Comparative physical mapping links conservation of microsynteny to chromosome structure and recombination in grasses

John E. Bowers*, Miguel A. Arias*, Rochelle Asher*, Jennifer A. Avise*, Robert T. Ball*, Gene A. Brewer*, Ryan W. Buss*, Amy H. Chen*, Thomas M. Edwards*, James C. Estill*, Heather E. Exum*, Valorie H. Goff*, Kristen L. Herrick*, Cassie L. James Steele*, Santhosh Karunakaran*, Gmerice K. Lafayette*, Cornelia Lemke*, Barry S. Marler*, Shelley L. Masters*, Joana M. McMillan*, Lisa K. Nelson*, Graham A. Newsome*, Chike C. Nwakanma*, Rosana N. Odeh*, Cynthia A. Phelps*, Elizabeth A. Rarick*, Carl J. Rogers*, Sean P. Ryan*, Keimun A. Slaughter*, Carol A. Soderlund†, Haibao Tang*, Rod A. Wing‡, and Andrew H. Paterson*§

*Plant Genome Mapping Laboratory, University of Georgia, Athens, GA 30602; and †Arizona Genomics Computational Laboratory and ‡Arizona Genomics Institute, University of Arizona, Tucson, AZ 95721

Edited by Ronald L. Phillips, University of Minnesota, St. Paul, MN, and approved July 28, 2005 (received for review March 22, 2005)

Nearly finished sequences for model organisms provide a foundation from which to explore genomic diversity among other taxonomic groups. We explore genome-wide microsynteny patterns between the rice sequence and two sorghum physical maps that integrate genetic markers, bacterial artificial chromosome (BAC) fingerprints, and BAC hybridization data. The sorghum maps largely tile a genomic component containing 41% of BACs but 80% of single-copy genes that shows conserved microsynteny with rice and partially tile a nonsyntenic component containing 46% of BACs but only 13% of single-copy genes. The remaining BACs are centromeric (4%) or unassigned (8%). The two genomic components correspond to cytologically discernible ''euchromatin'' and ''heterochromatin.'' Gene and repetitive DNA distributions support this classification. Greater microcolinearity in recombinogenic (euchromatic) than nonrecombinogenic (heterochromatic) regions is consistent with the hypothesis that genomic rearrangements are usually deleterious, thus more likely to persist in nonrecombinogenic regions by virtue of Muller's ratchet. Interchromosomal centromeric rearrangements may have fostered diploidization of a polyploid cereal progenitor. Model plant sequences better guide studies of related genomes in recombinogenic than nonrecombinogenic regions. Bridging of 35 physical gaps in the rice sequence by sorghum BAC contigs illustrates reciprocal benefits of comparative approaches that extend at least across the cereals and perhaps beyond.

comparative genomics | *Oryza* | synteny

The grasses (Poaceae) provide most of mankind's caloric intake and a growing share of our fuel. The best-studied grasses, leading cereal crops, shared a common paleopolyploid ancestor ≈42–47 million years ago (mya) (1). Cereals show much colinearity of genetic maps and often have important traits controlled by quantitative trait loci at corresponding locations (2). Despite these similarities, the cereals have diverged remarkably in genome size from ≈430 million base pairs (MBP) in rice (3) to 15,966 MBP in wheat (3), largely due to differential repetitive DNA amplification and elimination.

As a model for tropical grasses, sorghum [*Sorghum bicolor* (SB)] is a logical complement to rice (*Oryza*), in that it has biochemical and morphological specializations to improve carbon assimilation at high temperatures (C4 photosynthesis). By contrast, rice uses C3 photosynthesis more typical of temperate grasses. The ≈760-MBP (3) sorghum genome is a logical bridge to the ≈2,500-MBP (3) maize genome, and the ≈4,000-MBP (3) genome of sugarcane, the world's leading biomass/biofuels crop. Sorghum shared common ancestry with maize (12 mya) and sugarcane (5 mya), much more recently than rice (42–47 mya). The most recent whole-genome duplication in sorghum appears to be ≈70 mya (1) vs. ≈12 mya in maize (4) and <5 mya in sugarcane (5), promising a higher success rate in relating sorghum genes to phenotypes by knockouts than either maize or sugarcane genes. Comparison of SB and closely related *Sorghum propinquum* (SP) promises to shed new light on genes related to domestication (2) and expedite dissection of the genetic control of key features of "Johnsongrass" (*Sorghum halepense*), a polyploid descendant of SB and SP that is one of the world's most noxious weeds (6).

We describe two genetically anchored *Sorghum* physical maps. Anchoring of thousands of genetically mapped sequence-tagged sites (STSs) permits their detailed comparison with the rice sequence, to explore levels and patterns of conserved microsynteny on a genome-wide scale. Microsynteny is conserved in euchromatin but not in heterochromatin. A strong association of conserved microsynteny with recombinogenicity suggests that intergenic gene rearrangement (i.e., changes in gene order or location in the genome) is generally deleterious. Hybridization of conserved STS markers to bacterial artificial chromosome (BAC) libraries for other cereal and noncereal monocots may leverage conserved microsynteny to study their euchromatin. However, the high levels of colinearity suggested by cross-species comparisons of genetic recombination maps mask much rearrangement in recombinationally inert (heterochromatic) regions. Such regions may require physical rather than genetic characterization, with little guidance from cross-species data.

Materials and Methods

Probe Selection, Design, and Hybridization. Overgo (7) probes were designed from sequences of interest by BLAST comparison against all plant species to identify the most-conserved 40-bp segment, using a Microsoft VISUAL BASIC program. Repetitive probes were filtered by comparison with highly repetitive sorghum DNA (8) and The Institute for Genome Research (TIGR) grass repeat database.

Overgo probes were radioactively labeled, then hybridized in pools of up to 24 probes (7). Most experiments consisted of 576 probes, applied to macroarrays of 18,432 BACs each, in 72 hybridizations of 24 probes, each using row–column–diagonal pools.

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: mya, million years ago; MBP, million base pair; SB, *Sorghum bicolor*; SP, *Sorghum propinquum*; BAC, bacterial artificial chromosome.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. BZ691020–BZ697099, CL696780–CL705752, and CW784313–CW7926311).

§To whom correspondence should be addressed. E-mail: paterson@uga.edu.

Table 1. Overgo probes applied to monocots

Taxon	Genome coverage	No. of BACs	No. of overgos tested	Scored	Worked	No. of BACs anchored	Average no. of anchors per BAC
Zea mays	12.7×	221,184	2,623	2,623	2,161	85,437	0.39
SP	9.0×	55,296	7,573	7,143	5,683	148,758	2.69
SB cv. BT×623	11.0×	73,728	6,470	6,349	5,147	139,434	1.89
O. sativa cv. Nipponbare	20.0×	73,728	2,297	2,297	1,590	29,064	0.39
Saccharum spp. Cv. R570	14×*	92,160	5,974	5,574	4,639	80,535	0.87
Banana	6.0×	36,864	1,276	1,276	679	13,059	0.35
Total		571,392	30,607	26,200	19,899	476,254	0.83

Data are available at: www.plantgenome.uga.edu/bacman/BACManwww.php. The SB library is from cultivar BTX623. Others are as cited in the text.

*The sugarcane BAC library contains ≈14× coverage of the basic chromosome set, which equates to ≈1.3× coverage of each locus in the dodecaploid source cultivar.

Labeled filters were exposed to x-ray film for 2 weeks. Films were manually scored onto transparencies, which were scanned into a computer and read by using text recognition software (ABBYY FINEREADER 5.0, ABBYY Software, Moscow). Data were transferred into the database system "BACMan" (www.intl-pag.org/pag/11/abstracts/C01_C4_XI.html), which was used to deconvolute multiplexed results into individual probe to BAC assignments.

Physical Map Assembly. Physical maps of both SP and SB BACS were constructed by using HindIII agarose fingerprinting (9). Gel images were processed by using IMAGE (10) and assembled into contigs by using FPC, Ver. 7.2 (11), also incorporating overgo data. Automated assembly was performed with a cutoff of 10^{-11}, tolerance 7 (CpM, contigs plus markers) of 10^{-10}, 10^{-9}, and 10^{-8}. All contigs containing "Q" (questionable) clones were manually edited and split at higher thresholds until none remained. Two manual merge cycles (at cutoff 10^{-10} and 10^{-9}, respectively) were applied to contigs that were adjacent based on either the sorghum map or rice colinearity. Approximately 100 fingerprinting lanes believed (based on high band numbers or formation of "Q" contigs) to represent multiple and/or chimeric clones were voided.

BAC End Sequencing. A subset of contig-terminal SP BACs from a preliminary FPC assembly were end-sequenced as described (12), yielding 15,103 sequences.

Comparison of Sorghum with Rice. Sorghum contigs were anchored to a rice genome assembly (Ver. 2.0, www.tigr.org), based in part on sequence described elsewhere (13–15). Probe sequences were masked for repeats by using REPEATMASKER (www.repeatmasker.org) and compared with rice using BLAST (identifying the best hit at $e \leq 10^{-8}$ and using mismatch penalties -2 vs. the default of -3 to detect similarities down to nearly 66%). Based on these criteria, sequences useful in synteny analysis included 42.4% of hybridization probes detecting five contigs or fewer and 16.8% (2,625 of 15,103) of BAC ends showing no match to repetitive sequences and blasting to rice locations hit by <10 other BAC ends. Sorghum contigs were considered aligned to rice if they contained two or

Table 2. Chromosomal regions show different concentrations of repetitive probes

Category	Anchored to centromeres	Anchored to sorghum genetic map but not rice	Not anchored	Mitocondria or chloroplast	Syntenic to rice	BACs not in contigs	Total BACs
All BACs (SP physical map)	4.2	2.8	43	0.5	41	8.3	40,957
All contigs (SP physical map)	9.9	2.2	54	0.1	34	0	1,541 contigs
All BACs (SB physical map)	6.7	2.7	39	0	48	3.8	69,545
All contigs (SB physical map)	3.4	2.1	45	0.6	31	17.2	1,869 contigs
One locus probe (2,223 probes)	0.2	1.8	11	2.7	80	3.9	10,356
Ten locus probes (186 probes)	2.1	2.9	40	0	51	4	2,038
Probe pHind22	47.4	2.3	17	0.9	13	19.5	1,055
Probe SOG6774 (Sau3A10)	70	0.1	5.2	0.1	3.3	20.9	1,170
Probe SOG6773 (Sau3A9)	62.2	1.5	14	0.2	5.2	17.2	860
Probe CEN38	52.2	0.6	12	0.9	11	23.2	1,158
Probe SOG6741	1.5	6.4	70	0	18	3.9	1,024
Probe SOG6293	8.6	5.7	65	0	19	1.4	867
Probe SOG6066	2.9	3.5	73	0	19	1.7	812
Probe SOG2844	0.5	9.5	70	0	18	2.3	387
Probe SOG1581	0	1.8	17	0	79	2.7	664
Probe SOG3319	0	0.5	18	0.2	77	4.1	699
Probe SOG6772 (pHind12)	4.2	3.3	44	0	45	3.4	5,521
Probe RET6	3.4	2.1	45	0.6	31	17.2	2,038
Probe SOG6635	6.7	2.7	39	0	48	3.8	1,669
Probe SOG6667	2.1	2.9	40	0	51	4	644

Percentages of BACs and contigs from various genomic contexts (indicated at the top), that hybridized to different probes (or groups of probes, at the left). Unanchored contigs, and contigs anchored to centromeres, are presumed to be largely heterochromatin. Contigs syntenic to rice are enriched for euchromatin, as detailed in the text. Probe specific data correspond to the SP physical map. One locus and 10 locus probes are hybridized to 1 and 10 contigs, respectively.

GENETICS

more sequences with best-matching rice locations >5 but <500 kbp apart.

Intragenomic Analysis of Rice. A total of 59,712 putative rice genes annotated by The Institute for Genome Research (Ver. 2) were downloaded. A subset of 46,922 (44,766 transcription units) non-transposable element-related genes were blasted against one another, and 163,043 unique matches at $e < 10^{-6}$ used in synonymous substitution rate (K_s) calculations, as described (1).

Results

Assembly of a *Sorghum* Physical Map. Because the 2,514-locus *Sorghum* sequence-tagged site-based genetic map (16) is nearly saturated, with an average of 2.5 markers per centimorgan, further resolution of gene arrangement focused on BAC-based physical mapping. Overgo probes (7) corresponding to genetically mapped sequences, genes of interest, genes in corresponding rice regions of interest, or BAC end sequences were applied to various subsets of 571,392 BACs from SB, SP (17), maize (18), rice (19), sugarcane (20), and banana (21), yielding 476,254 anchors. There were an average of more than two markers per BAC in *Sorghum* (Table 1). Approximately 87% of overgos from *Sorghum* sequences detected BACs in one or both *Sorghum* libraries. Heterologous probes yielded lower success rates.

Overgo data were combined with fingerprints of both *Sorghum* BAC libraries (9) to assemble FPC contigs (10). A total of 40,957 SP BACs (7× genome coverage) were assembled into 1,541 contigs (averaging 24.4 BACs and 12.5 hybridization loci per contig), whereas 69,545 SB BACs (11× coverage) were assembled into 1,869 contigs (35.6 BACs, 10.1 hybridization loci). Singletons comprising 3,400 (8.3% of) SP BACs were gene-poor (3.9% of 1- and 4% of 10-locus probe anchors, Table 2) but rich in centromeric repeats (19–23% of all centromeric anchored BACs; Table 2). Most (81%) singletons had ≤14 fingerprint bands (vs. 3.5% of contigged BACs). BACs largely composed of tandem repeats with multiple copies of the same band size or small BACs with few bands cannot be reliably assembled by FPC. Many singletons may derive from such genomic regions that are recalcitrant to assembly. A sample of ordered sorghum FPC contigs is aligned to corresponding regions of the rice sequence and maize (22) contigs (Fig. 1).

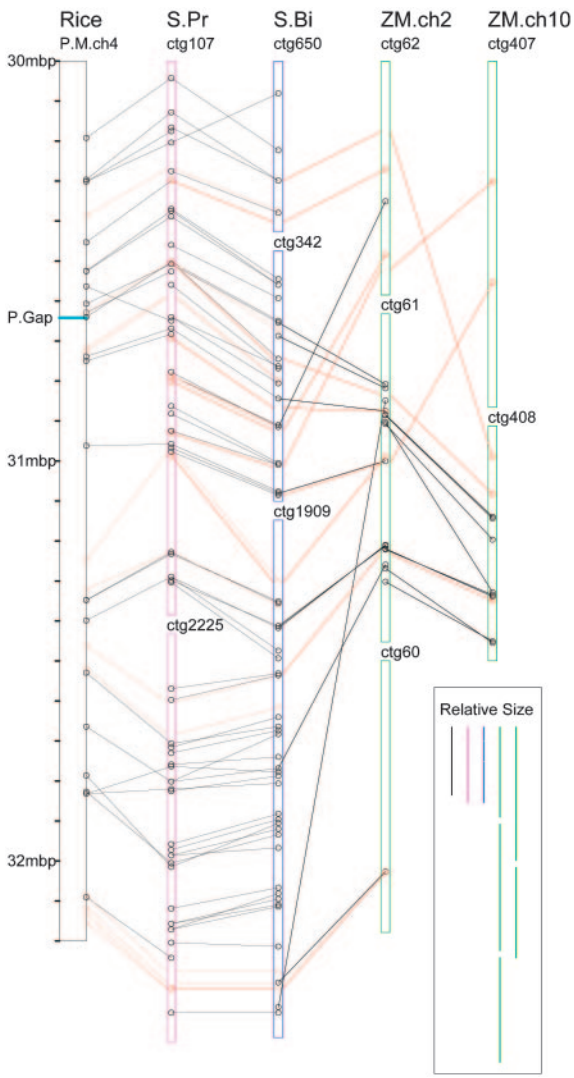

Fig. 1. Comparative physical maps. A segment of the rice chromosome 4 pseudomolecule is compared with BAC contigs from SP, SB, and *Zea mays* based on hybridization anchors (Table 1). The scales (MBP) of the rice pseudomolecule and sorghum contigs are equal, whereas maize contigs are shown at a 1:5 scale. Sorghum and maize contig lengths were estimated by multiplying the average FPC band size (4,740 bp, the observed average for rice) by the length of the contig in FPC consensus band (CB) units. Red lines represent cases for which loci were inferred in one genome (where no dot is shown) due to missing data. Maize contigs were from a recent release (www.genome.arizona.edu/fpc/maize, release 10/25/04), incorporating hybridization anchors from Table 1. Current sorghum contig assemblies are on-line (www.plantgenome.uga.edu/projects.htm).

Conserved–Sorghum–Rice Microsynteny. From sequences of 5,683 hybridization loci and 15,103 BAC ends, 1,644 and 903, respectively, could be used to anchor SP contigs to rice. An additional 35 SP and 42 SB contigs could be anchored to the sorghum genetic map by single-copy probes (16) but not to the rice pseudomolecule.

A total of 524 (34%) SP contigs containing 41% of BACs and a remarkable 80% of the 2,223 single-copy probes (that hybridized only to BACs from a single contig) could be aligned to rice, covering 39% of the sequence. The aligned sorghum contigs are only ≈30% longer than corresponding regions of rice, suggesting that most of the 70% difference in genome size between these taxa occurs elsewhere. The corresponding rice regions are high in EST-verified genes and low in repetitive DNA (23) (Fig. 2).

The chromosomal distribution of conserved sorghum–rice microsynteny was striking, virtually excluding cytologically defined heterochromatin (23) and pericentromeric regions (24, 25) (Fig. 2). This exclusion was not due to lack of coverage; 153 SP and 194 SB contigs are composed of >40% of BACs that hybridize to centromeric repeats [pHind22 (26), CEN38 (27), Sau3A9 (26), and Sau3A10 (26)]. Although the pericentromeric region is generally repeat-rich (Fig. 2), repetitive probe hybridizations reveal complex differences between the aligned contigs and 700 (54%) SP contigs that could not be aligned to the rice sequence; these nonaligned contigs include 46% of BACs but only 13% of single-copy probes. Based on SP data (which are the most complete), 13 probes that each hybridized to >100 BACs are strongly enriched in syntenic contigs (e.g., SOG1581 and SOG3319; Table 2), perhaps representing gene-associated repeats. A larger number, 37, probes are enriched in nonsyntenic contigs (e.g., SOG6066), suggesting that these contigs may account for much of the genome size difference between sorghum and rice. Some repetitive probes are roughly evenly distributed [e.g., SOG6772; RET6 (8)].

Conserved sorghum–rice synteny also corresponded closely with regions of ancient whole-genome duplication (1, 28–31). Overall, 57% of duplicated rice regions (per ref. 1) showed conserved synteny with sorghum contigs, vs. only 16% of regions in which ancient duplication could not be discerned. Tacitly assuming that whole chromosomes were duplicated at once, lower synteny conservation suggests that pericentromeric regions have restructured much more extensively than their surroundings because of the sorghum–rice divergence.

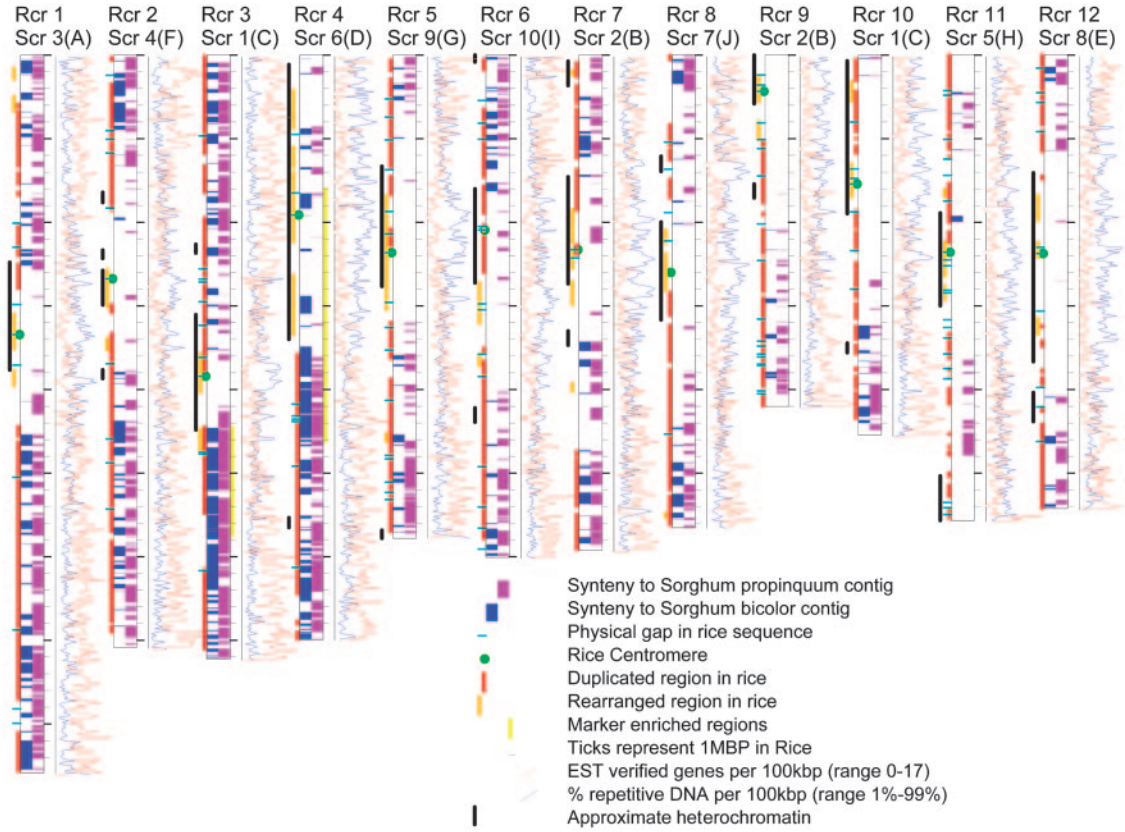

Fig. 2. Coverage of the rice genome by syntenic sorghum BAC contigs. Only hybridization markers that hit two or more BACs in the same contig were considered for microsynteny comparisons. Sorghum contigs were aligned to the rice sequence based on the criteria that two or more low copy probe loci (detecting five contigs or fewer) showed best matches between 5 and 500 kb apart in rice (anchoring 456 SP and 303 SB contigs). Sorghum chromosome (32) and linkage group designations (16) are as cited. Cytologically identified heterochromatic regions were approximated from refs. 19 and 25 based on relative distance from the centromere, assigning approximate base-pair locations without accounting for sequence gaps.

Conservation of sorghum–rice microsynteny is closely associated with recombinogenicity. All rice pericentromeric regions lack conserved synteny and show striking suppression of recombination (Fig. 3). Relative to the genome-wide average, increased synteny is associated with increased recombinational ($r = 0.49$) and physical ($r = 0.53$) length of sorghum chromosomes (32) (Table 3, which is published as supporting information on the PNAS web site). In the extreme, the number of BACs that could be anchored to sorghum chromosome 5 was <60% of that predicted based on its physical size. Similar enrichment was also observed in the sorghum restriction fragment length polymorphism map (16) and probably corresponds to higher gene content in sorghum chromosomes 1–4 than would be predicted by DNA content alone.

The boundary between at least one syntenic/nonsyntenic region appears rather discrete. For a 15.1-megabase (Mb) segment of rice chromosome 4, we enriched the above data with sorghum EST-based overgos corresponding to 272 rice genes spaced at ≈56-kb intervals. We found 72% detectable conserved synteny among the total of 307 loci (60 per Mb) along the distal 5.1 MBP and 2% among 117 loci [11.7 per megabase (Mb)] in the pericentromeric 10 MBP (Fig. 3). Although the marked drop in gene density complicates precise circumscription of the boundary, enrichment of this pericentromeric region to nearly the same marker density as the overall average (15.6/Mb, with 39% conserved synteny) illustrates the dramatic reduction in conserved synteny. Sequence data may be needed to determine whether synteny is merely reduced in such regions or totally eliminated.

Discussion

Gene Rearrangement Appears Generally Deleterious. Preferential conservation of microsynteny in recombinogenic regions suggests that gene rearrangement may generally be deleterious, an intuitive hypothesis that has previously lacked empirical support. Muller's ratchet (33) predicts that deleterious mutations may accumulate in and contribute to degeneration of nonrecombinogenic regions such as mammalian Y chromosomes or incipient plant sex chromosomes (34). This finding is consistent with a strong negative correlation ($r = -0.77$) between repetitive DNA content and recombinational length of rice chromosomes (Table 4, which is published as supporting information on the PNAS web site) and with the much greater abundance of repetitive DNA than EST-verified genes in the pericentromeric regions (Fig. 2). Despite the large differences in overall genome size among the grasses, the size of gene-rich regions remains similar (35). Accelerated gene loss in recombination-poor regions of wheat (36) and a propensity for small insertions in centromeric regions of mammalian genomes (37) support this idea.

Dating the Restructuring of Nonsyntenic Regions. The restructuring of nonsyntenic regions can be approximately dated based on levels of divergence between duplicated rice genes that are concentrated in these regions (Fig. 4). A high concentration of genes duplicated by ancient polyploidy falls near K_s 0.85 (1), whereas rice gene pairs with K_s 0.2–0.6 tend to be in peri/centromeric regions. Thus, shortly after polyploidization, a substantial restructuring of centromeric regions began that lasted until ≈16 MYA [based on the synonymous substitution rate used (38)]. Approximately 18% of the rice genome shows significant concentrations of matches ($P < 1 \times 10^{-5}$) in the K_s range of 0.2–0.6; only 14% of these regions showed conserved synteny with sorghum, whereas 45% of regions lacking concentrations of matches in this K_s range were syntenic (Fig. 4).

GENETICS

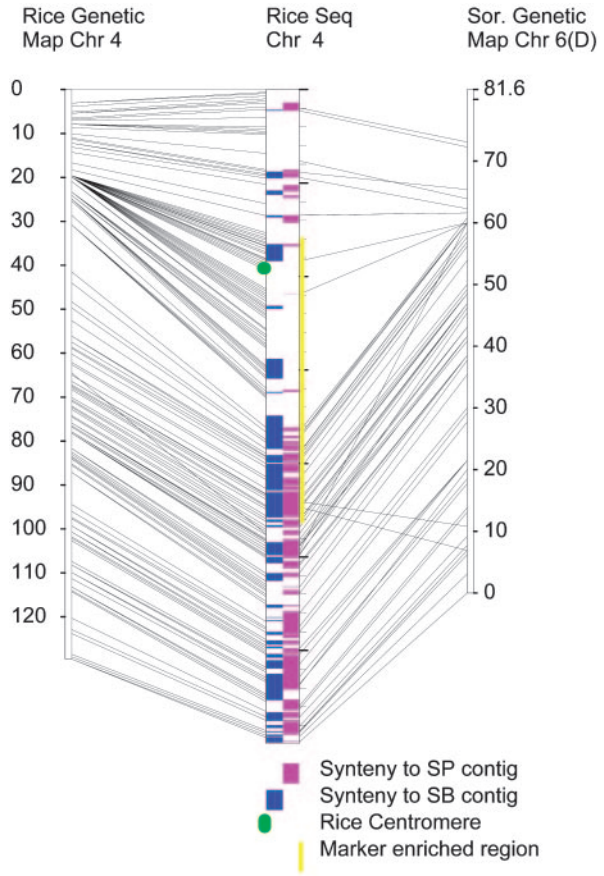

Fig. 3. Comparison of rice genetic map, rice sequence, and sorghum genetic map. BLAST similarity between sorghum probes and the rice sequence required hits with $E \leq 10^{-8}$ based on DNA BLAST. All other chromosomes are available (Figs. 15–26, which are published as supporting information on the PNAS web site). The rice genetic map is based on RGP restriction fragment length polymorphism 2000 (www.gramene.org).

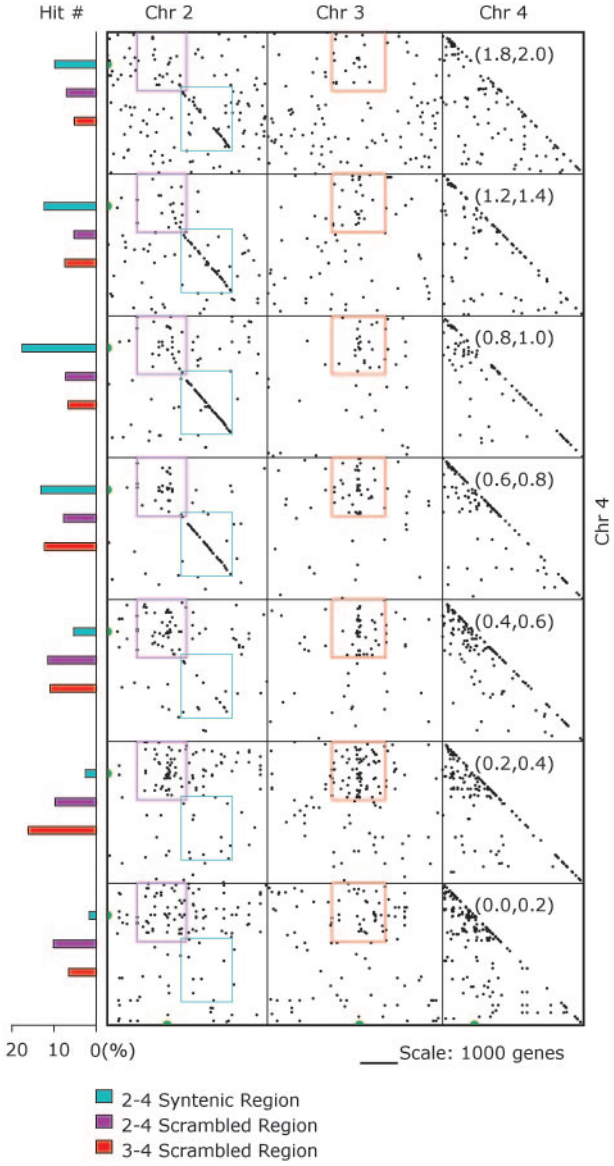

Fig. 4. Restructuring of rice pericentromeric regions over time. The genomic distributions of best-matching copies at differing degrees of synonymous divergence (K_s, in parentheses, upper right) are plotted. The numbers of K_s 0.2–0.6 pairs were counted in 100-gene sliding windows and compared with the random expectation using a χ^2 test. All significant regions are plotted in Fig. 2, and a sampling is shown in Fig. 4 in purple (chr. 2) or red (chr. 3) boxes. Duplicated regions (blue boxes) were described elsewhere (1). The bars to the left of the axis show the number of data points in each of the three indicated regions (two restructured, one duplicated) for each window of K_s values. Additional windows (making the set consecutive) for the whole genome are available (Figs. 5–14), with centromere locations indicated by red dots.

The concentrations of relatively recent single–locus duplications near the centromere may, but do not necessarily, reflect a higher duplication rate in these regions. Alternatively, recent duplications may more often be preserved in these regions, i.e., there may be more rapid loss of single-gene duplications in euchromatic regions. This is consistent with the distribution pattern in the most recent K_s window (0–0.2, Fig. 4), where relatively more gene duplications are in euchromatin.

The restructuring of nonsyntenic regions largely involves migration of DNA between pericentromeric regions of different chromosomes. Concentrations of recently duplicated nontandem genes are present in virtually all pairwise combinations of rice centromeres (Figs. 5–14, which are published as supporting information on the PNAS web site) and also in several noncentromeric but heterochromatic regions such as the upper half of rice chromosome 4. The corresponding region of sorghum also shows massive recombination suppression (Fig. 3). Because restructuring of the rice heterochromatin appears to have continued long after the rice–sorghum divergence (K_s 0.51:1), one would predict that parallel restructuring may have occurred in sorghum.

Restructuring of centromeric regions after ancient polyploidization may perhaps have been under selection, rather than merely being a passive consequence of Muller's ratchet (33). In a newly formed autopolyploid [arguably the most frequent type (39)] with sets of four homologous chromosomes, purging of deleterious alleles would be slow. Transition to diploid inheritance allows more rapid allele frequency changes and reduced genetic load, conferring a potential selective advantage to rapid centromeric divergence. The extensive gene duplication that remains soon after polyploidization may mitigate otherwise negative consequences of mechanisms for rapid structural divergence such as increased transposition. An explosion of repetitive sequence amplification followed polyploidization in maize, perhaps closely (4). Alternatively, centromeric regions may be periodically restructured independently of polyploidization. In the only plant centromere sequenced to date, most repeats are of recent origin [<9.4 million years (40)]. Many transposable element families that may be necessary for centromere maintenance (41) would be expected to degrade synteny over that time.

Leveraging Model Genomes. Identification of a well conserved component of the sorghum and rice genomes will guide contig and sequence assembly in recombinogenic regions of other grasses. Further enrichment of the sorghum physical map with overgos selected using rice as a guide will reduce the population of nonsyntenic BACs/contigs, some of which may just have too few loci to anchor. Targeted overgo enrichment in sorghum segments corresponding to parts of rice chromosomes 3 and 4 (Fig. 2) increased coverage of the corresponding regions of the rice sequence from the average 39% to 92% and 97%, respectively (combining SB and SP). In turn, sorghum contigs bridge gaps in the much larger maize genome (Fig. 1). These benefits promise to be reciprocal; the current sorghum contigs span 35 of 130 physical gaps in a recent rice pseudomolecule (14, 7, and 14 by SP, SB, and both; Table 5, which is published as supporting information on the PNAS web site).

The sorghum–rice comparison, together with hybridization data for other taxa (Table 1), permits us to begin to identify probable euchromatic regions of other genomes. For maize, these data merely add to a detailed physical map (22), but for sugarcane, this comprises a previously undescribed resource. The evolutionary lifespan of euchromatic regions remains unknown; however, hybridization of many of our probes to BACs from banana, separated from the cereals by an estimated 140 million years (1), sets the stage for further study.

Limitations of Microsynteny Approaches. Colinearity along recombinational (genetic) maps of related taxa masks much rearrangement in recombinationally inert regions, which contribute little to such maps (Fig. 3). Model genomes may be of little assistance in the study of nonrecombinogenic regions of other genomes (36). *De novo* characterization of these regions may be needed for each species, requiring new methods to deal with their highly repetitive nature.

Synthesis. Euchromatic and heterochromatic regions, as defined cytologically (19, 24, 25), appear to represent evolutionarily different components of genomes that can be independently detected by several comparative or *in silico* approaches. The structuring of cereal (and presumably other) genomes into these components has many implications for the evolution of individual genes and groups of genes. Differences in chromosomal context may resolve disparate conclusions regarding the degree to which colinearity and synteny persist over time, that have been drawn from comparative analyses of sets of individual BAC/contig sequences. Nonrecombinogenic domains may nurture the evolution of physically dispersed but genetically tightly linked coadapted gene complexes or supergenes that are predicted to be favored by domestication and have been observed in several cases (42). The genomic context of a gene may influence the degree to which information about gene function from model organisms extrapolates to less characterized organisms.

We thank the National Science Foundation Plant Genome Research Program, U.S. Department of Agriculture National Research Initiative Plant Genome Research Program, International Consortium for Sugarcane Biotechnology, and National Grain Sorghum Producers for support. We also thank many scientists who deposit and curate data in public databases.

1. Paterson, A. H., Bowers, J. E. & Chapman, B. A. (2004) *Proc. Natl. Acad. Sci. USA* **101,** 9903–9908.
2. Paterson, A. H., Lin, Y. R., Li, Z. K., Schertz, K. F., Doebley, J. F., Pinson, S. R. M., Liu, S. C., Stansel, J. W. & Irvine, J. E. (1995) *Science* **269,** 1714–1718.
3. Arumuganathan, K. & Earle, E. D. (1991) *Plant Mol. Biol. Rep.* **9,** 208–218.
4. Swigonova, Z., Lai, J. S., Ma, J. X., Ramakrishna, W., Llaca, M., Bennetzen, J. L. & Messing, J. (2004) *Comp. Funct. Genom.* **5,** 281–284.
5. Ming, R., Liu, S. C., Lin, Y. R., da Silva, J., Wilson, W., Braga, D., van Deynze, A., Wenslaff, T. F., Wu, K. K., Moore, P. H., *et al.* (1998) *Genetics* **150,** 1663–1682.
6. Paterson, A. H., Schertz, K. F., Lin, Y. R., Liu, S. C. & Chang, Y. L. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 6127–6131.
7. Cai, W. W., Reneker, J., Chow, C. W., Vaishnav, M. & Bradley, A. (1998) *Genomics* **54,** 387–397.
8. Peterson, D. G., Schulze, S. R., Sciara, E. B., Lee, S. A., Bowers, J. E., Nagel, A., Jiang, N., Tibbitts, D. C., Wessler, S. R. & Paterson, A. H. (2002) *Genome Res.* **12,** 795–807.
9. Marra, M. A., Kucaba, T. A., Dietrich, N. L., Green, E. D., Brownstein, B., Wilson, R. K., McDonald, K. M., Hillier, L. W., McPherson, J. D. & Waterston, R. H. (1997) *Genome Res.* **7,** 1072–1084.
10. Sulston, J., Mallett, F., Durbin, R. & Horsnell, T. (1989) *Comput. Appl. Biosci.* **5,** 101–106.
11. Soderlund, C., Humphray, S., Dunham, A. & French, L. (2000) *Genome Res.* **10,** 1772–1787.
12. Thon, M. R., Martin, S. L., Goff, S., Wing, R. A. & Dean, R. A. (2004) *Fungal Genet. Biol.* **41,** 657–666.
13. Feng, Q., Zhang, Y. J., Hao, P., Wang, S. Y., Fu, G., Huang, Y. C., Li, Y., Zhu, J. J., Liu, Y. L., Hu, X., *et al.* (2002) *Nature* **420,** 316–320.
14. Sasaki, T., Matsumoto, T., Yamamoto, K., Sakata, K., Baba, T., Katayose, Y., Wu, J. Z., Niimura, Y., Cheng, Z. K., Nagamura, Y., *et al.* (2002) *Nature* **420,** 312–316.
15. Yu, Y. S., Rambo, T., Currie, J., Saski, C., Kim, H. R., Collura, K., Thompson, S., Simmons, J., Yang, T. J., Nah, G., *et al.* (2003) *Science* **300,** 1566–1569.
16. Bowers, J. E., Abbey, C., Anderson, S., Chang, C., Draye, X., Hoppe, A. H., Jessup, R., Lemke, C., Lennington, J., Li, Z. K., *et al.* (2003) *Genetics* **165,** 367–386.
17. Lin, Y. R., Zhu, L. H., Ren, S. X., Yang, J. S., Schertz, K. F. & Paterson, A. H. (1999) *Mol. Breed.* **5,** 511–520.
18. Yim, Y. S., Davis, G. L., Duru, N. A., Musket, T. A., Linton, E. W., Messing, J. W., McMullen, M. D., Soderlund, C. A., Polacco, M. L., Gardiner, J. M., *et al.* (2002) *Plant Physiol.* **130,** 1686–1696.
19. Chen, M. S., Presting, G., Barbazuk, W. B., Goicoechea, J. L., Blackmon, B., Fang, F. C., Kim, H., Frisch, D., Yu, Y. S., Sun, S. H., *et al.* (2002) *Plant Cell* **14,** 537–545.
20. Tomkins, J. P., Yu, Y., Miller-Smith, H., Frisch, D. A., Woo, S. S. & Wing, R. A. (1999) *Theor. Appl. Genet.* **99,** 419–424.
21. Vilarinhos, A. D., Piffanelli, P., Lagoda, P., Thibivilliers, S., Sabau, X., Carreel, F. & D'Hont, A. (2003) *Theor. Appl. Genet.* **106,** 1102–1106.
22. Cone, K. C., McMullen, M. D., Bi, I. V., Davis, G. L., Yim, Y. S., Gardiner, J. M., Polacco, M. L., Sanchez-Villeda, H., Fang, Z. W., Schroeder, S. G., *et al.* (2002) *Plant Physiol.* **130,** 1598–1605.
23. Feltus, F. A., Wan, J., Schulze, S. R., Estill, J. C., Jiang, N. & Paterson, A. H. (2004) *Genome Res.* **14,** 1812–1819.
24. Cheng, Z. K., Buell, C. R., Wing, R. A., Gu, M. H. & Jiang, J. M. (2001) *Genome Res.* **11,** 2133–2141.
25. Jiao, Y., Jia, P., Wang, X., Su, N., Yu, S., Zhang, D., Ma, L., Feng, Q., Jin, Z., Li, L., *et al.* (2005) *Plant Cell* **17,** 1641–1657.
26. Miller, J. T., Dong, F. G., Jackson, S. A., Song, J. & Jiang, J. M. (1998) *Genetics* **150,** 1615–1623.
27. Zwick, M. S., Islam-Faridi, M. N., Zhang, H. B., Hodnett, G. L., Gomez, M. I., Kim, J. S., Price, H. J. & Stelly, D. M. (2000) *Am. J. Bot.* **87,** 1757–1764.
28. Goff, S. A., Ricke, D., Lan, T. H., Presting, G., Wang, R. L., Dunn, M., Glazebrook, J., Sessions, A., Oeller, P., Varma, H., *et al.* (2002) *Science* **296,** 92–100.
29. Kishimoto, N., Higo, H., Abe, K., Arai, S., Saito, A. & Higo, K. (1994) *Theor. Appl. Genet.* **88,** 722–726.
30. Wang, X. Y., Shi, X. L., Hao, B. L., Ge, S. & Luo, J. C. (2005) *New Phytol.* **165,** 937–946.
31. Nagamura, Y., Inoue, T., Antonio, B. A., Shimano, T., Kajiya, H., Shomura, A., Lin, S. Y., Kuboki, Y., Harushima, Y., Kurata, N., *et al.* (1995) *Breed. Sci.* **45,** 373–376.
32. Kim, J.-s., Klein, P. E., Klein, R. R., Price, H. J., Mullet, J. E. & Stelly, D. M. (2005) *Genetics* **169,** 1169–1173.
33. Muller, H. J. (1932) *Am. Nat.* **66,** 118–138.
34. Liu, Z. Y., Moore, P. H., Ma, H., Ackerman, C. M., Ragiba, M., Yu, Q. Y., Pearl, H. M., Kim, M. S., Charlton, J. W., Stiles, J. I., *et al.* (2004) *Nature* **427,** 348–352.
35. Feuillet, C. & Keller, B. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 8265–8270.
36. Akhunov, E. D., Goodyear, A. W., Geng, S., Qi, L. L., Echalier, B., Gill, B. S., Miftahudin, Gustafson, J. P., Lazo, G., Chao, S. M., *et al.* (2003) *Genome Res.* **13,** 753–763.
37. Bailey, J. A., Yavor, A. M., Massa, H. F., Trask, B. J. & Eichler, E. E. (2001) *Genome Res.* **11,** 1005–1017.
38. Lynch, M. & Conery, J. S. (2000) *Science* **290,** 1151–1155.
39. Harlan, J. R. & Dewet, J. M. J. (1975) *Bot. Rev.* **41,** 361–390.
40. Nagaki, K., Cheng, Z. K., Ouyang, S., Talbert, P. B., Kim, M., Jones, K. M., Henikoff, S., Buell, C. R. & Jiang, J. M. (2004) *Nat. Genet.* **36,** 138–145.
41. Wong, L. H. & Choo, K. H. A. (2004) *Trends Genet.* **20,** 611–616.
42. D'Ennequin, M. L., Toupance, B., Robert, T., Godelle, B. & Gouyon, P. H. (1999) *J. Evol. Biol.* **12,** 1138–1147.