEED2 (TS2) protein with NAD, NADP or NADPH (GxGxxG), subunit interaction or catalysis (YxxxK), inferred ligand position, and site potentially under positive selection indicated.

under positive selection identified with a significant posterior probability, with replacement of Asp^{32} (D) by Gln (Q) (P = 0.972). ω was also greater than 1 in analyses of the branch subtending the *Pharus* clade (branch 5), but only model B was significantly better than the site model and failed to identify any sites as being potentially under positive selection (Table 1). Model A identified two sites as being potentially under positive selection (Ala^{29} to Leu and Leu^{86} to Ala), but neither change had posterior probability ≥ 0.95 and the likelihood score for the branch site model was not significantly better than that for the site model (Table 1).

The Asp^{32} to Gln substitution on the branch subtending the herbaceous bambusoid clade requires substitutions at the first and third positions of the codon, and changes the charge of the residue. The tertiary structure of the *Lithachne humilis* Soderstrom TS2 protein shows that position 32 is not near either known active site (Fig. 6), although the Asp to Gln change could affect other as yet unidentified active or binding sites within the protein.

Other amino acid changes occur in parallel, and do not correlate with unisexuality. Ala^{66} changed to Ser along the unisexual panicoid branch, on the branch subtending the pooid clade (*Avena*, *Brachypodium*, *Psathyrostachys* and *Melica*) and on the branch leading to the chloridoid grass *Muhlenbergia*. Similarly, three nonsynonymous changes were estimated on the *Zea*/*Tripsacum* branch: Val^{20} to Ala, Leu^{112} to Arg, and Iso^{162} to Val. Identical changes at positions 20 and 112 occur on the branch subtending the *B. hirsuta* clade, and the change from Iso^{162} to Val was also inferred for the branches subtending the *Muhlenbergia*, *Spartina* and *Anomochloa* TS2 sequences. All these changes are located on the periphery of the TS2 protein, and none appears to affect either of the known active sites.

TS2 expression is not restricted to flowers

Models of *TS2* function imply that its primary role should be in flowers, and predict that it should be expressed primarily in inflorescences. We were unable to find any reports in the literature, however, that tested this assumption, and therefore performed the relevant RT-PCR with *TS2*-specific primers on inflorescence, culm, leaf and root RNA from *O. sativa*, *S. bicolor* and *Z. mays*. *Oryza sativa* has spikelets with one flower bisexual and fertile, and two strongly reduced and sterile; the latter do not initiate either gynoecium or androecium. If the primary role of *TS2* is to suppress gynoecium development, we would expect no expression in *O. sativa*. *Sorghum bicolor* has paired pedicellate and sessile spikelets, each with two flowers; the lower flower is always sterile and, like sterile flowers in rice, does not initiate floral organs. The upper floret of the sessile spikelet is bisexual, whereas that of the pedicellate spikelet is staminate or sterile. We expected some *TS2* expression in

New Phytologist

Table 1 Codon model parameter estimates for *TASSELSEED2* (*TS2*)

Model	*p*	ln	2Δ ln	Parameter estimates
M0 (one ratio)	1	−19206.608		$\omega = 0.0934$
Site models				
M1a (nearly neutral)	2	−19170.500		$p_0 = 0.971$ ($p_1 = 0.029$) $\omega_0 = 0.1259$
M2a (positive selection)	4	−19170.500	0	$p_0 = 0.968$, $p_1 = 0.032$ ($p_2 = 0$) $\omega_0 = 0.1031$, $\omega_2 = 25.997$
M3 (discrete, $K = 2$)	3	−18802.531	808.154**	$p_0 = 0.366$, $p_1 = 0.634$ $\omega_0 = 0.021$, $\omega_1 = 0.148$
M7 (beta)	2	−18712.816		$P = 0.897$, $q = 7.843$
M8 (beta and ω)	4	−18712.816	0	$p_0 = 1.00$, $p = 0.897$, $q = 7.843$ ($p_1 = 0.00$) $\omega = 3.067$
Branch site models				
Foreground 1–5				
Model A	4	−19170.373	0.254	$p_0 = 0.966$, $p_1 = 0.029$ ($p_2 + p_3 = 0.005$) $\omega_0 = 0.099$, $\omega_2 = 1.796$
Model B	5	−18799.271	6.62*	$p_0 = 0.347$, $p_1 = 0.636$ ($p_2 + p_3 = 0.0167$) $\omega_0 = 0.019$, $\omega_1 = 0.146$, $\omega_2 = 1.418$
Foreground 1 (*Zizania aquatica*)				
Model A	4	−19170.500	0	$p_0 = 0.971$, $p_1 = 0.029$ ($p_2 + p_3 = 0$) $\omega_0 = 0.100$
Model B	5	−18800.470	4.122	$p_0 = 0.113$, $p_1 = 0.205$ ($p_2 + p_3 = 0.6817$) $\omega_0 = 0.020$, $\omega_1 = 0.143$, $\omega_2 = 0$
Foreground 2 (herbaceous bamboo clade)				
Model A	4	−19165.557	9.886**	$p_0 = 0.962$, $p_1 = 0.029$ ($p_2 + p_3 = 0.009$) $\omega_0 = 0.099$, $\omega_2 = 999$
Model B	5	−18795.018	15.026**	$p_0 = 0.353$, $p_1 = 0.639$ ($p_2 + p_3 = 0.008$) $\omega_0 = 0.019$, $\omega_1 = 0.141$, $\omega_2 = 999.00$
Foreground 3 (*Bouteloua dimorpha*)				
Model A	4	−19170.450	0.1	$p_0 = 0.970$, $p_1 = 0.029$ ($p_2 + p_3 = 0$) $\omega_0 = 0.100$
Model B	5	−18800.743	3.576	$p_0 = 0$, $p_1 = 0$ ($p_2 + p_3 = 1$) $\omega_0 = 0.020$, $\omega_1 = 0.142$, $\omega_2 = 0$
Foreground 4 (Panicoid clade)				
Model A	4	−19170.450	0.1	$p_0 = 0.971$, $p_1 = 0.029$ ($p_2 + p_3 = 0$) $\omega_0 = 0.100$
Model B	5	−18802.360	0.342	$p_0 = 0.234$, $p_1 = 0.424$ ($p_2 + p_3 = 0.342$) $\omega_0 = 0.020$, $\omega_1 = 0.142$, $\omega_2 = 0$
Foreground 5 (*Pharus lappulaceous*)				
Model A	4	−19169.685	1.63	$p_0 = 0.962$, $p_1 = 0.029$ ($p_2 + p_3 = 0.009$) $\omega_0 = 0.099$, $\omega_2 = 6.797$
Model B	5	−18799.351	6.36*	$p_0 = 0.348$, $p_1 = 0.643$ ($p_2 + p_3 = 0.008$) $\omega_0 = 0.019$, $\omega_1 = 0.140$, $\omega_2 = 6.41$

Bayes empirical Bayes (BEB) tests detected no positively selected sites at > 0.95 posterior probability, except for branch site model A for the herbaceous bamboo lineage (branch 4; see text for details).
*$P < 0.05$; **$P < 0.01$.

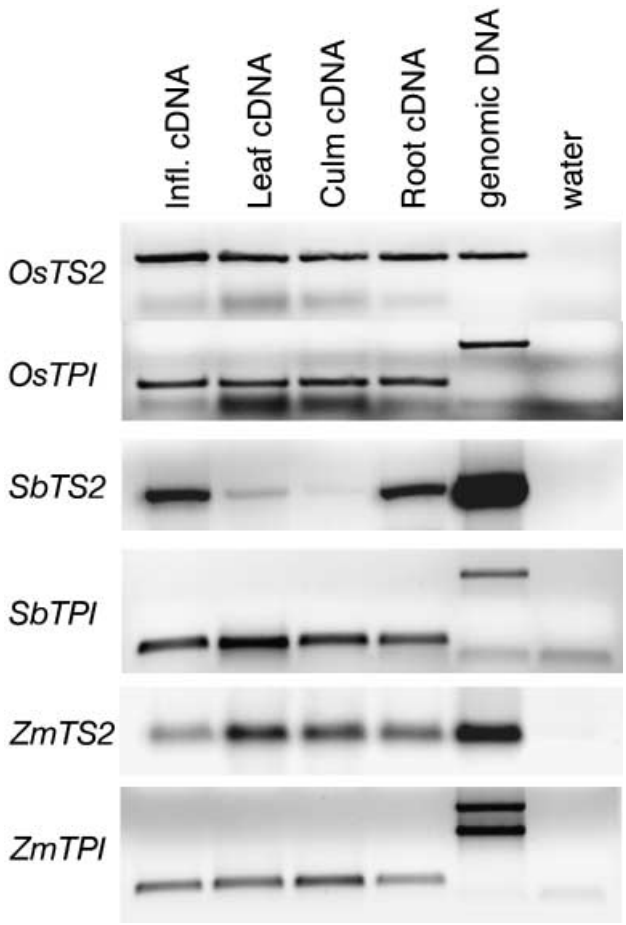

Fig. 7 *TASSELSEED2* (*TS2*) RNA RT-PCR expression patterns in inflorescences (Infl.), leaf, culm, and root tissues of *Oryza sativa* (Os), *Sorghum bicolor* (Sb) and *Zea mays* (Zm). Genomic DNA (gDNA) and *TRIOSE PHOSPHATE ISOMERASE* (*TPI*) RT-PCR were included as a positive control and to check for DNA contamination.

S. bicolor, consistent with suppression of the gynoecium in the upper flower of the pedicellate spikelet. As discussed above, *Z. mays* is monecious with staminate flowers borne in apical tassel inflorescences and pistillate flowers borne in axillary ear inflorescences. We expected *TS2* expression in both inflorescences, based on the mutant phenotypes.

Surprisingly, *TS2* is expressed in inflorescence, leaf, culm, and root tissue of all three species (Fig. 7). Amplification of *TPI*, which unlike *TS2* has introns, verified that none of the RNA extractions were contaminated by gDNA.

To investigate the expression pattern of *TS2* in the gynoecium of a species with bisexual florets, we conducted RNA *in situ* hybridization on developing rice flowers. The gene is expressed in the developing pistil and in anthers (Fig. 8). The pistil is functional in rice, indicating that, as in maize, *TS2* expression alone does not indicate which organs will undergo programmed cell death. Stamens in rice flowers appear to express *TS2* throughout development, and expression is particularly apparent in mature anthers (Fig. 8b).

Discussion

Sex expression has changed many times in the evolution of the grass family, and in disparate ways. Discussion of 'unisexual flowers' is thus somewhat misleading. The morphological diversity and repeated origins of unisexual flowers make it unlikely that there is a single master switch that affects transitions from unisexual to bisexual or vice versa. The variability of sex expression in the grasses suggests instead that multiple genes may be involved, and that modification of any one of them could lead to a different pattern of floral organ development.

Comparative morphology shows that androecial development is commonly decoupled from gynoecial development. For example, *Ixophorus unisetus* has two flowered spikelets, with the upper flower pistillate and lower one staminate (Kellogg *et al.*, 2004). The species is derived from ancestors in which the upper flower was bisexual; in the history of *Ixophorus* only stamen development in one flower of the spikelet was altered. In *Hyparrhenia hirta*, the upper flower of all pedicellate spikelets is staminate, but the upper flower of the sessile spikelets varies in sex expression depending on position in the inflorescence; sessile spikelets in the lower part of the inflorescence are staminate, whereas those in the upper part of the inflorescence

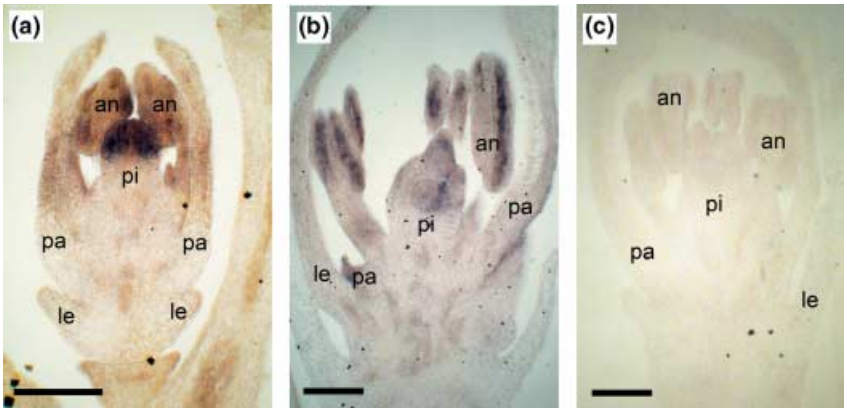

Fig. 8 *In situ* hybridization of *TASSELSEED2* (*TS2*) RNA expression in developing spikelets of *Oryza sativa* with antisense (a, b) and sense (c) digoxygenin-labeled RNA probes. (a, b) Developing spikelet with expression in anther (an), pistil (pi) and palea (pa). (c) Control section showing lack of hybridization signal using the sense probe. Bars, 100 μm.

are bisexual (Clayton & Renvoize, 1986). Space does not permit an exhaustive list of the possible combinations of staminate, pistillate, sterile, and bisexual flowers, but the many possibilities indicate precise genetic control and presumably ease of modification.

Stamen development is clearly controlled, at least in part, by GA. Blocking GA production in maize ear florets stops stamen abortion, producing the bisexual ear florets seen in *an1* and several *d* mutants (Phinney, 1961, 1984; Dellaporta & Calderon-Urrea, 1994). Stamen abortion in tassel florets of *ts2* mutants is also ascribed to perturbation of the GA pathway (Dellaporta & Calderon-Urrea, 1994), although this has not been rigorously tested. This suggests that formation of pistillate flowers may reflect modifications in availability of or sensitivity to GA during flower development.

Cytokinin is also apparently involved in regulating floral organ development. Ectopic expression of the cytokinin-synthesizing isopentyl transferase (*IPT*) permits the pistil in the lower floret in maize ear spikelets to develop, producing plants with two functional pistils in each spikelet (Young *et al.*, 2004). Curiously, the *IPT* transgene had no effect on floral development or sex expression in the tassel.

Gynoecial development is affected by TS1, TS2, TS5, and SK1, but the cellular and biochemical function of these proteins is unknown and only TS2 has been cloned. Although *TS2* was cloned over a decade ago, little information is available in the literature, perhaps in part because it is a difficult gene to work with. The GC content is high (~72%), and PCR amplification is unreliable without additives that reduce secondary structure such as DMSO and betaine, even with nondegenerate primers.

TS2 is a member of the large SDR protein superfamily, which comprises over 3000 members in over 1000 forms that act on diverse substrates including alcohols, sugars, steroids and aromatic compounds (Kallberg *et al.*, 2002). NAD, NADP or NADPH (GxxGxG) and subunit interaction or catalysis (YxxxK) motifs are conserved among all SDRs; the precise sequence of each motif is conserved among orthologs and differs among paralogs. Within the *TS2* clade, these motifs are GA/SRGIG and YTASK. The substrate is unknown, but its high sequence similarity to hydroxysteroid dehydrogenases suggests a gibberellin or steroid-like molecule (Calderon-Urrea & Dellaporta, 1999). The SDRs characterized in *A. thaliana* (*AtATA1* and *AtSDR2*) and *S. latifolia* (*STA1–12*) are not orthologous to any known monocot SDRs, including *TS2* itself, so function cannot be directly extrapolated from those two eudicots to the grass gene products. Functional studies will need to be conducted in the grasses themselves.

TS2 is expressed throughout the plant in maize, rice, and sorghum. It is expressed in both gynoecium and androecium in rice. In maize, it is expressed in all gynoecia, whether they develop fully or not (DeLong *et al.*, 1993; Calderon-Urrea & Dellaporta, 1999). We suggest that *TS2* probably has a general developmental role, and that its function in sex determination is ancillary and perhaps taxonomically restricted.

TS2 is necessary for abortion (i.e. normal development) of the gynoecium in staminate flowers of maize and *Tripsacum* (DeLong *et al.*, 1993; Li *et al.*, 1997). Expression of *TS2* in would-be staminate flowers correlates with death of cells – presumably the ones in which it is expressed – which lose their cytoplasm; death of these cells presumably prevents continued development of the gynoecium. A similar pattern of gynoecial cell death in staminate flower development is seen in other panicoid grasses (Le Roux & Kellogg, 1999).

If *TS2* had been recruited multiple times for specification of unisexual flowers, we might expect to find an elevated rate of amino acid replacements in one or more lineages. However, most data point to extensive purifying selection, indicating strong conservation of the protein sequence. Most of the amino acid replacements that do occur are conservative, occur more than once in the evolution of the grasses, and do not correlate with the origin of unisexual flowers. Only along the branch leading to the herbaceous bamboo lineage do we find evidence of selection on a single site, a nonconservative D to Q substitution.

In the case of the dioecious *B. dimorpha* clade, the origin of unisexual flowers does not correlate with any nonsynonymous changes at all. Kinney *et al.* (2003) examined the molecular evolution of 18 *TS2* alleles in *B. dimorpha* and concluded that the locus was evolving neutrally, which would suggest an elevated rate of nonsynonymous mutations. However, they used a population genetic test (Tajima's *D* statistic), rather than the likelihood ratio tests that directly address codon changes as reported here; they also amplified a longer *TS2* fragment so additional residues were certainly examined by their test. It is not therefore clear whether our results actually conflict.

Widespread expression of *TS2*, strong conservation of sequence, and lack of correlation of sequence variation and expression with unisexuality all suggest that TS2 is in fact not a sex determination protein. It seems more likely that *TS2* may have been co-opted for sex determination in one lineage of grasses (subfamily Panicoideae) but that it has multiple other roles. Additional insight may come from investigating the possible regulators of *TS2*, which have been identified by genetic, microarray, and bioinformatics approaches. At the same time, we should note that we find broad expression patterns and strong purifying selection on *TS2* even in *Zea*/*Tripsacum*, where the protein *is* believed to operate in sex expression. We therefore cannot completely rule out the possibility that it is involved in other lineages, but that we have no means to detect it.

TS2 expression is transcriptionally regulated by hormone levels, as shown by microarray experiments. Rice *TS2* is up-regulated in callus tissue exposed to GA and down-regulated when exposed to abscisic acid (ABA), indicating that the gene may play a role in the interchange between these two pathways (NCBI Gene Expression Omnibus accession GPL477; Yazaki *et al.*, 2003). The fact that *TS2* is expressed in callus tissue at all is evidence that its role is more general than

determination of sex expression in flowers. Although usually regarded as a stress hormone and growth inhibitor, ABA in *A. thaliana* also plays a key role in controlling fertility, promoting vegetative growth and determining organ size (Cheng *et al.*, 2002). Whether ABA functions in a similar way in grasses, and whether ABA and GA factor into *TS2*-mediated gynoecial cell death in panicoid grasses, is unknown.

The widespread expression of *TS2* RNA that we have observed points to possible post-transcriptional regulation. miRNAs within the 3′ UTR could block translation, as in APETALA 2 (*AP2*) in *A. thaliana* (Aukerman & Sakai, 2003) and as hypothesized for lineage (*lin*)*-4* in *Caenorhabditis elegans* (Bartel, 2004). Based on the output of MicroInspector (Rusinov *et al.*, 2005), the 3′ UTR of rice *TS2* has binding sites for *miR164* and *miR419*, whereas maize *TS2* has sites for *miR169*, *miR172*, *miR399*, *miR439* and *miR440*. Little is known about the role of these microRNAs, although *miR164* is known to regulate hormone response and *miR172* regulates translation (Kidner & Martienssen, 2005). microRNA expression profiles and TS2 protein levels could be compared to test whether *TS2* translation is regulated by microRNAs.

TS2 and the origin of unisexual flowers

Our data illustrate the power of a comparative approach for developing and testing hypotheses of gene function. At the same time, we show the perils of trying to explain the origin of a complex phenotype by studying the evolution of a single gene. Our data suggest that the role of *TS2* is much broader than simply killing cells in the gynoecium of maize, that it may be regulated post-transcriptionally as well as transcriptionally, and that the protein sequence is highly conserved among grasses.

When we began this study, we postulated that *TS2* might function in staminate flower specification in many lineages with unisexual flowers, but not in development of bisexual flowers. Expression data in rice, maize, and sorghum, and sequence data from multiple grasses all argue against this hypothesis. The protein sequence is conserved among grasses, indicating that any modifications of its developmental role are likely to be regulatory. The gene is expressed throughout the plant in the three disparate grasses investigated, indicating a function much more general than simple specification of gynoecial development. The gene is expressed in gynoecia that do develop fully as well as those that abort, indicating that even in gynoecial development it does not function as a simple on/off switch. We therefore postulate that TS2 may function in gynoecial development in all grasses, and its role in gynoecial abortion in the *Zea*/*Tripsacum* clade is a result of alterations in the activity or regulation of other genes in the gynoecial development pathway.

TS2 is regulated positively by *TS1* and apparently negatively by *SK1*, neither of which has been cloned. One tantalizing possibility is that *SK1* is a microRNA that blocks translation of *TS2*. Only if the microRNA were absent would the TS2 protein be present and functional.

Studies of the genetics of floral organ development in species other than maize will ultimately be necessary. The placement of staminate and pistillate flowers in separate inflorescences in maize adds a complication to interpretation of results, in that the fate of androecia or gynoecia in the tassel is not always the same as that of androecia or gynoecia in the ear. It would be of considerable interest to study floral organ mutations in species in which both sexes are in the same inflorescence. Fortunately, increasing numbers of genomic and genetic tools are available for multiple species of grasses. We can hope that over the next several years more detailed functional hypotheses will be available to test.

Acknowledgements

We thank the members of the Kellogg lab and three anonymous reviewers for comments and suggestions, Lynn Clark for insights into relationships and morphology of bamboos and leaf material of *Anomochloa* and *Flagellaria*, and Jessie Kossuth for help with sequencing *TS2* clones. This research was supported by grants from the University of Missouri, the National Science Foundation, and the E. Desmond Lee and Family Foundation.

References

Ainsworth C. 2000. Boys and girls come out to play: The molecular biology of dioecious plants. *Annals of Botany* **86**: 211–221.

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25**: 3389–3402.

Aukerman MJ, Sakai H. 2003. Regulation of flowering time and floral organ identity by a microRNA and its *APETALA2-like* target genes. *Plant Cell* **15**: 2730–2741.

Barrett SC. 2002. The evolution of plant sexual diversity. *Nature Reviews Genetics* **3**: 274–284.

Bartel DP. 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**: 281–297.

Calderon-Urrea A, Dellaporta SL. 1999. Cell death and cell protection genes determine the fate of pistils in maize. *Development* **126**: 435–441.

Chapman GP, ed. 1990. *Reproductive versatility in the grasses.* Cambridge, UK: Cambridge University Press.

Charlesworth D. 1985. Distribution of dioecy and self-incompatibility in angiosperms. In: Greenwood PJ, Harvey PH, Slatkin M, eds. *Evolution: essays in honour of John Maynard Smith.* Cambridge, UK: Cambridge University Press, 237–268.

Cheng WH, Endo A, Zhou L, Penney J, Chen HC, Arroyo A, Leon P, Nambara E, Asami T, Seo M, Koshiba T, Sheen J. 2002. A unique short-chain dehydrogenase/reductase in *Arabidopsis* glucose signaling and abscisic acid biosynthesis and functions. *Plant Cell* **14**: 2723–2743.

Cheng PC, Greyson RI, Walden DB. 1983. Organ initiation and the development of unisexual flowers in the tassel and ear of *Zea mays*. *American Journal of Botany* **70**: 450–462.

Clayton WD, Renvoize SA. 1986. *Genera Graminum: grasses of the world.* London, UK: Her Majesty's Stationery Office.

Coe EH, Neuffer MG, Hoisington DA. 1988. The genetics of corn. In: Sprague G, Dudley J, eds. *Corn and corn improvement.* Madison, WI, USA: American Society of Agronomy, Crop Science Society of America, and Soil Science Society of America, 81–258.

Columbus JT. 1999. An expanded circumscription of *Bouteloua* (Gramineae: Chloridoideae): New combinations and names. *Aliso* **18**: 61–65.

Connor HE. 1981. Evolution of reproductive systems in the Gramineae. *Annals of the Missouri Botanical Garden* **68**: 48–74.

Dellaporta SL. 1994. Plant DNA miniprep and microprep: versions 2.1–2.3. In: Freeling M, Walbot V, eds. *The maize handbook*. New York, NY, USA: Springer-Verlag, 522–525.

Dellaporta SL, Calderon-Urrea A. 1994. The sex determination process in maize. *Science* **266**: 1501–1505.

DeLong A, Calderon-Urrea A, Dellaporta SL. 1993. Sex determination gene *TASSELSEED2* of maize encodes a short-chain alcohol dehydrogenase required for stage-specific floral organ abortion. *Cell* **74**: 757–768.

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

Gadberry MD, Malcomber ST, Doust AN, Kellogg EA. 2005. Primaclade – a flexible tool to find conserved PCR primers across multiple species. *Bioinformatics* **21**: 1263–1264.

Giussani LM, Cota-Sanchez JH, Zuloaga FO, Kellogg EA. 2001. A molecular phylogeny of the grass subfamily Panicoideae (Poaceae) shows multiple origins of C_4 photosynthesis. *American Journal of Botany* **88**: 1993–2012.

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

Huelsenbeck JP, Ronquist F. 2001. MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**: 754–755.

Irish EE, Langdale J, Nelson TM. 1994. Interactions between *tasselseed* genes and other sex determining genes in maize. *Developmental Genetics* **15**: 155–171.

Jones DF. 1934. Unisexual maize plants and their bearing on sex determination in other plants and in animals. *Genetics* **19**: 552–567.

Jörnvall H, Hoog JO, Persson B. 1999. SDR and MDR: completed genome sequences show these protein families to be large, of old origin, and of complex nature. *FEBS Letters* **445**: 261–264.

Kallberg Y, Oppermann U, Jornvall H, Persson B. 2002. Short-chain dehydrogenase/reductase (SDR) relationships: a large family with eight clusters common to human, animal, and plant genomes. *Protein Science* **11**: 636–641.

Kellogg EA, Hiser KM, Doust AN. 2004. Taxonomy, phylogeny, and inflorescence development of the genus *Ixophorus* (Panicoideae: Poaceae). *International Journal of Plant Sciences* **165**: 1089–1105.

Kidner CA, Martienssen RA. 2005. The developmental role of microRNA in plants. *Current Opinion in Plant Biology* **8**: 38–44.

Kinney MS, Columbus JT, Friar EA. 2003. Molecular evolution of the maize sex-determining gene *TASSELSEED2. Bouteloua* (Poaceae). *Molecular Phylogenetics and Evolution* **29**: 519–528.

Lambert C, Leonard N, De Bolle X, Depiereux E. 2002. ESyPred3D: Prediction of proteins 3D structures. *Bioinformatics* **18**: 1250–1256.

Laurie DA, Pratchett N, Devos KM, Leitch IJ, Gale MD. 1993. The distribution of RFLP markers on chromosome 2 (2H) of barley in relation to the physical and genetic location of 5S rDNA. *Theoretical and Applied Genetics* **87**: 177–183.

Le Roux LG, Kellogg EA. 1999. Floral development and the formation of unisexual spikelets in the Andropogoneae (Poaceae). *American Journal of Botany* **86**: 354–366.

Lebel-Hardenack S, Ye D, Koutnikova H, Saedler H, Grant SR. 1997. Conserved expression of a *TASSELSEED2* homolog in the tapetum of the dioecious *Silene latifolia* and *Arabidopsis thaliana*. *Plant Journal* **12**: 515–526.

Li D, Blakey CA, Dewald C, Dellaporta SL. 1997. Evidence for a common sex determination mechanism for pistil abortion in maize and in its wild relative *Tripsacum*. *Proceedings of the National Academy of Sciences, USA* **94**: 4217–4222.

Maddison DR, Maddison WP. 2003. MacClade: Analysis of phylogeny and character evolution. Sunderland, MA, USA: Sinauer Associates.

Malcomber ST, Kellogg EA. 2004. Heterogeneous expression patterns and separate roles of the *SEPALLATA* gene *LEAFY HULL STERILE1* in grasses. *Plant Cell* **16**: 1692–1706.

Martz E. 2002. Protein Explorer: easy yet powerful macromolecular visualization. *Trends in Biochemical Science* **27**: 107–109.

Michelangeli FA, Davis JI, Stevenson DW. 2003. Phylogenetic relationships among Poaceae and related families as inferred from morphology, inversions in the plastid genome, and sequence data from the mitochondrial and plastid genomes. *American Journal of Botany* **90**: 93–106.

Mitchell CH, Diggle PK. 2005. The evolution of unisexual flowers: morphological and functional convergence results from diverse developmental transitions. *American Journal of Botany* **92**: 1068–1076.

Phinney BO. 1961. Dwarfing genes in *Zea mays* and their relation to the gibberellins. In: Klein RM, ed. *Plant growth regulation*. Ames, IA, USA: Iowa State University Press, 489–501.

Phinney BO. 1984. Gibberellin A_1, dwarfism and the control of shoot elongation and metabolism in higher plants. In: Crozier A, Hillman JR, eds. *The biosynthesis and metabolism of plant hormones*. Cambridge, UK: Cambridge University Press, 17–41.

Posada D, Crandall KA. 1998. MODELTEST: testing the model of DNA substitutions. *Bioinformatics* **14**: 817–818.

Richards AJ. 1997. *Plant breeding systems*. New York, NY, USA: Chapman & Hall.

Rudall PJ, Stuppy W, Cunniff J, Kellogg EA, Briggs BG. 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.

Rusinov V, Baev V, Minkov IN, Tabler M. 2005. MicroInspector: a web tool for detection of miRNA binding sites in an RNA sequence. *Nucleic Acids Research* **33**: W696–W700.

Stevens PF. 2001 (onwards). *Angiosperm phylogeny website*, version 6. http://www.mobot.org/MOBOT/research/APweb/

Strand AE, Leebens-Mack J, Milligan BG. 1997. Nuclear DNA-based markers for plant evolutionary biology. *Molecular Ecology* **6**: 113–118.

Swofford DL. 2000. *PAUP*: Phylogenetic Analysis Using Parsimony.* Sunderland, MA, USA: Sinauer Associates.

Tanurdzic M, Banks JA. 2004. Sex-determining mechanisms in land plants. *Plant Cell* **16**: S61–S71.

Veit B, Schmidt RJ, Hake S, Yanofsky MF. 1993. Maize floral development: New genes and old mutants. *Plant Cell* **5**: 1205–1215.

Watson L, Dallwitz MJ. 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/

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

Wong WS, Yang Z, Goldman N, Nielsen R. 2004. Accuracy and power of statistical methods for detecting adaptive evolution in protein coding sequences and for identifying positively selected sites. *Genetics* **168**: 1041–1051.

Yang Z. 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. *Computer Applications in the Biosciences* **13**: 555–556.

Yang Z, Nielsen R. 2002. Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. *Molecular Biology and Evolution* **19**: 908–917.

Yang Z, Wong WS, Nielsen R. 2005. Bayes empirical bayes inference of amino acid sites under positive selection. *Molecular Biology and Evolution* **22**: 1107–1118.

Yazaki J, Kishimoto N, Nagata Y, Ishikawa M, Fujii F, Hashimoto A,

Shimbo K, Shimatani Z, Kojima K, Suzuki K, Yamamoto M, Honda S, Endo A, Yoshida Y, Sato Y, Takeuchi K, Toyoshima K, Miyamoto C, Wu J, Sasaki T, Sakata K, Yamamoto K, Iba K, Oda T, Otomo Y, Murakami K, Matsubara K, Kawai J, Carninci P, Hayashizaki Y, Kikuchi S. 2003. Genomics approach to abscisic acid- and gibberellin-responsive genes in rice. *DNA Research* **10**: 249–261.

Young TE, Giesler-Lee J, Gallie DR. 2004. Senescence-induced expression of cytokinin reverses pistil abortion during maize flower development. *Plant Journal* **38**: 910–922.

Zaitchik BF, Le Roux LG, Kellogg EA. 2000. Development of male flowers in *Zizania aquatica* (North American Wild-Rice; Gramineae). *International Journal of Plant Sciences* **161**: 345–351.