MORPHOLOGICAL DIVERSITY AND GENETIC REGULATION OF INFLORESCENCE ABSCISSION ZONES IN GRASSES

ANDREW N. DOUST, MARGARITA MAURO-HERRERA, AMIE D. FRANCIS, AND LAURA C. SHAND

Botany Department, Oklahoma State University, Stillwater, Oklahoma 74078 USA

- **Premise of the study:** Variation in how seeds are dispersed in grasses is ecologically important, and selection for dispersal mechanisms has produced a great variety of dispersal structures (diaspores). Abscission (“shattering”) is necessary in wild grasses, but its elimination by selection on nonshattering mutants was a key component of the domestication syndrome in cereal grasses. A key question is whether a common genetic pathway controls abscission in wild grasses, and, if so, what genes in that pathway may have been selected upon during domestication. We summarize morphological and genetic information on abscission zones and disarticulation patterns in grasses and identify hypotheses to test the likelihood of a common genetic pathway.

- **Methods:** Morphological data on abscission zones for over 10 000 species of grasses were tabulated and analyzed using a tribal phylogeny of the grasses. The genomic location of quantitative trait loci (QTLs) and orthologs of genes controlling shattering were compared across species to ascertain whether the same loci might control shattering in different grass lineages.

- **Results and conclusions:** The simple trait of nonshattering is derived from a great diversity of shattering phenotypes. Several sets of QTLs from multiple species are syntenic yet many are not. Genes known to be involved in shattering in several species were found to have orthologs that sometimes colocalized with QTLs in different species, adding support to the hypothesis of retention of a common genetic pathway. These results are used to suggest a research plan that could test the common genetic pathway model more thoroughly.

**Key words:** abscission zone; diaspore; disarticulation; domestication; genetic regulation; grasses; inflorescence morphology; Poaceae; shattering; seed dispersal.

In wild plants, seeds must be dispersed for effective reproduction. Dispersal in most cases requires disarticulation of the seed or fruit from the body of the plant via means of the formation of an abscission zone (AZ) (Liljegren, 2012; Estornell et al., 2013). In the grasses, the seed is rarely dispersed alone, but rather with a fused or adherent fruit wall (making the characteristic grass fruit termed a caryopsis) and any of a variety of other structures (Clayton, 1990). Thus, the dispersal unit or diaspore (Howe and Smallwood, 1982) varies between groups of grasses, affecting the biology of dispersal and the vectors involved (e.g., wind, animals, water) (Davidse, 1987; Chapman, 1996).

The process of domestication in cereal crops has involved selection on mutants in which the AZ has been modified and disarticulation disrupted. The emergence of nonshattering forms is an essential component of the so-called “domestication syndrome” (Harlan, 1992), recognizable in all present-day cereal crops and in archeobotanical deposits of their ancestors (Fuller, 2007). Selection for nonshattering, so that seed can be easily collected, is also characteristic of other domestications, such as sunflower (Burke et al., 2002; Wills and Burke, 2007) and soybean (Gao and Zhu, 2013; Dong et al., 2014). Lack of shattering is also one of the first traits to be selected upon in recent domestinations, such as in wild rice (Zizania palustris) (Hayes et al., 1989; Kennard et al., 2002) and Microlaena stipoides, a potential new perennial cereal (Shapert et al., 2013).

Although the effect of selection is the same in all domesticated cereals (a nonshattering phenotype where the grain does not fall easily from the inflorescence), the wild progenitors of those cereals as well as grasses as a whole differ widely in their manner of disarticulation (Fig. 1). Dispersal units can vary from the seed (e.g., Sporobolus) to the whole inflorescence (e.g., Spinifex), and to almost everything in between. This diversity of dispersal units is achieved by the formation of an AZ in one or more of multiple locations within the inflorescence, leading to multiple forms of diaspores (cartoons of which are illustrated in Fig. 1 together with several illustrative photographs in Fig. 2).

Diaspore structure can often be directly related to dispersal method, although in many grasses diaspores simply fall off the plant when mature so that most seed falls relatively close to the parent (Clayton, 1990; Carey and Watkinson, 1993). Modifications to the diaspore can increase the chance of longer-distance dispersal, such as the caryopsis being dispersed together with floral parts modified into awns (e.g., the lemma awns of many pooid grasses that catch in the fur of animals). The hygroscopic movement of awns can also help to move the diaspore when it reaches the ground, by pushing it into crevices that will promote successful germination (Peart, 1979, 1981, 1984). Other grains are dispersed as groups of spikelets together with the branches they are attached to, such as in Cenchrus, where some species have branches modified into spines, while others have feathery branches. In species such as Spinifex and Chloris, the whole inflorescence detaches and is

---

1 Manuscript received 21 April 2014; revision accepted 21 August 2014.

We thank Derek Clayton, Maria Vorontsova, Elizabeth Kellogg, and Barbara Briggs for help in interpreting and clarifying aspects of abscission in grasses and the editors and two anonymous reviewers for many suggestions that improved the manuscript. We also thank Armond Swift, Jessica Stromski, and Kimberly Rogers for help in data collection.

2 Author for correspondence (e-mail: andrew.doust@okstate.edu)

doi:10.3732/ajb.1400186
A Spikelet structure

---

B Below glumes
C Above glumes
D Below florets
E Between florets, below floret insertion
F Between florets, above floret insertion
G Below seed
H Below caryopsis
I Inflorescence with many-flowered spikelets
J In rachis above spikelet
K In rachis below spikelet
L Inflorescence with one fertile floret per spikelet
M Disarticulation below glumes
N Inflorescence with paired spikelets, each spikelet with one fertile floret
O In rachis above sessile spikelet
P In rachis below sessile spikelet
Q Inflorescence of spikelets and sterile branches
R Below glumes
S At base of branch

---

* = abscission zone, where disarticulation occurs
The genetic control of loss of shattering has been studied in all major crops (Kandemir et al., 2000, 2004; Kennard et al., 2002; Poncet et al., 2002; Matsui et al., 2004; Konishi et al., 2006; Li et al., 2006; Nalam et al., 2006; Vidya Saraswathi transported by wind or water (Chapman, 1996). Thus, the single character of nonshattering that is part of the domestication syndrome in cereal crops is imposed on an evolutionary and ecologically diverse background.

Fig. 1. Inflorescence structure and disarticulation patterns in grasses. Diagrams in blue represent the inflorescence structure before disarticulation of dispersal units, diagrams in gray represent the variety of disarticulation patterns. Spikelet glumes are represented in green in all diagrams. Red dashed lines indicate the position of the AZs. (A) Spikelet structure. (B) Disarticulation below glumes, dispersal unit is spikelet. (C) Disarticulation above the glumes; dispersal unit is collection of florets. (D) Disarticulation below florets; dispersal unit is individual floret. (E) Disarticulation in rachilla above floret insertion; dispersal unit is floret attached to basilar rachilla internode. (G) Disarticulation below seed; dispersal unit is seed. (H) Disarticulation below caryopsis; dispersal unit is caryopsis. (I) Inflorescence with many-flowered spikelets. (J) Disarticulation in rachis above spikelet insertion; dispersal unit is spikelet attached to basilar rachis internode. (K) Disarticulation in rachis below spikelet insertion; dispersal unit is spikelet attached to distal rachis internode. (L) Inflorescence with one fertile floret per spikelet. (M) Disarticulation below glumes; dispersal unit is spikelet. (N) Inflorescence with paired spikelets, each spikelet pair with one fertile floret. (O) Disarticulation in rachis above spikelet pair insertion; dispersal unit is spikelet pair attached to distal rachis internode. (P) Disarticulation in rachis below spikelet pair insertion; dispersal unit is spikelet pair attached to distal rachis internode. (Q) Inflorescence of spikelets and sterile branches. (R) Disarticulation below glumes; dispersal unit is spikelet. (S) Disarticulation at base of branch, dispersal unit is whole branch.

Fig. 2. Representative disarticulation patterns in the grasses. (A) Avena fatua, disarticulation in the rachilla; dispersal unit a floret attached to a rachilla internode (Fig. 1E). (B) Hordeum pusillum, disarticulation in the rachis; dispersal unit a spikelet with three florets attached to a rachis internode (Fig. 1K). (C) Paspalum pubiflorum, disarticulation below the glumes, dispersal unit is a two-flowered spikelet with a single fertile floret (Fig. 1M). (D) Setaria viridis, inflorescence is constructed of spikelets and sterile branches, with the spikelets disarticulating individually below the glumes (Fig. 1R). (E) Cenchrus pauciflorus; dispersal unit is a fascicle of spikelets and sterile branches, with the branches modified into spines for animal dispersal (Fig. 1S). (F) Microstegium vimineum, disarticulation in the rachis (rame), each spikelet pair falling with an internode (Fig. 1P). Voucher information is listed in online Appendix S2. Abbreviations: a = awn, b = branch, f = floret, g = glume, i = rachilla internode, ps = pedicellate spikelet, r = rachis (or rame) internode, s = spikelet, sb = spikelet-branch compound structure, ss = sterile spikelet. White triangles: disarticulation zones. Scale bars: C, E, G, H = 5 mm; D = 3 mm; A, B, F = 10 mm.
et al., 2006; Watanabe et al., 2006; Gill et al., 2007; Onishi et al., 2007a; Weber et al., 2008; Larson and Kellogg, 2009; Zhang et al., 2009; Ji et al., 2010; Qin et al., 2010; Lin et al., 2012; Zhou et al., 2012; Tang et al., 2013), and genetic comparisons made between crops, first on the basis of shared molecular markers such as restriction fragment length polymorphisms (RFLPs) and more recently by genome comparisons. The monophyly of the grasses (GPWG II, 2012) and the high degree of synteny (relative order and identity of genes) between grass genomes (Moore et al., 1995; Gale and Devos, 2001; Devos, 2005) suggest that a shared genetic pathway may have been inherited from a common grass ancestor. A shared pathway was supported by the pioneering study of Paterson et al. (1995), who used genetic maps based on RFLP markers to compare the positions of quantitative trait loci (QTLs) for shattering, short-day flowering mutations, and seed size (Paterson et al., 1995). At the level of resolution afforded by these relatively low-density maps, they discovered that QTLs for these traits overlapped in rice, sorghum, and maize. Since QTLs represent regions of the genome where genotypic variation (e.g., differing alleles of causative genes) is significantly correlated with phenotypic variation, this suggests that overlapping QTLs from different species contain the same controlling genes. Several other studies have also shown that QTLs for shattering in different species overlap, including in Pennisetum (Poncet et al., 2002), Setaria (Doust et al., 2014) and Leymus ( Larson and Kellogg, 2009). However, more detailed mapping studies have often revealed QTLs that do not overlap, such as in comparisons between the domesticated Triticeae (wheat, barley, rye) and domesticated species in other tribes such as rice (Oryzaeae), sorghum (Andropogoneae) and maize (Andropogoneae) (for tribal relationships, see Fig. 3). For example, two main loci controlling shattering in wheat, Br1 and Br2, do not map to any QTLs in other species (Gill et al., 2007). Furthermore, the major genes identified for shattering in rice, including sh4 and gSH1, do not appear to have orthologs controlling shattering in the Triticeae, leading Li and Gill (2006) to declare that loss of shattering was achieved by mutation of different genes in the different lineages and that there was no direct evidence supporting the common gene hypothesis.

In addition to whether the same QTLs (or genes) are detected in multiple species, various studies have revealed that the number of QTLs controlling shattering differ among species, with sorghum having only one major locus while rice and maize have several. In addition, when studies on a particular species are considered, different sets of QTLs may be detected, suggesting that some QTLs merely reflect differences between the parents rather than differences between the species. Differences in QTLs also reflect differences in mapping population structure and size, potentially reducing the informativeness of QTL comparisons between species.

Recent studies in sorghum have provided new perspectives on whether there is a shared genetic pathway for shattering in grasses. In one study, a major locus controlling shattering in domesticated sorghum was identified and found to be syntenic to the QTL for shattering in foxtail millet, maize, and rice (Lin et al., 2012). The locus containing this gene (SH1) is the single major locus in sorghum but appears to be of minor effect in rice. This is the first gene-level evidence that the evolution of nonshattering during domestication is under common genetic control. In addition, the study showed that selection at different times and at different places had targeted the gene multiple times during the domestication of sorghum. However, another study on the control of nonshattering mutants in Sorghum propinquum, a wild relative of domesticated sorghum, found a separate causal gene mutation (SBWRKY) that is confined to the S. propinquum lineage, even though it is within 300kb of the SH1 gene (Tang et al., 2013). Thus, within a single region, sorghum contains two shattering loci, one of which may be part of a common pathway in grasses while the other is novel within one species.

There are multiple possibilities for genetic control of shattering in grasses, even if formation of an AZ is controlled by a common genetic pathway. For instance, if formation of an AZ always involves identical genes, a change in an upstream regulatory module may change when and where the AZ program is turned on. Given a common genetic pathway, selection during domestication may always act on the same gene in that pathway or on different genes. Conversely, there may be only a handful of conserved genes and many novel loci, but selection again could have worked on either or both of the conserved or novel genes. Therefore, there is potential for considerable genetic complexity, and more data and testable hypotheses are needed to judge the extent to which there are elements of a common genetic pathway for disarticulation in the grasses. Before this is possible, a better understanding of both diversity of disarticulation patterns and genetic evidence for shared pathways across the grasses is necessary. Here we examine variation in AZs across the grasses, along with colocalization of QTLs and genes presently known to control disarticulation. We then attempt to relate orthologs of genes known to control shattering in several species to QTLs for shattering identified in various cereal domestications and thus provide hypotheses to test a common pathway model.

MATERIALS AND METHODS

Morphological variation in shattering across the grasses—Morphological data on disarticulation patterns in each species of grass was derived primarily from GrassBase, a remarkable repository of morphology and other information on the grasses constructed and maintained by the Royal Botanic Gardens, Kew (Clayton et al., 2014). Other sources of information included Grass Genera of the World (Watson and Dallwitz, 1992 onward) and selected floras, including Flora of North America (Barkworth et al., 2003, 2007) and Flora of China (Chen et al., 2006). Only taxa in GrassBase identified as having flowering specimens were used (10/47 of 11,290 species). We chose a conservative approach to character state coding by concentrating on the position of the AZ, distinguishing between disarticulation of (1) whole inflorescences, (2) spikelets plus rachis or inflorescence branches, (3) spikelets (including glumes), and (4) spikelet fragments (including florets) (Appendix S1, see Supplemental Data with the online version of this article). Even though some species exhibit more than one AZ position, we focused on the primary, or initial, AZs, rather than the later breakup of subunits of the inflorescence. Herbarium specimens representing the majority of the tribes used in the analysis were examined at the Herbarium, Department of Botany, Oklahoma State University, and used in preparing illustrative photographs (Appendix S2, see online Supplemental Data). SPSS version 21 (SPSS, Chicago, Illinois, USA) was used for statistical analysis.

Generic and supergeneric delimitations—The GrassBase morphological database is actively curated, but changes in taxonomy in recent years mean that up to 10–17% of the names may need to be changed to comply with modern phylogenetic concepts (Vorontsova and Simon, 2012). Although species delimitations were not reassessed for this study, several generic and tribal delimitations were altered to fall in line with modern taxonomic and phylogenetic criteria (Sánchez-Ken and Clark, 2010; GPWG II, 2012). In particular, we have followed recent taxonomic treatments in reassessing generic delimitations for the previously large genus Panicus (Zuloaga et al., 1992; Morrone et al., 2008, 2012; Sede et al., 2008; Sede et al., 2009; Zuloaga et al., 2010) and have used the tribal classification delineated by Kellogg (in press) to map species and genera in GrassBase to the current tribal classification (Kellogg, in press). Relationships between these tribes follow that of the Grass Phylogeny Working
Fig. 3. Distribution of disarticulation patterns across the grasses. The distribution of different dispersal units are indicated for the various tribes within the grass family. Tribes in red are those that have domesticated cereal crops, with the crops listed to the right of the tribe. Dispersal units represented in graphs for each tribe include (1) entire inflorescences (black bars), (2) spikelets plus accessory structures (branches, pedicels, etc.) (pink bars), (3) single spikelets (blue bars), and (4) florets and spikelet parts (green bars).
Group II and represent all strongly supported nodes from recent analyses (GPWG II, 2012; Kellogg, in press).

Identification of QTLs and cross-genome comparisons of their location—

Literature on domestication in the grasses was searched for QTLs identified as controlling shattering. In particular, we focused on studies that mapped QTLs between domesticated species and their wild progenitors. The genomic coordinates of the flanking markers of QTLs were identified, where possible, using the Gramene database (http://www.gramene.org/, release 41), and only those for which physical genomic locations could be ascertained were used in cross-species comparisons. In several studies, only a single marker was identified as the QTL location, rather than flanking markers, but such QTLs were still included in the analysis. For comparing QTL locations, positions of rice orthologs for the flanking markers of QTLs from Setaria, Sorghum, and Zea were located on the rice genome and assessed for any overlap of QTLs among each other and with QTLs identified from rice.

Identification of shattering gene orthologs—Seven genes shown to be involved in the genetic regulation of shattering were identified from the literature (Table 1). We assessed homology initially by reciprocal nucleotide and protein BLAST searches for orthologs of the seven genes in each grass genome in the Phytozome database (Goodstein et al., 2012) and at the National Center for Biotechnology Information (NCBI). Gene orthology was estimated using a suite of programs incorporated into the Geneious workflow computing environment (Geneious version 6, Biomatters, Auckland, New Zealand, http://www.geneious.com/), including protein sequence alignment for each locus using the program MUSCLE (Edgar, 2004), followed by visual checks of the validity of the alignment, and maximum likelihood phylogenies using the program PhyML 3.0 (Guindon et al., 2010), with support assessed by analyses of 100 bootstrap data sets for each protein alignment. Model estimation was done using the program ProTest 3 (Abascal et al., 2005).

We used the SynMap function (Lyons et al., 2008) in the online CoGe portal to perform cross-species comparisons between the five sequenced grass genomes (Lyons and Freeling, 2008; Tang and Lyons, 2012). These genomes include Oryza sativa var. japonica (subfamily Ehrhartoideae, tribe Oryzeae, CoGE version id16890), Brachypodium distachyon (subfamily Pooidae, tribe Brachypodieae, CoGE version id8120), Setaria italica (foxtail millet, subfamily Panicoideae, tribe Paniceae, CoGE version id19491), Sorghum bicolor (subfamily Panicoideae, tribe Andropogoneae, CoGE version id95), and Zea mays (subfamily Panicoideae, tribe Andropogoneae, CoGE version id16905). We configured SynMap to assign gene pairs to classes based on their Ks values and thus to distinguish orthologous syntenic regions (resulting from speciation from a common ancestor) from paralogous syntenic regions that were the product of the much older pan-grass whole genome duplication (Paterson et al., 2004; Schnable and Lyons, 2011). Only orthologous syntenic regions were used for comparison, and we use the term syntenic to refer to such regions. Syntenic depth was set as 1:1 between the diploid species and 2:1 for comparisons between maize and diploid species, because of the additional whole genome duplication in maize (Blanc and Wolfe, 2004). The genomic positions of gene orthologs were then compared with QTL locations.

### Table 1. QTLs and genes identified from the literature and their genomic positions.

<table>
<thead>
<tr>
<th>Species</th>
<th>Chromosome</th>
<th>Reference</th>
<th>Flanking markers from original map</th>
<th>Estimated genomic position (chromosome: bp)</th>
<th>Colocalized orthlogs of genes controlling shattering, genomic position</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oryza sativa</td>
<td>1</td>
<td>Thomson et al., 2003</td>
<td>RM315–RM104 (RG331)</td>
<td>LOC_Os01g62920 (qSH1), 36445019–36449951</td>
<td></td>
</tr>
<tr>
<td>Oryza sativa</td>
<td>1</td>
<td>Onishi et al., 2007a, b</td>
<td>RM8278</td>
<td>36621939–36622082</td>
<td></td>
</tr>
<tr>
<td>Oryza sativa</td>
<td>1</td>
<td>Cai and Morishima, 2000</td>
<td>G393–OP</td>
<td>31047047–35000000</td>
<td></td>
</tr>
<tr>
<td>Oryza sativa</td>
<td>3</td>
<td>Doebel and Scel, 1991; Onishi et al., 2007a, b</td>
<td>RM16</td>
<td>23126064–23126231</td>
<td></td>
</tr>
<tr>
<td>Oryza sativa</td>
<td>3</td>
<td>Fukata and Yagi, 1998</td>
<td>R250–C746</td>
<td>25115568–27128223</td>
<td></td>
</tr>
<tr>
<td>Oryza sativa</td>
<td>3</td>
<td>Gu et al., 2005a</td>
<td>RM487–RM520</td>
<td>22019474–28442656</td>
<td></td>
</tr>
<tr>
<td>Oryza sativa</td>
<td>3</td>
<td>Li et al., 2006</td>
<td>RM168–RM293</td>
<td>28091534–31657418</td>
<td></td>
</tr>
<tr>
<td>Setaria italica</td>
<td>5</td>
<td>Douet al., 2014</td>
<td>UGSF367–UGSF379</td>
<td>Si000753m (Os_qSH1), 41418123–41422628</td>
<td></td>
</tr>
<tr>
<td>Setaria italica</td>
<td>9</td>
<td>Douet al., 2014</td>
<td>UGSF23–UGSF40</td>
<td>Si037789m (Sb_WRKY), 91073047–10077808</td>
<td></td>
</tr>
<tr>
<td>Sorghum bicolor</td>
<td>1</td>
<td>Lin et al., 2012; Tang et al., 2013</td>
<td>SOG0251–SOG1273</td>
<td>Sb01g03270 (SSH1), 1:12285545–12287480</td>
<td></td>
</tr>
<tr>
<td>Zea mays</td>
<td>1</td>
<td>Doebel and Scel, 1991</td>
<td>UMC83–UMC107</td>
<td>GRMZM2G142919 (SwWRKY), 1:258955026–258956380</td>
<td></td>
</tr>
<tr>
<td>Zea mays</td>
<td>2</td>
<td>Doebel and Scel, 1991</td>
<td>BNLS6–BNL298 (UMC42A)</td>
<td>3:146023–30994638</td>
<td></td>
</tr>
<tr>
<td>Zea mays</td>
<td>3</td>
<td>Doebel and Scel, 1993</td>
<td>Cu5328–UMC38</td>
<td>1:17951203–179516809</td>
<td></td>
</tr>
</tbody>
</table>

**Notes:** This table lists QTL regions in each species that colocalize with the positions of orthologs of genes controlling shattering (For a full list of QTLs, see online Appendix S4). Where possible, the genomic locations of QTLs were estimated from flanking markers. However, in several cases, only the marker closest to the peak of the QTL was reported; in other cases, it was not possible to find the genomic position of markers. In several of those cases, the marker that was originally reported is given in parentheses, and the nearest marker that could be localized was listed and used to determine the position of the QTL. Thus, QTL regions are approximate only. Colocalized orthologs have the ortholog first, followed by the query gene in parentheses, and then the genomic position of the ortholog. Separate studies within a species that identify the same QTL have been grouped together, as indicated by shading.
RESULTS

Morphological variation across the grasses—Multiple disarticulation patterns are found across the grasses (Figs. 1, 2), but we concentrated on four main patterns, distributed to varying degrees throughout the family (Fig. 3; online Appendix S3). The first pattern is quite rare and consists of the whole inflorescence disarticulating to become the diaspore. In several cases, secondary abscission of the spikelets or florets occurs, facilitating widespread dispersion of the seed. Dispersal of these inflorescences is often by wind, such as in Spinifex and Chloris species. A second pattern is for diaspores to consist of spikelets falling while attached to other structures. Within this category are a number of overlapping morphologies, with unclear homology, such as spikelets falling alone but retaining their pedicel, spikelets dispersed in groups joined by their fused pedicels, and spikelets falling attached to part of the rachis (branch). Spikelet–branch combinations may be unmodified, or some branches may be sterile and become modified into plumose or spiny structures to enhance dispersal by wind or animals (e.g., Cenchrus). A third pattern, particularly prominent in the panicoide grasses, is disarticulation of whole spikelets, including glumes. The fourth pattern, the most common across the grasses, is for parts of the spikelet to disarticulate, whether that be fertile florets, florets attached to rachilla internodes, or groups of florets. Usually the florets are dispersed together with the lemma and palea, and sometimes with the lemma and/or palea modified into an awn to aid in dispersal. Other variations include genera such as Eragrostis, where the seed and fruit wall (caryopsis) is the dispersal unit, and Sporobolus, where the seed is the dispersal unit. Species may possess more than one method of disarticulation, such as the whole inflorescence followed by spikelet disarticulation, or spikelet followed by individual floret disarticulation.

QTL identification and comparison—Data from 13 QTL studies were used to assess the extent to which QTL colocalize within species (from different mapping studies) and between species. These studies included a total of 43 QTLs, several of which mapped to the same genomic positions (online Appendix S4). After similar QTL positions were merged, there were 14 distinct QTL regions in rice, two in Setaria, one in sorghum, and 11 in maize (Fig. 4). However, no one study identified more than six QTLs. The greater number of QTLs identified in rice and maize is correlated with the greater number of independent populations examined, and there is the possibility that the creation of new mapping populations in Setaria and sorghum may lead to the discovery of further shattering QTLs.

When QTLs from all species are considered, there are approximately 20 distinct QTL regions, of which five are shared across more than one species (Fig. 4). Of these, five appear to colocalize with orthologs of shattering genes, on chromosome 1 (OsqSH1), chromosome 3 (SbSH1 and SbWRKY), and chromosome 4 (OsSHAT1) (Table 1, Fig. 4). There are several cases of QTLs from other species colocalizing with shattering gene orthologs. In Setaria, one QTL colocalizes with an ortholog of the SH1 gene from sorghum, while the other is very close to the qSH1 gene from rice. In rice, one QTL colocalizes with the ortholog of the SH1 gene from sorghum, as previously reported (Lin et al., 2012), as well as the ortholog of the WRKY gene from sorghum, a gene that is within 500 kb of the SH1 ortholog (Table 1). Two of the QTLs in maize colocalize with orthologs of SH1 and WRKY from sorghum on chromosomes 5 and 1, respectively. Due to their location on separate chromosomes, orthologs of the two genes may have separate effects in maize, whereas in sorghum they are so close to each other that it may not be possible to separate their effects. The population used to map the QTLs in maize had over 800 recombinant inbred lines (RILs), giving great statistical power and allowing very precise localization of the QTLs (Lin et al., 2012). In the case of maize, several of the large QTLs from early studies (e.g., Doebley and Stec, 1991, 1993) localized to more than one chromosome in rice, adding uncertainty to the precise location of those QTLs.

Gene orthology and syntology—Gene orthology was assessed by a reciprocal BLAST search of query and orthologous sequences (online Appendix S5), and through construction of gene trees using protein sequences (online Appendix S6). To assess the synteny of orthologs, we analyzed the positions of known orthologs of shattering genes in pairwise genomic comparisons using the SynMap and SynFind modules in CoGe (Appendix S5). Most genes had orthologs in syntenous positions across the genomes although several genes were not found in orthologous positions in all genomes, and the orthologs of OsCPL1 were not found in syntetic positions in any genome (Appendix S5).

DISCUSSION

Outgroups of the grasses (GPWG II, 2012) shed their seed as a fleshy fruit (Joinvilleaceae, Flagellariaceae) or as a hard nut or loculicidal capsule (Ecdieiocoleaceae) (Rudall et al., 2005). These groups do not have the unique spikelet structure found in almost all grasses (Kellogg, 2001). Our analyses of variation across the grasses indicate that the presence of the spikelet has likely enhanced the variety of dispersal structures that can be produced. The survey of AZ position and the resulting diaspore structure reveals that most groups show a high percentage of species with AZs within the spikelet itself, giving rise most commonly to a diaspore consisting of the caryopsis with associated lemma and palea (floral parts) or to these three plus part of the internode of the rachilla (Fig. 3). The only sizeable group that does not have a large percentage of species where AZs are found inside the spikelet is the subfamily Panicoideae, where spikelets contain two florets, of which one or both are fertile. In this group, disarticulation of the spikelet is often equivalent to dispersal of a single seed, making the spikelet diaspore in the Panicoideae functionally equivalent to the floret diaspore in other groups.

The diversity of structures created by the initiation of the AZ in different locations within the inflorescence has led to a variety of modes of dispersal. Two trends can be discerned, one is the change in position of the zone itself, which can create new morphological possibilities, and the other is the elaboration of resulting morphological combinations to varied dispersal vectors. An example of the interplay between AZ location and elaboration of diaspore structure occurs in the “bristle clade”, a monophyletic group within the tribe Paniceae, where all species produce sterile branches in the inflorescence as well as spikelets (Zuloaga et al., 2000; Doust and Kellogg, 2002; Sato et al., 2013) (Fig. 2). In most species in the clade, the diaspore is the spikelet itself, with the AZ located underneath the glumes, but in Cenchrus the AZ has shifted to the base of the fascicle (primary branch arising from culm axis), so that the diaspore consists of a number of orders of sterile branchlets as well as one to few spikelets. Fascicles have evolved in two directions, one
One striking difference between grasses and their outgroups is the occurrence of AZs in inflorescence axes that are not those that directly bear the spikelet or floret. Shattering of the rachis axis occurs in several widely separated grass groups, including the tribe Andropogoneae (subfamily Panicoideae) and tribe Triticeae (subfamily Pooideae), two tribes who last had a common ancestor more than 60 million years ago (Vicentini et al., 2008).

Many Triticeae, along with many other pooid grasses, also exhibit shattering of the rachilla axis within the spikelet, leading to the dispersal of individual florets attached to internode segments. It is unknown whether the same developmental genetic process controls AZ formation in both the rachis and rachilla or whether the different groups that exhibit this mode of disarticulation do so using the same genetic machinery. Likewise, the functional effect of diaspores that contain both spikelet and rachis internode or floret and rachilla internode is poorly known. However, placement of the AZ to produce a diaspore that has either the basal or distal internode attached to the spikelet produces strikingly different diaspores in the wheat and barley relatives of the tribe Triticeae, with basal internodes producing a characteristic wedge-shaped diaspore (e.g., *Hordeum* spp.) and distal internodes a barrel-shaped diaspore (e.g., *Aegilops speltoides*).

Not surprisingly, there is much less information about genetic regulation of disarticulation than about its morphological diversification. While multiple QTLs have been identified from different cereal domestications, evidence remains correlative that the same loci are involved in multiple domestication events. Our analyses of approximately 20 distinct QTLs identified from the literature showed that five of these comprised QTLs from multiple species and three of these colocalize with orthologs of genes controlling
shattering. This provides some evidence for a common genetic pathway for control of disarticulation in grasses. This hypothesis is strengthened by the QTLs identified in Setaria, where both QTLs are associated with orthologs of shattering genes, one from sorghum and one from rice.

The pattern of colocalization of QTLs in Setaria with orthologs of genes identified from sorghum and rice is intriguing because Setaria is in the same subfamily as sorghum, but the AZ is located in the pedicel under the spikelet, as in a number of other grass groups, including rice. In sorghum, the AZ usually forms at the base of the sessile spikelet of the triplet spikelet unit (one sessile and two pedicillate), so that the disperose consists of the three spikelets attached to the internode above. This condition is also found in several other groups. However, there appears to be no overlap between the positions of QTL for shattering in the Triticaceae and in sorghum and other Andropogoneae (Gill et al., 2007). As well, the loci governing rachis abscission in the Triticaceae, such as btr1 and btr2 in barley (Azhaguvel et al., 2006; Sakuma et al., 2011) and Br1, Br2 (Nalam et al., 2006), and Q (Simons et al., 2006) in wheat have not been found to colocalize with QTLs for either culm or branch disarticulation in other species such as rice or sorghum (Sakuma et al., 2011).

In terms of the genetic regulation of disarticulation, the evidence for a common genetic pathway is tantalizing but incomplete. Almost all investigations of shattering to date have been concerned with identifying the major cause of shattering in individual species, hoping to find the genes responsible for controlling shattering in the wild relatives of particular domesticated crops. This approach is not ideal for illuminating a common genetic pathway, especially since the importance of any particular locus for domestication may change in the different grass lineages. There is, as yet, little direct evidence that any gene has been involved in the evolution of shattering in more than one domestication event. The evidence gathered so far has been correlative, being either the shared genomic colocalization of QTLs for shattering across multiple species (Paterson et al., 1995) or colocalization of gene orthologs with QTLs, as discussed by multiple authors (Larson and Kellogg, 2009; Lin et al., 2012; Doust et al., 2014). Therefore, the next step in understanding the evolution of disarticulation patterns in grasses and the extent to which there is a common genetic pathway is to investigate the role of orthologs of shattering genes in multiple grass species. We suggest that the three colocalizations of QTLs with shattering genes identified in this analysis are excellent starting points for such an investigation and represent testable hypotheses of the presence of a common genetic pathway for disarticulation. Several of these species have transformation systems that would enable gene knockdown and over-expression studies to test whether these genes are involved in disarticulation in multiple species.

Some evidence for widely conserved genetic elements already exists. For example, similar BEL1-like proteins are involved in AZ formation in rice (encoded by qSh1) and in the Arabidopsis silique (encoded by REPLUMLESS), and in both cases, the same cis-regulatory element has been modified (Arnaud et al., 2011). As well, when TaqSH1, the wheat ortholog of the rice gene qSh1, is overexpressed in Arabidopsis thaliana, the formation of the AZ zone is delayed in both the silique and at the base of the petals (Zhang et al., 2013). Both of these results suggest that some elements of a common genetic pathway for AZ formation may be distributed more widely in plants. In the grasses, investigation of whether known genes have effects on disarticulation across multiple lineages and the elucidation of the genetic network in which these genes act will allow us to judge the extent to which a common genetic pathway for shattering exists. In addition, this will provide new targets for control and fine-tuning of the shattering response, potentially reducing harvest losses and providing opportunities for selection in emerging domesticated crops.

LITERATURE CITED


Doust et al.—Inflorescence abscission zones in grasses


