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, and James E. Irvine
Convergent Domestication of Cereal Crops by Independent Mutations at Corresponding Genetic Loci


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-
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 5 // M c 8]; [M c 1 // M c 5 // S LG C/R c 9]; and [M c 4 (FP 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 oft-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
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) seeds. Dissection was measured as described for sorghum (11) and maize (12); for rice, each of 2415 F2-derived F1 plants were scored 0 (nonshattering), 1, or 2 (highly shattering) and F2 family mean were calculated. Genetic maps are as described for sorghum (13), rice (10), and maize (11). Experiments were performed in two laboratories (H. C. 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. 6:157–6:169). The hypergeometric probability function [H. J. Larsen and L. M. 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 Ma4; 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 5) and a putative QTL below the significance threshold (LGs D 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 (3) 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 crossed between heterologous plants that differed in photo- and periodic flowering response. In S. officinarum cv. Green German × S. spontaneum cv. IND 81-146, IND alleles at three of four segregating loci (gS118Sbl, b, or c1) were associated with flowering (F = 0.021, 0.025, chi square independence tests). In S. spontaneum cv. PIN 84-1 × S. officinarum cv. Mun- tok Java, Muntk alleles at one of two segregating loci (gS11SblBar) were associated with flowering (F = 0.025).
19. All three phytochrome genes that have been isolated in the Poaceous phyA-C) map to different chromosomes from all mapped short-day flowering loci (Fig. 1; sorghum LG C) [M.-C. Pratt, G. Koehert, V.-R. Lin, A. H. Paterson, L. Pratt, in preparation].
25. W. Beavis, Proceedings of the 49th Annual Corn and Sorghum Research Conference (American Seed


27 April 1995; accepted 27 July 1995