Van der Straeten, D. A. Djudzman, W. Van Caeneghem, J. Smalle, M. Van Montagu, *Plant Physiol.* **102**, 401 (1993); G. Roman, B. Lubarsky, J. J. Keiber, M. Rothenberg, J. R. Ecker, *Genetics* **139**, 1393 (1995).
5. C. Chang, S. F. Kwok, A. B. Bleecker, E. M. Meyerowitz, *Science* **262**, 539 (1993).
6. J. B. Stock, A. M. Stock, J. M. Mottonen, *Nature* **344**, 395 (1990); J. S. Parkinson and E. C. Kofoid, *Annu. Rev. Genet.* **26**, 71 (1992).
7. J. J. Kieber, M. Rothenberg, G. Roman, K. A. Feldmann, J. R. Ecker, *Cell* **72**, 427 (1993).
8. T. Maeda, S. M. Wurgler-Murphy, H. Saito, *Nature* **369**, 242 (1994).
9. Genomic DNA isolated from *Arabidopsis* (ecotype Columbia) was partially digested with Sau 3A, then partially filled in, ligated to Xho I half-site arms of λ GEM11 (Promega), and packaged in vitro. We screened this library representing four to five genomes using the *ETR1* C5 cDNA (encoding residues 270 to 738) as a probe. We obtained 10 positive clones using a wash of 2× standard saline phosphate EDTA [SSPE; 1× is 150 mM NaCl, 10 mM NaH_2PO_4, and 1 mM EDTA] at 57°C. Five of them were reisolations of the *ETR1* gene, and the other five represented a newly discovered gene.
10. We screened approximately 1×10^5 plaque forming units of an *Arabidopsis* flower cDNA library in lambda ZAPII [D. Weigel, J. Alvarez, D. R. Smyth, M. F. Yanofsky, E. M. Meyerowitz, *Cell* **69**, 843 (1992)] at high stringency (65°C, 0.1 × SSPE), using a 1.2-kb Sac I–Sph I *ERS* genomic DNA fragment as a probe. Nine positive clones were obtained, and they were converted to plasmids by in vitro excision according to the manufacturer's instructions (Stratagene). They represented five cDNAs starting at different 5′ positions.
11. J. Hua, C. Chang, Q. Sun, E. M. Meyerowitz, unpublished data.
12. An *ERS* genomic DNA fragment was mapped by restriction fragment length polymorphism analysis in a Columbia-Landsberg mapping population, and linkage analysis was done with the MAPMAKER program [C. Chang, J. L. Bowman, A. W. DeJohn, E. S. Lander, E. M. Meyerowitz, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 6856 (1988)].
13. We determined nucleotide sequences by the dideoxynucleotide chain termination method using a Sequenase II kit (U.S. Biochemical) according to the manufacturer's instructions.
14. I. M. Ota and A. Varshavsky, *Science* **262**, 566 (1993).
15. R. V. Swanson, L. A. Alex, M. I. Simon, *Trends Biochem.* **19**, 485 (1994).
16. A full-length *ERS* gene was cloned into the Hind III and Sac I sites of plasmid pGEM7Z(+) (Promega) by ligation of a 4-kb Hind III fragment and a 2-kb Hind III–Sac I fragment. In vitro mutagenesis was done with the use of a Muta-Gene Phagemid in vitro mutagenesis kit (Bio-Rad). A primer (5′-GGAGCCTTTTTCATTCTC-3′) containing the desired A to T change was used.
17. A 5-kb Kpn I genomic DNA fragment was cloned into the Kpn I site of pCGN1547 [K. E. McBride and K. R. Summerfelt, *Plant Mol. Biol.* **14**, 269 (1990)]. *Agrobacterium tumefaciens* strain ASE [R. T. Fraley *et al.*, *Biotechnology* **3**, 629 (1985)] carrying constructs in pCGN1547 was used to transform roots of *Arabidopsis* (ecotype No-0) according to standard protocols [D. Valvekens, M. Van Montagu, M. Van Lijsebettens, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 5536 (1988); L. Marton and J. Browse, *Plant Cell Rep.* **10**, 235 (1991)].
18. We performed incubations in ethylene by placing plates or pots in a sealed chamber with an ethylene flow giving a concentration of approximately 10 μl/liter.
19. The last two rosette leaves produced by the plants before bolting were excised, prepared, and examined as described (*7*). Ten representative areas from the upper leaf epidermis per leaf were photographed. Epidermal cell numbers in each photograph were counted, and the cell concentration was deduced. Seven leaves each from *ers-1* and No-0 plants were analyzed.
20. G. L. Funke, F. DeCoeyer, A. DeDecker, J. Maton, *Biol. Jaarb.* **5**, 335 (1938); O. M. Van Andel and D. R. Verkerke, *J. Exp. Bot.* **29**, 638 (1978).
21. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
22. We thank E. Koh for carrying out plant transformations; L. Medrano for RFLP mapping; and X. Chen, S. Jacobsen, B. Krizek, J. Levin, Z. Liu, J.-L. Reichmann, M. Running, R. Sablowski, H. Sakai, and B. Williams for critical comments on the manuscript. Funded by U.S. Department of Energy grant FG03-88ER13873 to E.M.M.

12 May 1995; accepted 4 August 1995

Convergent Domestication of Cereal Crops by Independent Mutations at Corresponding Genetic Loci

Andrew H. Paterson,* Yann-Rong Lin, Zhikang Li, Keith F. Schertz, John F. Doebley, Shannon R. M. Pinson, Sin-Chieh Liu, James W. Stansel, James E. Irvine

Independent domestication of sorghum, rice, and maize involved convergent selection for large seeds, reduced disarticulation of the mature inflorescence, and daylength-insensitive flowering. These similar phenotypes are largely determined by a small number of quantitative trait loci (QTLs) that correspond closely in the three taxa. The correspondence of these QTLs transcends 65 million years of reproductive isolation. This finding supports models of quantitative inheritance that invoke relatively few genes, obviates difficulties in map-based cloning of QTLs, and impels the comparative mapping of complex phenotypes across large evolutionary distances, such as those that separate humans from rodents and domesticated mammals.

Most calories consumed by humans and livestock derive from the major cereals: rice, wheat, maize, millet, and sorghum. The cereals are members of the grass family (Poaceae) and were each domesticated from their wild relatives between 7000 and 12,000 years ago (*1*), quite recently in human history. Independent domestication of the four major cereal complexes—in Africa (sorghum and millet), Asia (rice) [but see (*1*)], the Near East (wheat, barley, oats, and rye) and America (maize)—produced similar results: In all cases, small-seeded wild grasses with natural seed dispersal were converted into large-grained symbionts that depended on farmers to harvest and sow their seed (*1*).

Although conservation of gene order along the chromosomes is well known to transgress species boundaries, the extent of correspondence in the QTLs that account for variation in complex phenotypes has been a point of conjecture. Comparative maps reveal a common order of genes and monogenic phenotypes over large chromosomal tracts in grasses after 65 million years (My) of divergence (*2*) and in mammals after 100 My of divergence (*3*). However, models for quantitative inheritance that invoke many genes, each with only tiny effects (*4*), would predict little correspondence of QTLs across taxa. Recent QTL mapping studies suggested a simpler basis for quantitative inheritance (*5*), although such studies detected only a subset of QTLs with relatively large effects (*6*). Correspondence of QTLs in congeneric plant species has been suggested (*7*), but the promiscuity of plants carries the possibility of recent genetic exchange between species.

If QTLs in disparate taxa mapped to corresponding locations more often than would be expected by chance, such a finding would strongly suggest that corresponding genes were involved in the evolution of the relevant phenotypes. To investigate this hypothesis, we assessed correspondence between QTLs that affect seed mass and disarticulation of the mature inflorescence (shattering) in crosses between cultivated and wild sorghum species, between cultivated and wild maize species, and between divergent subspecies of cultivated rice (*8*). Correspondence among short-day flowering mutations was evaluated in these and additional taxa. Use of interspecific (sorghum, maize) or subspecific (rice) crosses maximized segregation for allelic variants at both QTLs and DNA markers.

We studied three traits that were inde-

A. H. Paterson, Y.-R. Lin, Z. Li, S.-C. Liu, Department of Soil and Crop Science, Texas A&M University, College Station, TX 77843, USA.
K. F. Schertz, Agricultural Research Service, U.S. Department of Agriculture (USDA-ARS), College Station, TX 77843, USA.
J. F. Doebley, Department of Plant Biology, University of Minnesota, St. Paul, MN 55108, USA.
S. R. M. Pinson, USDA-ARS, Beaumont, TX 77713, USA.
J. W. Stansel, Texas A&M Research and Extension Center, Beaumont, TX 77713, USA.
J. E. Irvine, Texas A&M Research and Extension Center, Weslaco, TX 78596, USA.

*To whom correspondence should be addressed.

pendently altered during the domestication of grasses. The small seed mass (size) of wild grasses permits more rapid maturity, wider distribution, and deeper penetration of the soil, but seedling vigor is reduced and mechanical harvesting is more difficult relative to the larger seeds of the cereals. Wild grasses disperse seed by shattering, but nonshattering mutants were selected during domestication because humans could more efficiently harvest grains that adhered to the plant. Short-day flowering enables coordination of seed development with the season of optimal rainfall in the semiarid tropics, but it fails to maximally exploit incident solar radiation in temperate latitudes (9).

Three QTLs that affect seed mass (size) correspond closely in sorghum, rice, and maize, and at least five additional QTLs correspond between two of these genera (Fig. 1). Among seven QTLs that account for 52% of phenotypic variance explained (PVE) in sorghum seed mass, five [on linkage groups (LGs) A, C, E, F, and I] correspond to five of the eight QTLs that account for 78% of PVE in rice (*10*). Four of the sorghum QTLs (on LGs A, B, C, and F) correspond to four of the eight QTLs that account for 69% of PVE in maize (*11*). Five maize QTLs correspond to rice QTLs. Only four QTLs [two on maize chromosome 2, one on rice chromosome 3, and one on sorghum LG J (not shown)] showed no correspondence. The probabilities that seed mass QTLs in sorghum, rice, and maize would correspond so frequently by chance are conservatively estimated as 0.1 to 0.8% (*8*).

QTLs that affect seed dispersal show similar correspondence across taxa. Shattering mapped to a single locus (~100% PVE) in sorghum, three loci (24% PVE) in rice, and 10 loci (60% PVE) in maize. The discrete sorghum locus (*12*) corresponds to a rice QTL on chromosome 9 and to maize QTLs on duplicated regions of chromosomes 1 and 5 (*11*). Rice QTLs on chromosomes 2 and 3 correspond to maize QTLs on chromosomes 4 and 1. Six additional QTLs influence shattering in maize but not in rice or sorghum (Fig. 1).

The ability of many cultivated cereals to flower in the long days of the temperate summer may be largely the result of mutations at a single ancestral locus. A sorghum LG D QTL (probably *Ma1*:13) explains ~86% of PVE in flowering time and accounts for the dichotomy of F_2 phenotypes in our day-neutral (*S. bicolor*) × short-day (*S. propinquum*) cross; it also accounts for short-day flowering in each of the five races of *S. bicolor* (*13*). Short-day flowering of sugarcane is closely associated with the DNA probe *pSB188* (*8*), which lies near *Ma1*. The corresponding region of maize chromosome 10 accounts for up to 26% of PVE in the flowering of a temperate × tropical cross (*14*). The corresponding regions in wheat and barley, the short arm of the group 2 homeologs (chromosomes that do not normally pair with each other but are descended from a common ancestral chromosome), all harbor photoperiodic flowering mutants (*15*). In rice, the orthologous (directly descended from a common ancestral locus) region on chromosome 4 harbors no known flowering mutants; however, short-day flowering mutations *Se1* and *Se3* both map to a region of chromosome 6 (*16, 17*) that is orthologous to sorghum LG I and paralogous (derived by duplication and subsequent divergence from a common ancestral locus) to the sorghum LG D region of *Ma1* (Fig. 1). The *Se1*/*Se3* region of rice corresponds to a region of maize chromosome 9 that harbors QTLs that affect flowering in at least four populations (*13*). This model implies ancient duplication of regions of maize chromosomes 9 and 10 (which published data neither confirm nor contradict) and regions of rice chromosomes 4 and 6 [equivocally supported by the correspondence of *Pi-2* and *Pi-5t* genes that influence rice blast reaction (Fig. 1) (*16, 18*)]. These day-length-insensitive flowering mutations are not in any of at least three genes for phytochrome, a key regulator of photomorphogenesis (*19*).

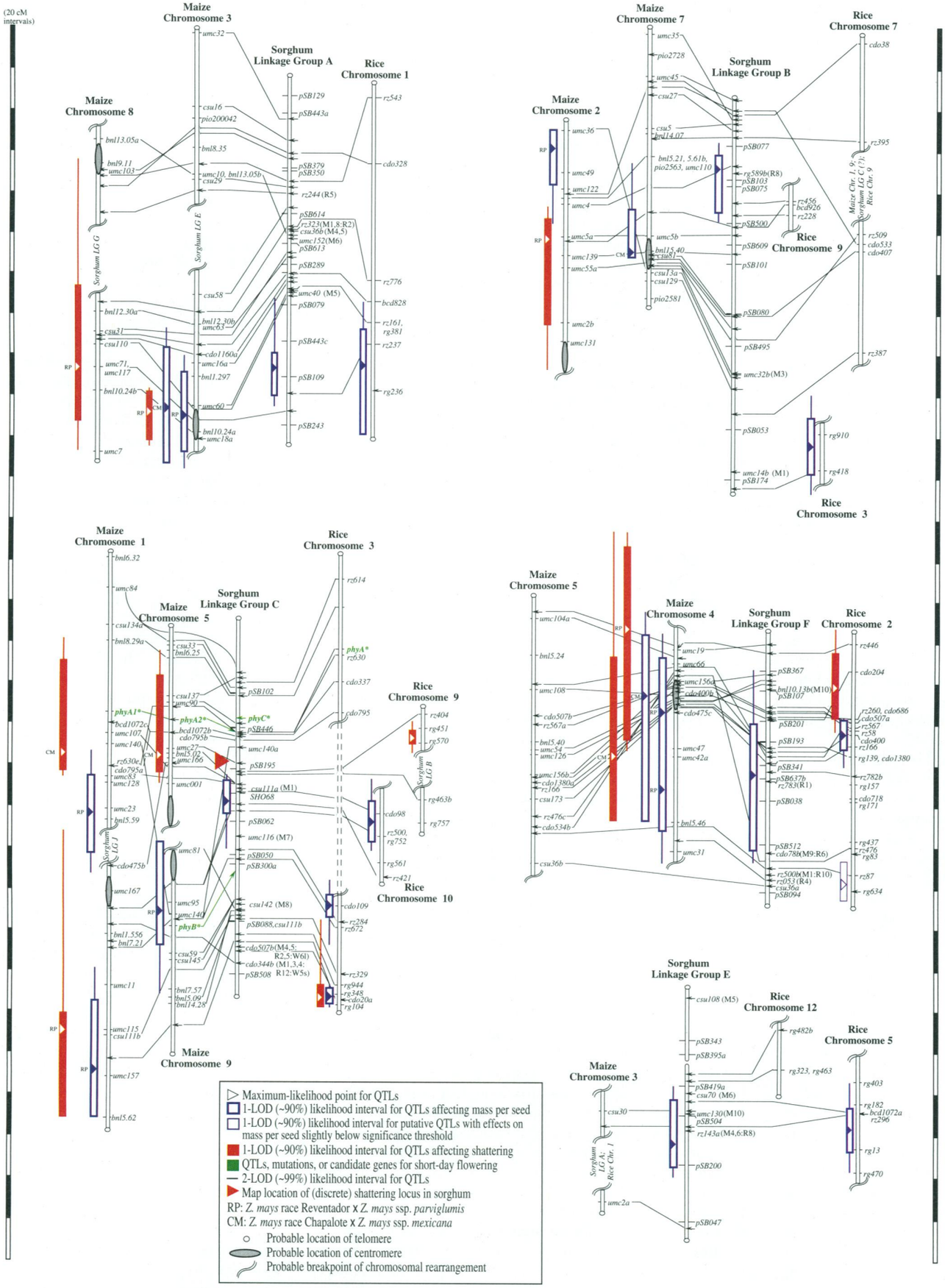
Convergent domestication of sorghum, rice, and maize appears to result from mutations at corresponding genetic loci, which would suggest that few genes with large effects determine the phenotypes studied. This finding supports punctuational evolutionary models proposed for other taxonomic lineages, such as the transformation of the berrylike ovary of wild nightshades into the tomato "fruit" (*5–7*) and the transformation of the teosinte inflorescence into the maize "ear" (*20*). Moreover, it suggests that cereal domestication may have happened rather quickly, perhaps within a century. Correspondence in the location of QTLs in different taxa does not prove that the underlying genes are identical, but does suggest that some of them may be identical; this suggestion is reinforced by the tendency of corresponding QTLs to show similar gene action (*10, 11, 21*). Domestication of other cereals such as wheat may have involved mutations at loci corresponding to those reported here. Ongoing crop improvement may be associated with yet additional mutations in these same genes or with mutations in new genes that become rate-limiting only after the initial suite of mutations is fixed.

Chromosomal duplications within taxa may partly account for polygenic inheritance, a finding which and would reinforce the theme of functional correspondence between genetic loci that have evolved independently for millions of years. Such duplications occurred in many angiosperms that now have more than six to nine gametic chromosomes (*22*), and engineered duplications have been implicated in mammalian quantitative phenotypes (*23*). Putatively duplicated maize QTLs that affect shattering are found on chromosomes 1 and 5 and on chromosomes 3 and 8, respectively. Possible duplication of QTLs that affect the height of sorghum and maize has also been reported (*13*).

The lack of correspondence of occasional QTLs across taxa may result from heterogeneity within taxa, pleiotropic effects of mutations in other traits (*8*), environmental effects (*7, 20*), differences in "genetic background" (*7, 20, 24*) [perhaps as a result of epistasis (*10*)], or false negatives in small study populations (*25*). Stringent statistical criteria make it unlikely that the QTLs reported here are false positives (*5*). These occasional incongruent QTLs may simply reflect that 65 My of divergence, followed by three independent episodes of domestication on different continents, did involve some genetic changes that were different.

Some incongruities between QTLs may also illustrate the role of contingency in evolution. For example, the order in which mutations happened to occur may influence the selective advantage accruing to subsequent mutations. The African domesticators of sorghum may simply have been fortunate to find a mutant in a critical step leading to grain abscission (*Sh1*) that "turned off" the pathway, accounting for ~100% of PVE in crosses between shattering and nonshattering types. The American domesticators of maize may not have been as lucky but still succeeded in reducing disarticulation by "pyramiding" mutations with smaller effects on several distinct steps. The Asian domesticators of rice intentionally selected for intermediate degrees of nonshattering that would reduce field losses but not preclude hand-threshing; this may account for the smaller PVE in rice than in maize or sorghum. Moreover, independent and random occurrence of mutations in related genes may have formed new alleles with very different phenotypic consequences. For example, putatively homeologous nonshattering mutations on maize chromosomes 3 and 8 explain grossly different portions of PVE in the same population (*11*).

Map-based cloning of QTLs, previously a refractory objective (*26*), appears more feasible in view of our results. Independent mutations of the same genes in different taxa provide multiple mutant alleles that can be used to seek correlations between loss-of-function phenotype and mutant genotype. Repetitive DNA element families are a common obstacle in chromosome


(20 cM intervals)
Maize Chromosome 3
Sorghum Linkage Group A
Rice Chromosome 1
Maize Chromosome 8
Maize Chromosome 7
Rice Chromosome 7
Sorghum Linkage Group B
Maize Chromosome 2
Rice Chromosome 9
Rice Chromosome 3
Maize Chromosome 1
Rice Chromosome 3
Maize Chromosome 5
Sorghum Linkage Group C
Rice Chromosome 9
Rice Chromosome 10
Maize Chromosome 9
Maize Chromosome 5
Maize Chromosome 4
Sorghum Linkage Group F
Rice Chromosome 2
Sorghum Linkage Group E
Rice Chromosome 12
Rice Chromosome 5
Maize Chromosome 3
Maximum-likelihood point for QTLs
1-LOD (~90%) likelihood interval for QTLs affecting mass per seed
1-LOD (~90%) likelihood interval for putative QTLs with effects on mass per seed slightly below significance threshold
1-LOD (~90%) likelihood interval for QTLs affecting shattering
QTLs, mutations, or candidate genes for short-day flowering
2-LOD (~99%) likelihood interval for QTLs
Map location of (discrete) shattering locus in sorghum
RP: Z. mays race Reventador x Z. mays ssp. parviglumis
CM: Z. mays race Chapalote x Z. mays ssp. mexicana
Probable location of telomere
Probable location of centromere
Probable breakpoint of chromosomal rearrangement

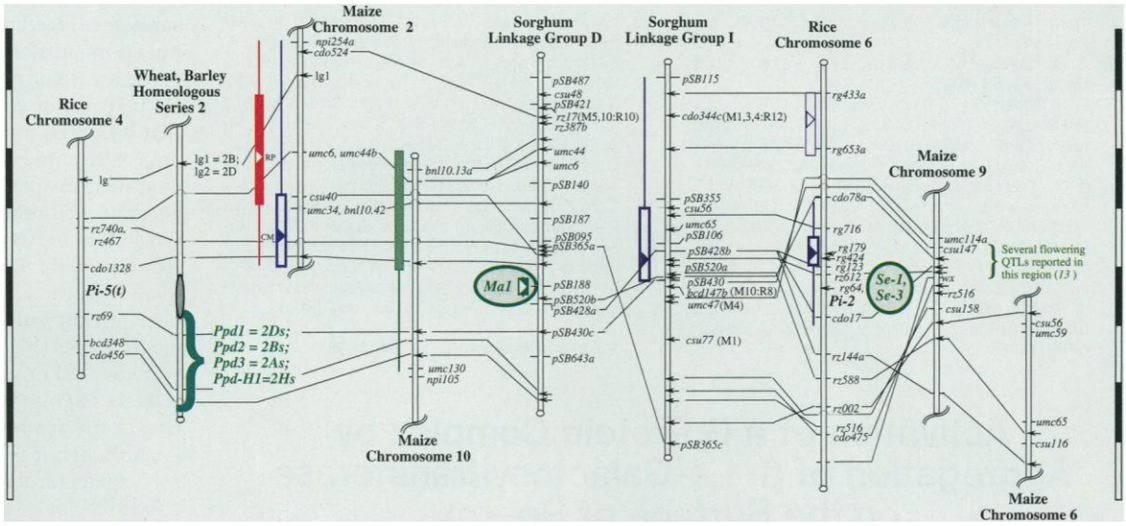


Fig. 1. Comparative mapping of QTLs that affect seed mass, shattering, and short-day (photoperiodic) flowering in sorghum, rice, maize, and other Poaceae. The genetic maps are based on orders and recombinational distances that were determined as described (*8*). DNA marker loci indicated by a line (–) were directly mapped in the cited populations; those indicated by an arrow were mapped in other populations and the appropriate locations were inferred from the map positions of flanking markers. Inferred locations are based on our own results for sorghum (*13*) and rice (*10*) and on published results for maize (*2*). For DNA markers that conflict with the most parsimonious interpretation of chromosomal correspondence between taxa, map positions are indicated in parentheses adjacent to the mapped sorghum locus. Comparative markers were mapped in as few as 56 individuals; thus, reversals in the order of closely linked loci (<3 cM) are not conclusive evidence of chromosomal rearrangement. QTL likelihood intervals (1-LOD and 2-LOD) were drawn as described (*5*). In cases where the QTL likelihood interval spanned a chromosomal rearrangement, the QTL was drawn along the chromosome that contained the likelihood peak. Correspondence among QTLs for seed mass is inferred as follows (M, maize; S, sorghum; R, rice; and c, chromosome): [M c 1 // S LG A//R c 1]; [M c 7 // S LG B]; [M c 1 // S LG C // R c 10]; [M c 1 // R c 3]; [S LG E//R c 5]; [M c 9 // R c 3]; [M c 4 // S LG F // R c 2]; and [S LG I // R c 6]. Correspondence among QTLs for shattering is inferred as follows: [M c 3 // M c 8]; [M c 1 // M c 5 // S LG C // R c 9]; and [M c 4 (RP only) // R c 2].

walks, but because they are often species-specific, "parallel walks" to corresponding genes in different species may bridge such obstacles. Choosing a system that shows discrete variation (*27*), such as shattering in sorghum, simplifies a complex trait more quickly than is possible with breeding approaches (*6*, *20*). Finally, the quantity of DNA in the maize genome is about four times that of sorghum and about six times that of rice; hence, maize QTLs might be most expediently cloned by first isolating the corresponding sorghum or rice genes.

Our results impel comparative mapping of complex phenotypes in many other biota. For example, genes that affect diabetes, hypertension, obesity, alcohol and drug addiction, and other medically important phenotypes have been mapped in mouse, rat, pig, cow, and sheep (*28*); our findings support the often tacit assumption that these genes may be relevant to corresponding human phenotypes, despite 100 My of divergence (*3*). The conservation of gene order provides a framework for the comparative analysis of complex phenotypes in mammals and other biota, as shown herein for the cereals.

REFERENCES AND NOTES

1. J. R. Harlan, *Crops and Man* (American Society of Agronomy, Madison, WI, 1975); J. M. J. deWet, *Ann. Mo. Bot. Gard.* **68**, 87 (1981); N.W. Simmonds, *Evolution of Crop Plants* (Longman, Essex, UK, 1976), pp. 98, 112, 128.
2. S. H. Hulbert *et al., Proc. Natl. Acad. Sci. U.S.A.* **87**, 4251 (1990); S. Ahn and S. D. Tanksley, *ibid.* **90**, 7980 (1993); S. Ahn *et al., Mol. Gen. Genet.* **241**, 483 (1993); J. L. Bennetzen and M. A. Freeling, *Trends Genet.* **9**, 259 (1993); T. Helentjaris, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 8308 (1993); R. Shields, *Nature* **365**, 297 (1993); N. Kurata *et al., Biotechnology* **12**, 276 (1994).
3. S. J. O'Brien *et al., Nature Genet.* **3**, 103 (1993).
4. R. Lande, *Genet. Res.* **26**, 221 (1975); M. Turelli, *Theor. Popul. Biol.* **25**, 138 (1984).
5. A. H. Paterson *et al., Nature* **335**, 721 (1988); E. S. Lander and D. Botstein, *Genetics* **121**, 185 (1989).
6. A. H. Paterson, J. DeVerna, B. Lanini, S. D. Tanksley, *Genetics* **124**, 735 (1990).
7. A. H. Paterson *et al., ibid.* **127**, 181 (1991); C. A. Fatokun, D. I. Menacio-Hautea, D. Danesh, N. D. Young, *ibid.* **132**, 841 (1992).
8. The sorghum, rice, and maize crosses, plant care, and lab techniques are described elsewhere (*10–13*). Seed mass was calculated from the average mass of 1000 (sorghum), 100 (rice), or 10 (maize) seed. Disarticulation was measured as described for sorghum (*11*) and maize (*12*); for rice, each of 2418 F_2-derived F_4 plants were scored 0 (nonshattering), 1, or 2 (highly shattering) and F_2 family means were calculated. Genetic maps are as described for sorghum (*13*), rice (*10, 16*), and maize [E. H. Coe and M. G. Neuffer, in *Genetic Maps,* S. J. O'Brien, Ed. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, ed. 6, 1993), pp. 6.157–6.189]. The hypergeometric probability function [R. J. Larsen and M. L. Marx, *An Introduction to Probability and its Applications* (Prentice-Hall, Englewood Cliffs, NJ, 1985), pp. 152–160] affords evaluation of the correspondence between seed mass QTLs (J. Taylor, personal communication), as described (*13*). A match was declared when QTL likelihood intervals (1-LOD) for two taxa overlapped. Estimates of correspondence exclude a seed mass QTL on sorghum LG D, which may be a pleiotropic effect of *Ma1*: this eliminated a match with maize. Further, likelihood intervals for two sorghum QTLs (LGs F and I) were too long to be attributable to single genes; each spanned one rice QTL with a significant effect (chromosomes 2 and 6) and a putative QTL below the significance threshold (LODs of 2.15 and 2.02). The maize chromosome 4 QTL corresponding to sorghum LG F also showed two likelihood peaks that could not be resolved (*12*). Correspondence of diffuse likelihood intervals across taxa supports the suggestion (*5*) that they span multiple QTLs. We have conservatively counted such intervals as including only one QTL in each taxon. The two sugarcane populations were each crosses between heterozygous plants that differed in photoperiodic flowering response. In *S. officinarum* cv. Green German × *S. spontaneum* cv. IND 81-146, IND alleles at three of four segregating loci (pSB188aI, bI, or cI) were associated with flowering (P = 0.020 to 0.026, χ^2 independence tests). In *S. spontaneum* cv. PIN 84-1 × *S. officinarum* cv. Muntok Java, Muntok alleles at one of two segregating loci (pSB188aM) were associated with flowering (P = 0.025).
9. J. L. Harper, *Plant Population Biology* (Academic Press, London, 1977).
10. Z. Li, S. R. M. Pinson, W. D. Park, A. H. Paterson, J. Stansel, in preparation.
11. J. Doebley *et al., J. Hered.* **85**, 191 (1994).
12. A. H. Paterson *et al., Proc. Natl. Acad. Sci. U.S.A.* **92**, 6127 (1995).
13. Y. R. Lin, K. F. Schertz, A. H. Paterson, *Genetics* **141**, 391 (1995).
14. R. P. Koester *et al., Crop Sci.* **33**, 1209 (1993).
15. D. A. Laurie, N. Pratchett, J. H. Bezant, J. W. Snape, *Heredity* **72**, 619 (1994).
16. M. A. Causse *et al., Genetics* **138**, 1251 (1994).
17. D. Mackill, M. Salam, Z. Wang, S. Tanksley, *Theor. Appl. Genet.* **85**, 536 (1993).
18. T. Kinoshita and M. Takahashi, *J. Fac. Agric. Hokkaido Univ.* **65**, 1 (1991).
19. All three phytochrome genes that have been isolated in the Poaceae (*phyA-C*) map to different chromosomes from all mapped short-day flowering loci (Fig. 1; sorghum LG C) (M.-M. C. Pratt, G. Kochert, Y.-R. Lin, A. H. Paterson, L. Pratt, in preparation).
20. J. Doebley *et al., Proc. Natl. Acad. Sci. U.S.A.* **87**, 9888 (1990); J. Doebley and A. Stec, *Genetics* **134**, 559 (1993); J. Dorweiler, A. Stec, J. Kermicle, J. Doebley, *Science* **262**, 233 (1993).
21. A. H. Paterson *et al.*, in preparation.
22. G. L. Stebbins, *Science* **152**, 1463 (1966); J. Masterson, *ibid.*, **264**, 421 (1994).
23. O. Smithies and H.-S. Kim, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 3612 (1994).
24. S. D. Tanksley and J. D. Hewitt, *Theor. Appl. Genet.* **75**, 811 (1988).
25. W. Beavis, *Proceedings of the 49th Annual Corn and Sorghum Research Conference* (American Seed

Trade Association, Washington, DC, 1994), pp. 250–266.
26. S. D. Tanksley, M. W. Ganal, G. B. Martin, *Trends Genet.* **11**, 63 (1995).
27. D. S. Robertson, *J. Theor. Biol.* **117**, 1 (1985).
28. P. Hilbert *et al.*, *Nature* **353**, 521 (1991); H. J. Jacob *et al.*, *Cell* **67**, 213 (1991); M. L. Rise, W. N. Frankel, J. M. Coffin, T. N. Seyfried, *Science* **253**, 669 (1991); J. A. Todd *et al.*, *Nature* **351**, 542 (1991); M. Georges *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 1058 (1993); L. Andersson *et al.*, *Science* **263**, 1771 (1994); N. Cockett *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 3019 (1994); J. C. Crabbe, J. K. Belknap, K. J. Buck, *Science* **264**, 1715 (1994); Y.-S. Ho, *Am. J. Physiol.* **266**, L319 (1994); T. W. Kurtz, *Lancet* **344**, 167 (1994); M. Paul *et al.*, *Annu. Rev. Physiol.* **56**, 811 (1994).

29. We thank J. Taylor, M. Pratt, L. Pratt, and G. Kochert for unpublished results; N. Borlaug, S. Davis, K. Feldmann, and J. Wendel for comments; and E. Coe, J. Gardiner, M. Sorrells, and S. Tanksley for probes. Supported by USDA grants 92-37300-6570 (A.H.P. and K.F.S.) and 93-37300-8773 (J.F.D.), Pioneer Hibred (A.H.P. and K.F.S.), American Sugar Cane League, Australia Sugar Research and Development Corp., Cenicana, Copersucar Technologies Corp., Florida Sugar Cane League, Hawaiian Sugar Planters Association, Mauritius Sugar Industry Research Institute (A.H.P. and J.E.I.), Texas Higher Education Coordinating Board (A.H.P.), USDA-ARS (K.F.S. and S.R.M.P.), and Texas Agricultural Experiment Station (A.H.P., Z.L., J.E.I., and J.W.S.).

27 April 1995; accepted 27 July 1995

Activation of a G Protein Complex by Aggregation of β-1,4-Galactosyltransferase on the Surface of Sperm

Xiaohai Gong,* Daniel H. Dubois,* David J. Miller,† Barry D. Shur‡

Fertilization is initiated by the species-specific binding of sperm to the extracellular coat of the egg. One sperm receptor for the mouse egg is β-1,4-galactosyltransferase (GalTase), which binds O-linked oligosaccharides on the egg coat glycoprotein ZP3. ZP3 binding induces acrosomal exocytosis through the activation of a pertussis toxin–sensitive heterotrimeric guanine nucleotide–binding protein (G protein). The cytoplasmic domain of sperm surface GalTase bound to and activated a heterotrimeric G protein complex that contained the $G_{i\alpha}$ subunit. Aggregation of GalTase by multivalent ligands elicited G protein activation. Sperm from transgenic mice that overexpressed GalTase had higher rates of G protein activation than did wild-type sperm, which rendered transgenic sperm hypersensitive to their ZP3 ligand. Thus, the cytoplasmic domain of cell surface GalTase appears to enable it to function as a signal-transducing receptor for extracellular oligosaccharide ligands.

The species-specific binding of sperm to the egg coat initiates a series of events that culminates in the formation of the zygote. In the mouse, sperm recognition of the egg coat requires the binding of GalTase on the sperm plasma membrane to O-linked oligosaccharide ligands on the zona pellucida glycoprotein ZP3 (*1*). After egg activation, the block to polyspermy is facilitated by the release of hexosaminidase from the egg cortical granules, which removes the binding site on ZP3 oligosaccharides for sperm GalTase (*1*).

GalTase is a biosynthetic component of the Golgi complex in somatic cells. However, in both somatic cells and sperm, GalTase is also expressed on the cell surface, where it functions as a receptor for extracellular oligosaccharide ligands. GalTase's dual subcellular distribution results, at least in part, from the synthesis of two GalTase isoforms with different cytoplasmic domains (*2*). The shorter isoform, which contains an 11–amino acid cytoplasmic domain, functions biosynthetically in the Golgi complex; the longer isoform contains an additional 13–amino acid sequence that overrides the Golgi retention signal and targets some GalTase molecules to the cell surface. The longer GalTase isoform mediates the binding of sperm to ZP3 oligosaccharides in the zona pellucida and is the only isoform found in mature sperm (*1*, *3*). Two other ZP3-binding proteins have been identified on mouse sperm in addition to GalTase. One appears to be a cell surface protein with intrinsic hexokinase activity (*4*); the other behaves as a peripheral membrane protein (*5*).

The binding of sperm to ZP3 induces the acrosome reaction, in which hydrolytic enzymes are released that aid penetration of sperm through the zona pellucida. The acrosome reaction appears to be induced by multivalent oligosaccharides on ZP3 that aggregate sperm-bound receptors, thus eliciting intracellular signals. ZP3 glycopeptides that bind to sperm are unable to induce the acrosome reaction unless they are cross-linked by antiserum to ZP3 (*6*). Induction of the acrosome reaction by ZP3 is inhibited by pertussis toxin (PTx), which suggests that a G protein complex that contains G_i or G_o is required (*7*). A $G_{i\alpha}$ subunit has been identified in sperm, as assessed by PTx-dependent adenosine diphosphate (ADP) ribosylation and immunoblotting of a 41-kD sperm protein; mature sperm do not appear to contain the G_o subunit (*7*). Other signal transduction pathways elicited by ZP3 or by progesterone (*8*) that involve calcium mobilization (*9*) or phosphorylation of tyrosine residues (*4*) also likely contribute to induction of the acrosome reaction.

Although it is clear that sperm GalTase participates in fertilization by binding glycoside ligands on ZP3, and that ZP3 binding elicits a $G_{i\alpha}$ cascade to induce the acrosome reaction, the relation between sperm GalTase and activation of the $G_{i\alpha}$ cascade remains unknown. We determined whether sperm GalTase binds to and activates a $G_{i\alpha}$ cascade by testing four predictions. Previous studies showed that intact antibodies to sperm GalTase (anti-GalTase) induce the acrosome reaction whereas Fab fragments do not unless they are cross-linked with a secondary antibody (*10*). Thus cross-linking GalTase mimics the effect of ZP3 binding. Because ZP3-induced acrosome reactions are inhibited by PTx (*7*), the first prediction to be tested is that acrosome reactions induced by anti-GalTase might also be sensitive to PTx, which would suggest the involvement of a $G_{i\alpha}$ cascade. Second, if GalTase functions through a PTx-sensitive cascade, then the cytoplasmic domain of GalTase might associate, directly or indirectly, with a $G_{i\alpha}$ subunit as well as with the βγ subunit. Third, because ZP3 binding to sperm membranes induces G protein activation (*11*), the aggregation of sperm GalTase by anti-GalTase should also elicit G protein activation. Finally, transgenic sperm that have elevated amounts of surface GalTase (*12*), and which bind more ZP3 ligand than do wild-type sperm, should show enhanced G protein activation.

We tested the ability of PTx to inhibit acrosomal exocytosis induced by anti-GalTase (Fig. 1). Positive controls included sperm treated with ionophore or with solubilized zona glycoproteins. Anti-GalTase and zona glycoproteins induced the acrosome reaction to a similar degree. Neither preimmune immunoglobulin G (IgG) nor anti-GalTase Fab fragments had any significant effect on acrosomal exocytosis (Fig. 1) (*10*). The addition of PTx inhibited zona glycoprotein–induced and anti-GalTase–induced acrosome reactions to background rates. Because PTx inhibits only "physiological" acrosome reactions, that is, those induced by a ZP3-dependent cascade (*7*), these

Department of Biochemistry and Molecular Biology, Box 117, University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030, USA.

*These authors contributed equally to this work.
†Present address: Department of Animal Sciences, University of Illinois, Urbana, IL 61801, USA.
‡To whom correspondence should be addressed.