

## RESEARCH ARTICLES

# Seventy Million Years of Concerted Evolution of a Homeologous Chromosome Pair, in Parallel, in Major Poaceae Lineages <sup>W</sup>

Xiyin Wang<sup>a,b</sup>, Haibao Tang<sup>a,c</sup>, and Andrew H. Paterson<sup>a,c,1</sup>

<sup>a</sup> Plant Genome Mapping Laboratory, University of Georgia, Athens, Georgia 30602

<sup>b</sup> Center for Genomics and Computational Biology, School of Life Sciences, and School of Sciences, Hebei United University, Tangshan, Hebei 063000, China

<sup>c</sup> Department of Plant Biology, University of Georgia, Athens, Georgia 30602

**Whole genome duplication ~70 million years ago provided raw material for Poaceae (grass) diversification. Comparison of rice (*Oryza sativa*), sorghum (*Sorghum bicolor*), maize (*Zea mays*), and *Brachypodium distachyon* genomes revealed that one paleo-duplicated chromosome pair has experienced very different evolution than all the others. For tens of millions of years, the two chromosomes have experienced illegitimate recombination that has been temporally restricted in a stepwise manner, producing structural stratification in the chromosomes. These strata formed independently in different grass lineages, with their similarities (low sequence divergence between paleo-duplicated genes) preserved in parallel for millions of years since the divergence of these lineages. The pericentromeric region of this homeologous chromosome pair accounts for two-thirds of the gene content differences between the modern chromosomes. Both intriguing and perplexing is a distal chromosomal region with the greatest DNA similarity between surviving duplicated genes but also with the highest concentration of lineage-specific gene pairs found anywhere in these genomes and with a significantly elevated gene evolutionary rate. Intragenomic similarity near this chromosomal terminus may be important in hom(e)ologous chromosome pairing. Chromosome structural stratification, together with enrichment of autoimmune response-related (nucleotide binding site–leucine-rich repeat) genes and accelerated DNA rearrangement and gene loss, confer a striking resemblance of this grass chromosome pair to the sex chromosomes of other taxa.**

## INTRODUCTION

Whole genome duplications (WGDs) have occurred in most if not all angiosperm genomes and have provided raw material for natural selection to act upon for tens of millions of years (Tang et al., 2008a). Evolution of the grass family, Poaceae, which provides much of our food, feed, fodder, and fuel, was profoundly influenced by a WGD that occurred ~70 million years ago (mya; Paterson et al., 2003, 2004; Wang et al., 2005; Yu et al., 2005). However, most chromosomes duplicated ~70 mya were already highly differentiated from one another by ~50 mya when the rice (*Oryza sativa*) and sorghum (*Sorghum bicolor*) lineages diverged (Paterson et al., 2009), with both members of the paleo-duplicated gene pairs remaining extant in only ~17% of cases.

Whereas the two members of most paleo-duplicated chromosome pairs have evolved largely independently of one another, one pair of duplicated chromosomes is a striking exception, as compared with all other chromosomes (Wang et al., 2009). Rice chromosomes 11 and 12 (hereafter R11 and R12) share an ~3-

Mb duplicated DNA segment at the termini of their short arms, dated based on synonymous substitutions to ~5 to 7 mya and once suspected to represent a segmental duplication more recent than the pan-grass WGD (International Rice Genome Sequencing Project, 2005; Wang et al., 2005; Yu et al., 2005). Remarkably, the corresponding regions of the sorghum genome (S5 and S8, respectively) also contain such an apparently recent duplication (Paterson et al., 2009). Physical and genetic maps suggest shared terminal segments of the corresponding chromosomes in wheat (*Triticum aestivum*; 4 and 5), foxtail millet (*Setaria italica*; VII and VIII), and pearl millet (*Pennisetum americanum*; linkage groups 1 and 4; Devos et al., 2000; Singh et al., 2007). It would be exceedingly unlikely for segmental duplications to each happen independently at such closely corresponding locations in several reproductively isolated lineages. A much more parsimonious hypothesis is that the R11/12 and S5/8 regions each resulted from the pan-grass duplication 70 mya but have an unusual evolutionary history (Paterson et al., 2009).

Here, we performed a detailed analysis of these chromosomes and their homeologs from *Brachypodium distachyon* (Brachypodium Genome Sequencing Initiative, 2010) and maize (*Zea mays*; Schnable et al., 2009). Homeologous and illegitimate recombination has continued for millions of years after the divergence of these homeologs and remains ongoing in rice and perhaps other grasses. Homeologous recombination has resulted in prominent concentration of lineage-specific gene losses and significantly

<sup>1</sup> Address correspondence to paterson@uga.edu.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is: Andrew H. Paterson (paterson@uga.edu).

<sup>W</sup>Online version contains Web-only data.  
www.plantcell.org/cgi/doi/10.1105/tpc.110.080622

elevated gene divergence in the affected regions, which may potentially contribute to speciation (Scannell et al., 2006). Gradual and step-by-step restrictions on recombination, starting from the pericentromeric regions around the time of the WGD, have resulted in chromosome structural stratification. A significant enrichment of disease resistance genes near the pericentromeric regions is another prominent hallmark of these chromosomes, possibly contributing to their singular evolution. The structural stratification of these chromosomes bears a striking resemblance to that of sex chromosomes in other diverse organisms (Lahn and Page, 1999; Charlesworth, 2002) and provides a unique opportunity to study homologous chromosomal pairing when there is competition from homeologous chromosomes.

## RESULTS

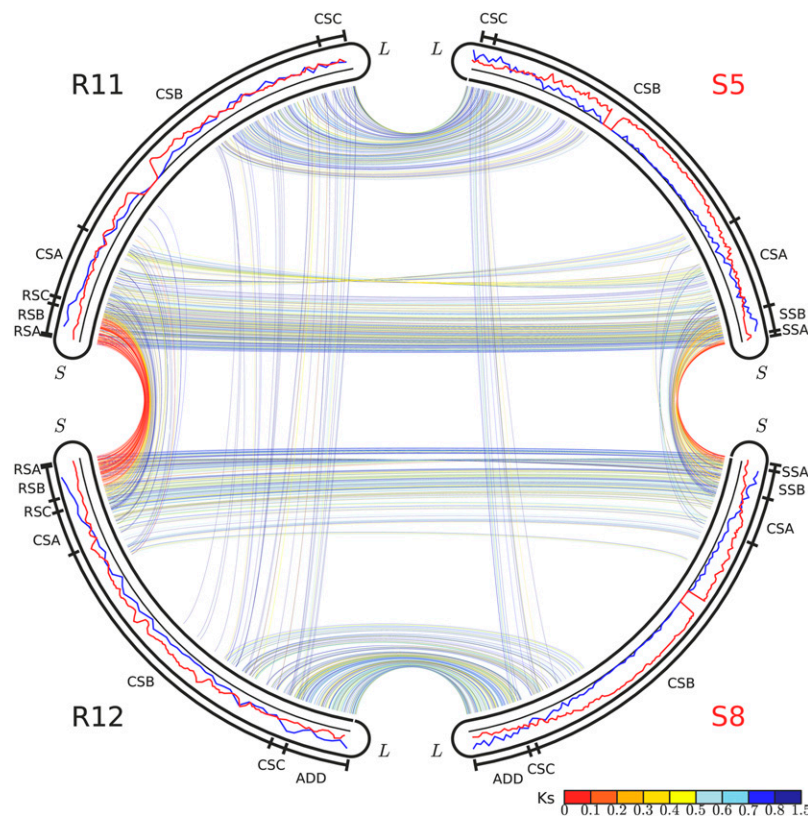
### Stepwise Restriction of Homeologous Recombination

Using multiple chromosome sequence alignment (Tang et al., 2008b), we revealed homology between R11 and R12 and their respective orthologs S5 and S8 (Figure 1), along virtually their

entire lengths except for a terminal segment translocated to the common ancestor of R12/S8 from R3/S1 (see Supplemental Data Sets 1A–1C and Supplemental Figure 1 online). Syntenic gene density is highest on the distal short arms and lowest in pericentromeric chromosomal regions of each species. Structural comparison revealed an ~5-Mb DNA inversion on S5 (12–17 Mb) and a 1.86-Mb inversion on S8 (0.94–2.8 Mb).

Concerted evolution has occurred between paleo-duplicated genes in both rice and sorghum, with R11/12 and S5/8 being by far the most affected chromosomes (Wang et al., 2009). Among 149 gene “quartets” (duplicates in both rice and sorghum; see Supplemental Data Set 1D online) that could be aligned on these chromosomes, 55.2% (R11/12) and 49.7% (S5/8) show evidence of gene conversion since the rice–sorghum divergence (Table 1). A remarkable 90% and 83% of paleo-duplicated genes on the distal short arms of R11/12 and S5/8, respectively, show evidence of conversion, declining to ~6% in the interstitial regions (similar to the whole genome average of ~8%) and virtually to zero near pericentromeric regions.

Intragenomic divergence of paleo-duplicated genes along both R11/12 and S5/8 is closely related to their physical locations, being lowest on the distal short arms and increasing toward the



**Figure 1.** Homology Pattern and Evolutionary Model of Chromosomes R11 and R12 from Rice, S5 and S8 from Sorghum, and Their Common Ancestor.

Strata (chromosomal segments RSA-RSC, SSA-SSB, and CSA-CSC) are displayed in subfigures. RS indicates strata formed after rice–sorghum divergence along R11 and R12; SS indicates those along S5 and S8; CS indicates common strata on rice and sorghum chromosomes that might have formed before the divergence of two species. Additional segments on R12 and S8 are noted with ADD. Blue and red curves along each chromosome show gene and repetitive sequences densities, respectively. S and L indicate short and long arms. Lines between chromosomes connect syntenic genes, and colors correspond to Ks values.

**Table 1.** Gene Conversion Rates and Ages of Homologous Strata Inferred on Chromosomes R11, R12, S5, and S8

Strata	Chromosome	Physical Location (Mb)	Gene No.	Gene Density <sup>a</sup>	Colinear Genes <sup>b</sup>	Quartet No. <sup>c</sup>	Conversion Rate <sup>d</sup>	Ks Median	Date (mya)	P Value <sup>e</sup>
Rice-specific										
RSA	R11	0.00–0.14	17	123.38	11 <sup>f</sup>	4	1	0.005	<0.5	RSA–RSB 3.89E-04
	R12	0.00–0.13	13	96.42						
RSB	R11	0.14–2.19	245	119.72	131	67	0.91	0.115	9.4	RSB–RSC 7.37E-12
	R12	0.14–2.22	256	122.89						
RSC	R11	2.21–2.72	59	114.35	29	17	0.47	0.477	39.1	RSC–CSA 1.75E-07
	R12	2.23–2.98	70	93.38						
Sorghum-specific										
SSA	S5	0.03–0.67	79	123.81	39	27	0.93	0.163	13.4	SSA–SSB 5.83E-13
	S8	0.01–0.84	87	104.37						
SSB	S5	0.73–4.23	359	102.45	67	61	0.56	0.486	59.7	SSB–CSA 6.86E-11
	S8	0.86–4.29	351	102.4						
Common to rice and sorghum										
CSA	R11	2.74–7.76	354	70.5	38	34	0.08	0.794	65.1	CSA–CSB 5.00E-02
	R12	2.98–5.71	198	72.65						
	S5	4.24–8.90 12.06–16.93 <sup>g</sup>	503	52.78	41					
CSB	S8	4.30–10.14	327	55.92		13	0	0.985	80.7	CSB–CSC 2.03E-01
	R11	7.79–28.86	971	46.08	37					
	R12	5.76–22.84	712	41.67						
CSC	S5	11.62–12.06 18.66–59.55	1149	27.8	11	10	0	0.653	53.5	
	S8	10.73–47.52	636	17.29						
	R11	29.03–30.66	117	71.89	8					
	R12	22.87–23.80	93	100.81		9	0			
	S5	60.21–61.37	119	102.75						
	S8	47.85–48.58	53	72.31						

<sup>a</sup>Gene number per Mb.<sup>b</sup>Numbers of those colinear genes estimated with their Ks values.<sup>c</sup>Quartets with good alignment.<sup>d</sup>Bootstrap percentage > 0.8.<sup>e</sup>Wilcoxon rank-sum test.<sup>f</sup>Three additional genes annotated by the authors.<sup>g</sup>An inversion in S5 produced two segments of CSA therein.

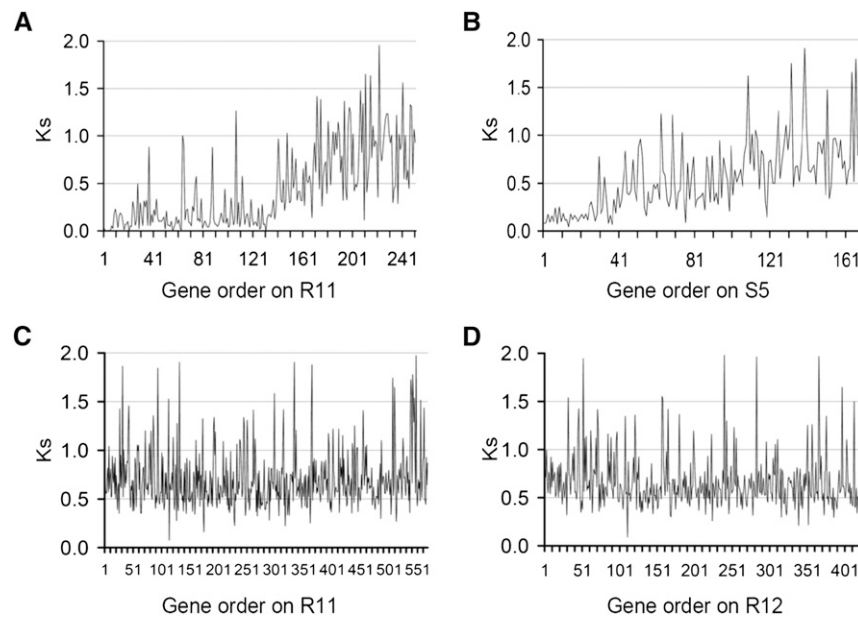
pericentromeric region (Figures 2A and 2B). This is in contrast to more uniform intergenomic (rice–sorghum) divergence patterns between orthologs (Figures 2C and 2D). Based on 254 R11/12 and 167 S5/8 homeologs, median Ks values (frequencies of synonymous substitutions at synonymous nucleotides, after removing the too-divergent homeologs without Ks estimated) near the short arms' termini are 0.005 in rice and 0.163 in sorghum, gradually increasing to ~1 at the other termini (see Supplemental Data Sets 1B and 1C online). A sharp decrease in gene conversion rate, and an increase in Ks, occur near pericentromeres in both genomes, delineated by homologous genes Os11g0157200, Os12g0159600, Sb05g003630, and Sb08g003720.

For further study, we divided the homeologous pairs into highly converted (HC) and little converted (LC) subgroups in both species (Table 1). On R11/12, 171 HC and 83 LC homeologs are highly divergent in Ks (0.22 versus 0.86;  $P < 2.2e-16$ ), with HC homeologs appearing much younger and LC homeologs much older ( $P < 2.2e-16$  for both) than rice–sorghum divergence (50 mya; expected Ks = 0.61). On S5/8, 106 HC and 61 LC homeologs, respectively, also appear much younger and much older than expected (Ks = 0.39 versus 0.83;  $P = 2.94e-14$ ).

One-dimensional spatial autocorrelation analysis of Ks values differentiating paleo-duplicated genes indicates that R11/12 HC homeologs can be clustered into three strata, RSA, RSB, and RSC (Table 1, Figure 1), extending 140/140 kb, 2.5/2.5 Mb, and 540/750 kb on R11/12 and containing 11, 131, and 29 homeologs with median Ks < 0.01, 0.12, and 0.48, respectively. Sorghum HC homeologs can be clustered into two strata, SSA and SSB, extending 638/830 kb and 3.5/3.5 Mb on S5/8 and containing 39 and 67 homeologs with median Ks of 0.16 and 0.49, respectively. Neither the RSA/B nor RSB/C transition precisely coincides with the SSA/B transition, indicating that these strata formed independently in the two lineages and that their similarities (low Ks of paleo-duplicated genes) have been preserved in parallel.

### Chromosomal Stratification and Molecular Dating

LC gene pairs appear to have formed before the rice–sorghum divergence. While implementing our algorithm to find strata, we first removed homeologs (four on R11/12 and two on S5/8) with Ks < 0.4, perhaps affected by infrequent conversion. After strata



**Figure 2.** Distribution of  $K_s$  between Syntenic Gene Pairs along Chromosomes.

$K_s$  values are arranged according to gene order on one of the two involved chromosomes. Pairs of  $K_s > 2$ , mostly from young strata, are removed from display.

**(A)** R11–R12 paralogs.

**(B)** S5–S8 paralogs.

**(C)** R11–S5 orthologs.

**(D)** R12–S8 orthologs.

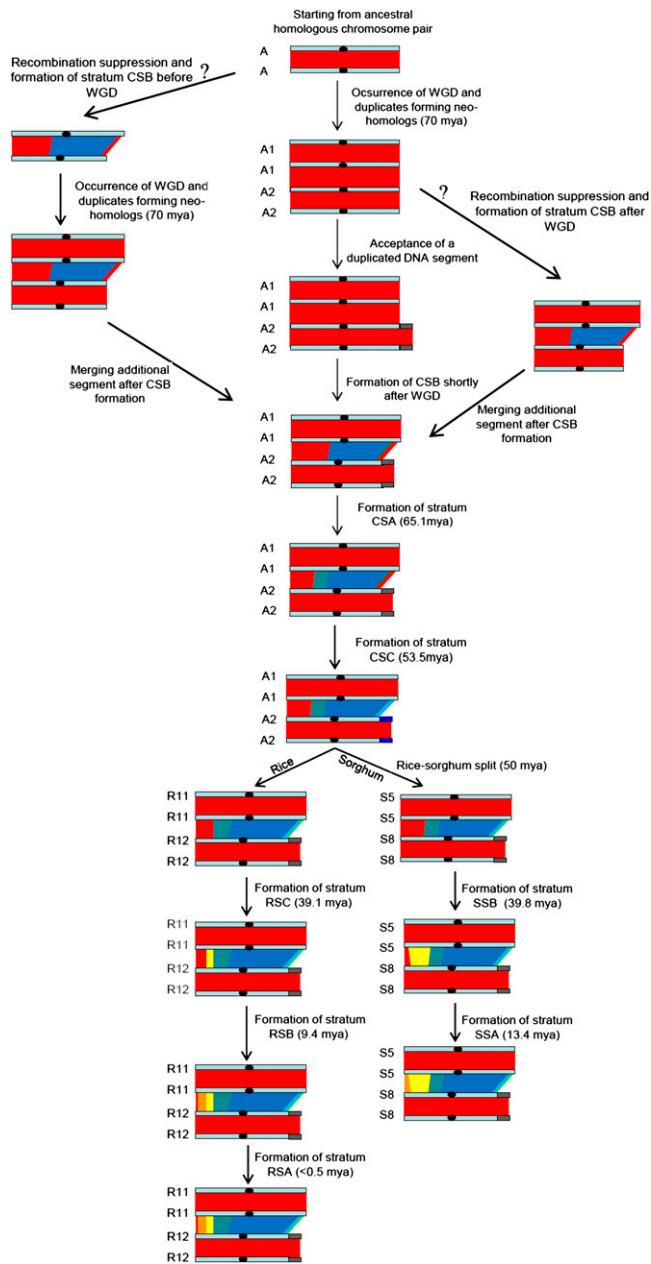
were initially inferred, we added back the six homeologs. Eventually, we clustered 144 LC homeologs into three strata common in both species, CSA, CSB, and CSC (Table 1). Each stratum comprises a quartet of two homeologous regions in each of the two genomes. CSA covers  $\sim 5/3$  Mb containing 38 homeologs on R11/12 and  $\sim 4/5.8$  Mb containing 41 homeologs on S5/8, with median  $K_s$  of 0.79 between homeologs, significantly different from adjacent RSC and SSB. CSB crosses the gene-poor pericentromeric regions in both species, covering  $\sim 20/17$  Mb containing 37 homeologs on R11/12 and 42/37 Mb containing 11 homeologs on S5/8, with median  $K_s$  (0.98) that is significantly larger than in CSA. CSC covers 1.6/1 Mb containing 8 homeologs on R11/12 and 1.2/0.7 Mb containing 9 homeologs on S5/8, with median  $K_s$  (0.65) smaller (but not significantly so, due to small gene numbers) than CSB.

Because concerted evolution makes some strata appear artificially younger (Wang et al., 2009), the time of formation of the underlying homeologous chromosome pair is best indicated by the oldest stratum, CSB, in which median  $K_s$  of 0.98 implies an age of 80.7 mya (95% confidence interval, 72.3–91.6 mya), seemingly older ( $P < 2.2 \times 10^{-16}$ ) than the WGD (95% confidence interval, 67.7–70 mya). Despite this statistical evidence, we are cautious about asserting that CSB duplicated before the WGD (and reflect this uncertainty in our evolutionary model; Figure 3), since a high frequency of pseudogenes or other rate-distorting factors may inflate  $K_s$  values and the exclusion of possibly converted genes in CSB may have induced statistical bias.

### DNA Loss and Rearrangement Leading to Recombination Suppression

To try to explain the perplexing evolution of these chromosomes, we propose that R11/12 and S5/8 originated from a pair of homeologous chromosomes produced by the WGD from a common ancestor that we will refer to as “chromosome A” (Figure 3). In the modern rice and sorghum genomes, CSB regions of R12/S8 have much less DNA and fewer genes than their homeologs R11/S5, accounting for two-thirds of gene content differences between R11/12 and S5/8 (Table 1). Because numerous gene gains are unlikely, this suggests fractionation (and degeneration) of the ancestral chromosome of R12/S8. Since R12 and S8 closely resemble one another and each differs from its homeolog R11 and S5, this fractionation presumably preceded the rice–sorghum divergence. Indeed, the high median  $K_s$  of CSB homeologs (above) suggests that this fractionation process may have occurred near the WGD.

Extensive gene/DNA loss on one member of heteromorphic chromosome pairs is closely associated with suppressed recombination, by Muller’s ratchet (Charlesworth, 2002). There is considerable gene number difference between homeologous regions of CSB but not the younger CSA or CSC strata, whereas the distal short arms of all four chromosomes remained very similar. In that R11 and R12 have experienced concerted evolution more recently than the divergence of rice subspecies *japonica-indica* only  $\sim 0.4$  mya (Wang et al., 2007), occasional illegitimate recombination between homeologs (formerly homologs) seems likely to have



**Figure 3.** Evolutionary Model of Chromosomes R11 and R12 and Their Respective Orthologs S5 and S8.

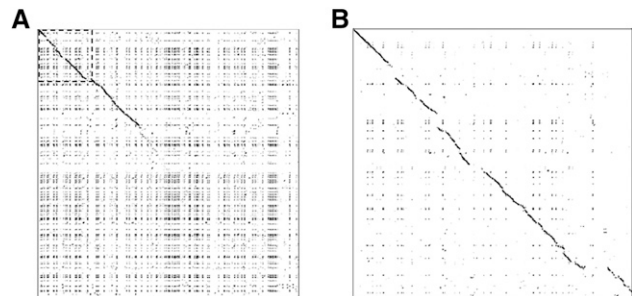
Chromosomes are shown with rectangles, between which a range of colors show the formations of the inferred strata and the recombination levels at present, with red showing active or normal recombination, blue showing the least recombination, and intermediate colors showing intermediate degrees of recombination suppression. The divergence of homologous chromosomes may have started before or after the occurrence of the WGD, which cannot be settled by sequence divergence analysis. We have shown this alternative option in the figure, and a question mark is used to highlight this question. The acceptance of a duplicated block may have occurred before or after the formation of stratum CSB. This unsettled controversy is also highlighted with a question mark.

continued in the newly polyploid grass ancestor as well as in each lineage after their divergence. Such illegitimate recombination may perhaps have contributed to the formation of CSA and CSC before the rice–sorghum divergence and of RSB-C and SSA-B after the rice–sorghum divergence (Figure 3).

Additional factors may have reduced recombination. An ~5-Mb DNA inversion on S5 (12–17 Mb physically) and a 1.86-Mb inversion on S8, both after the WGD, may have suppressed homeologous recombination. The S8 inversion in particular is in SSB, only ~300 kb from the boundary with SSA, and may have suppressed homeologous recombination in much of SSB, which would have resulted in larger Ks values than those of SSA, as indicated (Figures 1 and 2B). Moreover, the borders of the R11/12 strata are enriched with repeats (see Supplemental Data Set 1E online), associated with genome instability such as DNA rearrangement (Caceres et al., 1999; Fu et al., 2002) that may lead to suppressed recombination. For example, the interval inferred to contain the RSA-B break point on R12 extends 80 kb, and 68% of the sequences are long terminal repeat retrotransposon-like elements, far higher than the average content (~12%) in the surrounding regions (Figure 4). Further DNA rearrangements have occurred in *Brachypodium* and even more in maize, which contribute to the relatively low gene conversion rates in these species (see below).

**Intragenomic Similarity Coupled with Intergenomic Diversity**

Concerted evolution acting on the terminal strata of R11/12 and S5/8 appears to have reduced intragenomic diversity while increasing intergenomic divergence. As noted above, RSA and SSA have the highest rates of gene conversion and greatest similarity between homeologous genes found anywhere in the respective genomes. Paradoxically, they also have by far the highest concentrations of species-specific gene pairs. There are 33 and 23 rice- and sorghum-specific gene pairs on these chromosomes (see Supplemental Data Set 1F online), with respectively 100% and 90% in the young strata (Figure 5). Both



**Figure 4.** Dot Map between Chromosomes R11 and R12.

Highly similar strings, reflected by dots in the figure, from two chromosomes were used to make the dot map. The suffix tree algorithm was used to construct dot maps with string size set to be 16 bp or greater. Dot maps show the initial ~5-Mb (A) and ~1-Mb (B) nucleotide sequences on termini of short arms.

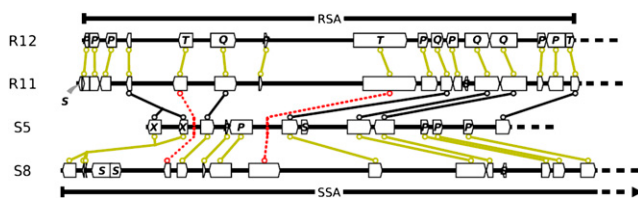
members of a remarkable 50% of the 16 duplicated RSA gene pairs are absent from sorghum and 15 (38%) of 39 SSA pairs are absent from rice. RSB also has a high frequency of taxon-specific duplicated genes (Table 1).

Gene losses on either one of a pair of homeologs experiencing concerted evolution may be commuted to the other, perhaps explaining the more than 10-fold higher gene loss in the RSA and SSA regions than the genome-wide averages of 1.8% in rice and 3.1% in sorghum since their divergence  $\sim 50$  mya (Paterson et al., 2009). The alternative hypothesis, that two corresponding genes would each be recently arrived (in the taxon retaining the gene pair) at homeologous locations tracing to duplication in the grass ancestor, is implausible. Moreover, gene losses are widespread in duplicated genomic regions in many organisms after WGD (Tang et al., 2008a).

Gene conversion is also associated with increased divergence of orthologs between the respective species. HC orthologs had larger  $K_s$  and  $K_a$  (nonsynonymous nucleotide substitution) values than did LC orthologs (0.66 versus 0.57 [ $P = 0.01$ ], 0.17 versus 0.12 [ $P = 0.006$ ]), although  $K_a:K_s$  ratios suggest no difference in the type of selective pressure experienced by the respective gene sets.

#### Further Evidence from *Brachypodium* and Maize

With the availability of maize and *Brachypodium* genome sequences, we retrieved colinear genes on corresponding homeologous chromosomes (Figure 6; see Supplemental Data Set 1A and Supplemental Figure 1 online) and investigated the possibility that homeologous gene conversion has occurred in these additional genomes. There has been a maize-specific WGD, after which large-scale genome reshuffling occurred (Schnable et al., 2009). The *Brachypodium* genome has also experienced widespread rearrangement (Brachypodium Genome Sequencing Initiative, 2010). In both maize and *Brachypodium*, these extensive rearrangements have led to reductions in chromosome number. Comparatively, rice and sorghum genomes have preserved much of the common ancestral genome structure of grasses (Salse et al., 2009). In spite of genome reshuffling in both species, it was possible to reveal the regions that are respectively orthologous to R11/S5 and R12/S8 (Figure 6). Among 58 rice–



**Figure 5.** Rice RSA Stratum and Associated Portion of Sorghum SSA.

This figure illustrates four rice–sorghum quartets (Q), three triplets (T; losses on S5 [2] and S8), 13 pairs of taxon-specific genes (P; five sorghum and eight rice), eight taxon-specific singletons (S; six sorghum and two rice), and one lineage-specific duplication (X; two terminal genes on S5 and S8 share a common ancestor with single rice genes). Classification letters for a gene family appear in only one member (members connected by lines).

*Brachypodium* homologous quartets (see Supplemental Data Set 1G online), we found that  $\sim 30\%$  of *Brachypodium* homeologs have been converted by one another since rice–*Brachypodium* divergence. The lower conversion rate in *Brachypodium* than rice or sorghum (49.7–55.2%) may be due to more DNA rearrangement in the *Brachypodium* genome, breaking homeologous segments into small pieces nested within other large DNA patches and perhaps reducing their chance to recombine (Figure 6). Nonetheless, the conversion pattern corresponds to those in rice and sorghum with regard to chromosomal distribution, albeit showing a less distinctive gradient change in DNA similarity. A comparison between sorghum and maize homologous quartets showed that 39% (24/62) and 23% (14/62) of duplicated genes have been converted after the divergence of these species, respectively. However, extensive DNA rearrangements and losses in the maize genome, very likely occurring after its genome-specific WGD, make it difficult to align the homeologous chromosomal segments and to perform the comprehensive analysis possible in other species.

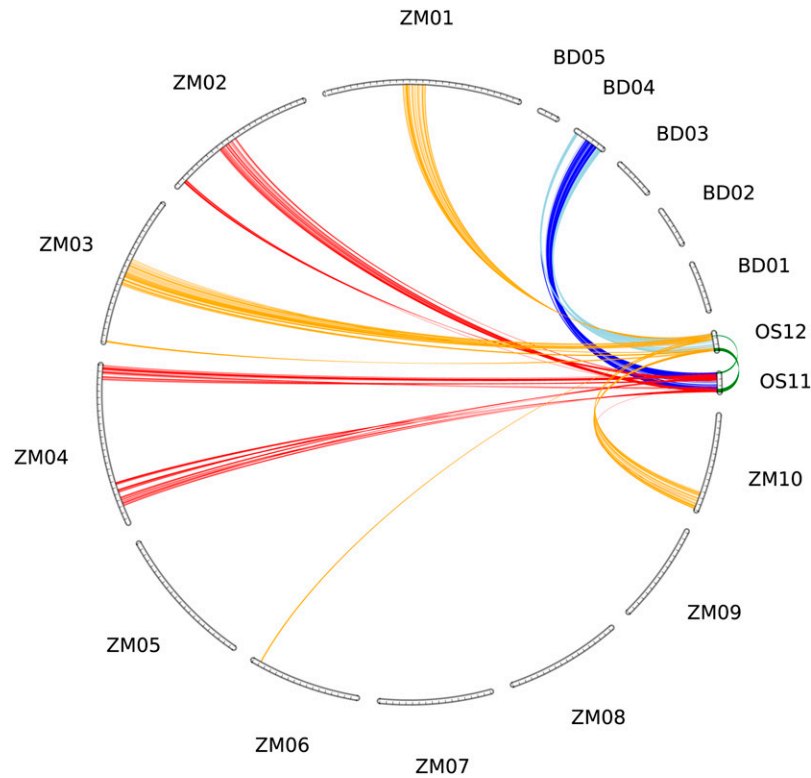
## DISCUSSION

### A Singular Evolutionary History

With the availability of a dozen plant genome sequences, it has become clear that widespread and even recursive WGDs have been of central importance in angiosperm evolution (Bowers et al., 2003; Tang et al., 2008a). Immediately following a WGD, a genome may experience a period of instability characterized by extensive chromosomal rearrangements and numerous gene losses (Paterson et al., 2004; Wang et al., 2005; Thomas et al., 2006; Sankoff et al., 2010), with these changes eventually contributing to the formation of a new diploid karyotype (Bowers et al., 2005; Lysak et al., 2006; Salse et al., 2009). The fractionation process after WGD may be explained by a short-DNA deletion mechanism such as that inferred using the maize genome sequence (Woodhouse et al., 2010).

The derivatives of grass chromosome A, including R11/12 and S5/8, had a singular evolutionary history (Figure 3). The ancestor of R12 and S8 lost much of its pericentromeric gene content, perhaps accompanying recombination suppression near the time of the WGD. This implies a history of unbalanced gene losses between the duplicated chromosomes, a phenomenon that is common (Wang et al., 2005). Such imbalance may imply a fitness advantage to preservation of some genes at linked loci on at least one copy of the duplicated chromosomes or perhaps epigenetic factors (Freeling and Thomas, 2006). Divergence in specific regions might have begun to reduce recombination between the homeologs (Bowers et al., 2005). On the derivatives of chromosome A, recombination suppression may have started from the pericentromeric regions and progressively expanded to other regions in a series of evolutionary episodes.

Factors including repetitive sequence accumulation, chromosomal rearrangement, and gene loss may have contributed to the singular evolutionary history of chromosome A and its derivatives. The grass WGD presumably produced identical (autopolyploid) or similar (allopolyploid) homeologous chromosomes. If the similarity



**Figure 6.** Homology between Chromosomes R11 (OS11) and R12 (OS12) and Maize (ZM) and *Brachypodium* (BD) Genomes.

The green lines exterior to the circle of chromosomes show paralogy relationships between two rice chromosomes; orthologous maize regions of Os11 and Os12, respectively, are shown in red and orange, and those in *Brachypodium* are shown in blue and light blue.

between homeologs was sufficient that recombination could occur frequently, then the resulting genome could have been unstable, forming complex structures here and there. A segregation process in which DNA breakages and rearrangements are frequent could contribute to high rates of gene loss such as those reported in extant genomes of ancient polyploids (Simillion et al., 2002; Bowers et al., 2003; Paterson et al., 2004; Wang et al., 2005) and in artificial polyploids soon after their formation (Feldman et al., 1997; Gaeta et al., 2007). Our analysis indicated that each episode of recombination suppression may have been preceded by DNA rearrangement, especially DNA inversion. DNA rearrangements and losses may contribute to the divergence of homeologs from one another in structure and DNA content, consequently “stabilizing” a newly formed polyploid by restricting further homeologous pairing and recombination. Accumulation of repetitive sequences near the break points that divide the strata could perhaps be directly involved in the recombination suppression mechanism or could be a passive consequence of suppressed recombination (i.e., Muller’s ratchet).

The mechanics of chromosome pairing may have contributed to the patterns of intragenomic variation along the derivatives of grass chromosome A. Homologous chromosome pairing in early meiotic prophase is accompanied by dynamic repositioning of chromosomes in the nucleus and formation of a cytological structure called the telomere bouquet (i.e., chromosomes that are bundled at the telomere to form a bouquet-like arrangement;

Ding et al., 2004; Bozza and Pawlowski, 2008). It has been proposed that the telomere-clustering bouquet, nuclear oscillation, and recombination are three key steps to facilitate homologous chromosome pairing (Ding et al., 2004). This implies that if duplicated chromosomes preserve the telomeres and the proximal chromosomal regions, they may preserve the ability to pair and recombine with one another like homologous chromosomes. However, such illegitimate pairing near the termini may increase genomic instability due to multivalent formation and segregation disorder. Therefore, it is not exceptional to find that at least one member of duplicated chromosome pairs have terminal breakages/inversions (see chromosomal dot plots in Paterson et al., 2009; *Brachypodium* Genome Sequencing Initiative, 2010), which may reduce homeologous chromosome pairing and contribute to genomic stability. The genome structure near one terminus of chromosome A appears to be quite old. We recently reported that the grass genome appears to have been affected by additional WGD(s) before the pan-grass event, perhaps as much as 130 mya (Tang et al., 2010). The derivatives of grass chromosome A, including R11/12 and S5/8, are unusual in having preserved relatively more of the ancient pre-WGD chromosomal structure than other taxa, particularly near the short arm termini. Rice is particularly striking, with median  $K_s < 0.01$ , among duplicated genes in the terminal strata of R11/12. The chromosomal termini of other grasses show greater

divergence, but it is not always accompanied by structural rearrangement, suggesting that other factors also contribute to the singular evolution of the derivatives of grass chromosome A.

### Homeologous Recombination and Gene Evolution

Though compelling evidence indicates a central importance of recombination during the evolution of plant genomes, our knowledge of its rates and patterns is far from comprehensive (Gaut et al., 2007). Here, we revealed a perplexing and intriguing pattern of gene evolution on chromosome A and its descendants. Homeologous genes on these chromosomes have much larger rates of gene conversion and evolve in concert to a much greater degree than do those on other chromosomes (Wang et al., 2009). Previous reports dated the ages of the duplicated segments based on the molecular distance between homeologs, which led to misleading estimates that duplication was relatively recent (~5–25 mya; Goff et al., 2002; Rice Chromosomes 11 and 12 Sequencing Consortia, 2005; Wang et al., 2005; Yu et al., 2005). Here, we showed that the molecular distance between homeologs reflects the frequency of gene conversion rather than the age of the duplication event producing the homeologs, with different segments (strata) of the chromosomes evolving restrictions on recombination at different ages.

Gene conversion was previously shown to affect gene evolution with whole genome data sets (Gao and Innan, 2004; Kellis et al., 2004; Teshima and Innan, 2004; Sugino and Innan, 2005; Wang et al., 2007, 2009) and often proposed to occur between members of very conservative gene families (Liu et al., 1987; Liao, 2000; Galtier, 2003). Therefore, gene conversion may be related to gene conservation. Indeed, gene conversion helps to keep homeologous sequences similar, which may increase the likelihood of further conversion events between the affected homeologs. However, whereas homeologs experiencing illegitimate recombination tend to have a small molecular distance from one another, the distance from their respective orthologs in other species tends to be larger than those of genes not affected by illegitimate recombination (Wang et al., 2009). This is seemingly incongruous; however, it can be explained by classical theory about duplicated gene evolution (Innan and Kondrashov, 2010). Gene conversion keeps homeologs similar in sequence, which may preserve functional redundancy. When there is such functional redundancy, a mutation in one gene may be less likely to affect an organism's fitness and, therefore, more likely to be fixed in offspring. Thus, gene conversion may increase gene evolutionary rates as shown in our previous analysis (Wang et al., 2009). Sequential episodes of gene conversion may each accelerate gene evolution. Such conversion episodes may occur more frequently in the highly similar homeologous strata revealed in the present analysis than in the more divergent strata or in other regions of the genome. This agrees with a recent proposition in *Oryza* species suggesting continuous gene conversion in the so-called young strata here (Jacquemin et al., 2009) as well as with our earlier report that rice genes have experienced conversion in the past 400,000 years since the divergence of subspecies *japonica* and *indica* (Wang et al., 2007).

Recently, a comparative analysis of the R11 and R12 distal regions of the short arms was performed in the genus *Oryza*

(Jacquemin et al., 2009). Homeologous pairs were isolated for nine genes in all *Oryza* genome types as well as in two outgroup species. Phylogenetic analysis of these genes supported our proposition of recurrent and ongoing homeologous recombination in the RSA stratum, keeping the homeologous sequences highly similar. We reanalyzed phylogenetic trees constructed for nine duplicated genes (Jacquemin et al., 2009) and found increased intergenomic divergence among *Oryza* genomes. A total of nine gene losses may have occurred independently in the analyzed genomes, suggesting the RSA regions to have experienced rapid evolution accelerated by homeologous recombination. This finding is consistent with what we have shown above for the rice and sorghum genomes.

### Possible Factors Underlying Unusual Evolution of One Chromosome Pair?

What structural or functional properties could be responsible for the singular evolutionary history of R11/12 and S5/8 chromosomes and their orthologs in other cereals, relative to all other chromosomes in the same nuclei? Several characteristics of the chromosomes noted above suggest two directions that (in our opinion) are high priorities for further investigation.

First, extensive gene loss in RSA-B and SSA, greater than anywhere else in either the rice or sorghum genome and accompanied by elevated evolutionary rates of the remaining genes, may facilitate speciation in that the loss of alternative copies of duplicated genes leads to reproductive isolation (Werth and Windham, 1991; Lynch and Force, 2000). The recently suggested interrelationship between reproductive isolation and autoimmune responses (Bombliet et al., 2007; Yin et al., 2008) draws attention to the finding that orthologs R11 and S5 each contains ~25% of the nucleotide binding site–Leu-rich repeat (nonself recognition) genes still under expansion through duplication (see Supplemental Data Set 1H online) in their respective genomes (Zhou et al., 2004; Paterson et al., 2009).

Second, a high level of concerted evolution, associated stratification of chromosomal segments, and extensive homeologous gene loss are all characteristics of sex chromosomes in humans (Lahn and Page, 1999), chickens (Lawson Handley et al., 2006), fungi (Charlesworth, 2002), and plants (Ming and Moore, 2007). For example, the evolution of human sex chromosomes from ancestral autosomes appears to have involved stepwise segmental recombination suppression, producing four distinct strata with ongoing recombination at one terminus during male meiosis (Lahn and Page, 1999). If it proves true that CSB formed shortly before the 70-mya grass genome duplication (as our data tenuously suggest), then its formation may have involved divergence among homologs rather than homeologs (Figure 3), yet another parallel with the evolution of heteromorphic sex chromosomes.

## METHODS

### Materials

Rice (*Oryza sativa*), sorghum (*Sorghum bicolor*; version 1.0), and *Brachypodium distachyon* sequences were downloaded from the RAP2 database (<http://rgp.dna.affrc.go.jp/E/index.html>) and the Department of



Energy Joint Genome Institute website (<http://www.jgi.doe.gov/>). BACs on rice fingerprint contig 77 (<http://www.genome.arizona.edu/fpc/rice/>) were downloaded from GenBank ([www.ncbi.nlm.nih.gov/Genbank/](http://www.ncbi.nlm.nih.gov/Genbank/)). Maize (*Zea mays*) genomes and annotations were downloaded from the Arizona Genome Institute (<http://ftp.maizesequence.org/>). By searching R11 genes predicted with BLAT with default parameters (Kent, 2002), we found three highly similar homeologs on R12, which were missed from the RAP2 data set and involved in our following analysis. The three genes were referred as Os12g0100305, Os12g0100805, and Os12g0101355, as to their R11 homeologs Os11g0100300, Os11g0100800, and Os11g0101350 (see Supplemental Data Set 2 online).

### Inferring Colinear Homologs

We used BLASTP to search for potential anchors (E value < 1e-5; top five matches) between every possible pair of chromosomes in multiple genomes. The homologous pairs are used as the input for MCscan (Tang et al., 2008b), which is a new program to infer multiple chromosomal colinearity. The built-in scoring scheme for MCscan is  $\min(-\log_{10}E \text{ value}, 50)$  for every matching gene pair and  $-1$  for each 10-kb distance between anchors. The resulting syntenic chains were evaluated using a procedure adopted by ColinearScan (Wang et al., 2006), and E value < 1e-10 was used as a significance cutoff. We enriched the colinear gene data set by inferring more small homologous blocks by running ColinearScan to detect pairwise chromosome homology. The parameter, the maximum gap length, was set to be 40 intervening genes between neighboring genes in colinearity on both chromosomes. The homologous blocks containing three colinear genes were merged into the previous data set. Gene clusters that contain 10 or more genes in a chromosome were removed from the present analysis, for gene redundancy may lead to an aberrantly fast evolutionary rate (e.g., *R* genes; Michelmore and Meyers, 1998). Eventually, we obtained data sets of the colinear gene pairs between any two chromosomes and homologous quartets residing on all four chromosomes of R11, R12, S5, and S8.

### Evolutionary Analysis

Protein sequences were aligned using ClustalW (Thompson et al., 1994), and the protein alignment was translated into DNA coding sequence alignment. Only the aligned homologs having >40% amino acid identity and >50 amino acid no-gap columns were involved in the present analysis. Ks and Ka were estimated using an evolutionary pathway approach (Nei and Gojobori, 1986) implemented in PAML (Yang, 1997). Gepard (Krumstiek et al., 2007), implementing a suffix tree algorithm, was used to make the dot plots of homeologous sequences (word size set to be 16 bp).

### Detecting Gene Conversion

We performed a phylogenetic analysis of the homologous quartets to infer gene conversion (Wang et al., 2009) and consequently infer DNA recombination. Since the duplicated chromosomal segments were created in the common ancestor of grasses, if there was no gene conversion we anticipate that the rice-sorghum orthologs are more similar to one another than to their respective homeolog in each species. However, if there was conversion between two homeologs, we would find they are more similar to one another than to their respective ortholog, resulting in aberrant tree topology. Sequence divergent percentage was estimated to characterize gene tree topology. Bootstrap tests of 1000 repetitive random samples were performed based on aligned DNA coding sequences.

### Inferring Homeologous Strata

To investigate recombination suppression between homeologous chromosomes, we devised a one-dimensional spatial clustering algorithm to

cluster the homeologous gene pairs into subgroups along the chromosomes according to Ks values. Initially, we defined an array with Ks as its elements and divided it into  $L$  subgroups, each containing  $n$  elements except the last one, which may contain  $<n$  elements. The clustering procedure is composed of two subprocedures: the iterative element shifting procedure, which is to moderately adjust the boundaries between neighboring groups by conditionally shifting the elements to its neighboring subgroup, and the iterative merging procedure, which is to conditionally merge neighboring subgroups into a larger subgroup. For example, to adjust the left boundary of subgroup  $S_i$ , in the shifting procedure if the averaged distance of the left boundary element  $b$  in the subgroup to the other elements in this subgroup

$$d(b, S_i - b) = |b - \sum_{x \in S_i/b} x / (|S_i| - 1)|$$

is larger than that to the left neighboring subgroup

$$d(b, S_{i-1}) = |b - \sum_{x \in S_{i-1}} x / |S_{i-1}||$$

the element  $b$  will be shifted to the left neighboring subgroup. The shifting procedure is iteratively implemented until no boundary could be altered. Only  $B$  left/right boundary elements could be adjusted for each subgroup. In the merging procedure, we merge two neighboring subgroups according to two criteria to accommodate absolute and relative differences, respectively. First, we calculate the distance between two neighboring subgroups,  $S_{i-1}$  and  $S_i$ ,

$$d(S_{i-1}, S_i) = | \sum_{x \in S_{i-1}} x / |S_{i-1}| - \sum_{y \in S_i} y / |S_i| |$$

and the merging ratio between them

$$r(S_{i-1}, S_i) = d(S_{i-1}, S_i) / \min \left( \sum_{x \in S_{i-1}} x / |S_{i-1}|, \sum_{y \in S_i} y / |S_i| \right)$$

If the merging ratio is  $>R$ , the neighboring groups  $S_{i-1}$  and  $S_i$  that have the smallest distance would be merged if  $d(S_{i-1}, S_i) < D$ . The absolute and relative criteria adopted here ensure reasonable clustering of subgroups with small and large averaged element values, which is often evolutionarily interesting. This merging procedure is also iteratively implemented until no neighboring groups could be merged.

The spatial clustering algorithm can be formulated as follows.

Step 1. Define an element array  $A$  and subdivide it into subgroups  $A = (S_1, S_2, \dots, S_L)$ .

Step 2. Iterative shifting procedure: adjusting the boundary elements between neighboring subgroups until no boundary change can be made.

Step 3. Iterative merging procedure: merging the nearest neighboring groups according to the absolute and relative criteria.

Step 4. If the present subgroups are different from those when step 2 starts, go to step 2; or go to step 5.

Step 5. Stop.

When the above algorithm was implemented to cluster Ks of homeologous gene pairs, we set the parameters:  $B = 2$ ,  $D = 0.2$ ,  $R = 2$ ,  $n = 10$ . To reduce the effect from random variation, only the Ks < 2 values were involved and the largest 10% Ks values were not used to calculate the mean of Ks. Finally, the subgroups containing fewer than eight genes were manually merged into neighboring subgroups having the closer mean Ks value.

### Dating Evolutionary Strata

The median Ks values for homeologous gene pairs in these respective strata were used to date each strata by assuming the grass gene synonymous substitution rate to be 6.1e-8 per site per year (Gaut, 1998).

### R-Like Genes

RGAs on chromosomes 11 and 12 in the present data set were located by searching against the previously reported two data sets at BLAST E value < 1e-10 (Zhou et al., 2004; Rice Chromosomes 11 and 12 Sequencing Consortia, 2005).

### Newly Annotated Genes

High DNA similarity between the termini of R11/12 short arms facilitates the annotation of three additional genes (Os12g0100305, Os12g0100805, Os12g0101355) on R12 according to their homeologs on R11. Their sequences are in Supplemental Data Set 2 online.

### Accession Numbers

Sequence data from this article can be found in the GenBank/EMBL databases under the accession numbers given in Supplemental Data Sets 1 and 2 online.

### Supplemental Data

The following materials are available in the online version of this article.

**Supplemental Figure 1.** Alignment of Homologous Chromosomes.

**Supplemental Data Set 1A.** Alignment of Genes on Rice Chromosomes 11 and 12 and Their Respective Homologous Chromosomes or Chromosomal Segments from Sorghum, Maize, and *Brachypodium*.

**Supplemental Data Set 1B.** Strata and Syntenic Homeologous Genes on Chromosomes R11/12.

**Supplemental Data Set 1C.** Strata and Syntenic Homeologous Genes on S5/8.

**Supplemental Data Set 1D.** Strata and Homologous Quartets on R11/12 and S5/8.

**Supplemental Data Set 1E.** Repetitive Sequences around Break Points at the Boundaries of Young Rice Strata.

**Supplemental Data Set 1F.** Species-Specific Gene Pairs.

**Supplemental Data Set 1G.** Quartets of Homologous Genes in Synteny between R11, R12, and Their *Brachypodium* Quartets.

**Supplemental Data Set 1H.** Nucleotide Binding Site–Leu-Rich Repeat–Resistant Gene Analogs Affected by Recent Duplication Events on Chromosomes R11/12 and S5/8.

**Supplemental Data Set 2.** Newly Annotated Genes Based on Sequence Similarity.

### ACKNOWLEDGMENTS

We thank Lifeng Lin for art work. We appreciate financial support from the U.S. National Science Foundation (Grant MCB–1021718) and the J.S. Guggenheim Foundation to A.H.P. and from the China National Science Foundation (Grant 30971611) and the China-Hebei New Century 100 Creative Talents Project to X.W.

Received October 20, 2010; revised December 6, 2010; accepted January 5, 2011; published January 25, 2011.

### REFERENCES

- Bombliès, K., Lempe, J., Eppe, P., Warthmann, N., Lanz, C., Dangel, J.L., and Weigel, D.** (2007). Autoimmune response as a mechanism for a Dobzhansky-Muller-type incompatibility syndrome in plants. *PLoS Biol.* **5**: e236.
- Bowers, J.E., et al.** (2005). Comparative physical mapping links conservation of microsynteny to chromosome structure and recombination in grasses. *Proc. Natl. Acad. Sci. USA* **102**: 13206–13211.
- Bowers, J.E., Chapman, B.A., Rong, J., and Paterson, A.H.** (2003). Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature* **422**: 433–438.
- Bozza, C.G., and Pawlowski, W.P.** (2008). The cytogenetics of homologous chromosome pairing in meiosis in plants. *Cytogenet. Genome Res.* **120**: 313–319.
- Brachypodium Genome Sequencing Initiative.** (2010). Genome sequencing and analysis of the model grass *Brachypodium distachyon*. *Nature* **463**: 763–768.
- Caceres, M., Ranz, J.M., Barbadilla, A., Long, M., and Ruiz, A.** (1999). Generation of a widespread *Drosophila* inversion by a transposable element. *Science* **285**: 415–418.
- Charlesworth, B.** (2002). The evolution of chromosomal sex determination. *Novartis Found. Symp.* **244**: 207–219; discussion 220–224, 253–257.
- Devos, K.M., Pittaway, T.S., Reynolds, A., and Gale, M.D.** (2000). Comparative mapping reveals a complex relationship between the pearl millet genome and those of foxtail millet and rice TAG. *Theor. Appl. Genet.* **100**: 190–198.
- Ding, D.Q., Yamamoto, A., Haraguchi, T., and Hiraoka, Y.** (2004). Dynamics of homologous chromosome pairing during meiotic prophase in fission yeast. *Dev. Cell* **6**: 329–341.
- Feldman, M., Liu, B., Segal, G., Abbo, S., Levy, A.A., and Vega, J.M.** (1997). Rapid elimination of low-copy DNA sequences in polyploid wheat: a possible mechanism for differentiation of homoeologous chromosomes. *Genetics* **147**: 1381–1387.
- Freeling, M., and Thomas, B.C.** (2006). Gene-balanced duplications, like tetraploidy, provide predictable drive to increase morphological complexity. *Genome Res.* **16**: 805–814.
- Fu, H., Zheng, Z., and Dooner, H.K.** (2002). Recombination rates between adjacent genic and retrotransposon regions in maize vary by 2 orders of magnitude. *Proc. Natl. Acad. Sci. USA* **99**: 1082–1087.
- Gaeta, R.T., Pires, J.C., Iniguez-Luy, F., Leon, E., and Osborn, T.C.** (2007). Genomic changes in resynthesized *Brassica napus* and their effect on gene expression and phenotype. *Plant Cell* **19**: 3403–3417.
- Galtier, N.** (2003). Gene conversion drives GC content evolution in mammalian histones. *Trends Genet.* **19**: 65–68.
- Gao, L.Z., and Innan, H.** (2004). Very low gene duplication rate in the yeast genome. *Science* **306**: 1367–1370.
- Gaut, B.S.** (1998). Molecular clocks and nucleotide substitution rates in higher plants. *Evol. Biol.* **30**: 93–120.
- Gaut, B.S., Wright, S.I., Rizzon, C., Dvorak, J., and Anderson, L.K.** (2007). Recombination: an underappreciated factor in the evolution of plant genomes. *Natl. Rev.* **8**: 77–84.
- Goff, S.A., et al.** (2002). A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science* **296**: 92–100.
- Innan, H.K., and Kondrashov, F.** (2010). The evolution of gene duplications: classifying and distinguishing between models. *Nat. Rev. Genet.* **11**: 97–108.
- International Rice Genome Sequencing Project** (2005). The map-based sequence of the rice genome. *Nature* **436**: 793–800.
- Jacquemin, J., Laudíé, M., and Cooke, R.** (2009). A recent duplication

- revisited: phylogenetic analysis reveals an ancestral duplication highly-conserved throughout the *Oryza* genus and beyond. *BMC Plant Biol.* **9**: 146.
- Kellis, M., Birren, B.W., and Lander, E.S.** (2004). Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. *Nature* **428**: 617–624.
- Kent, W.J.** (2002). BLAT—The BLAST-Like Alignment Tool. *Genome Res.* **4**: 656–664.
- Krumsiek, J., Arnold, R., and Rattei, T.** (2007). Gepard: A rapid and sensitive tool for creating dotplots on genome scale. *Bioinformatics* **23**: 1026–1028.
- Lahn, B.T., and Page, D.C.** (1999). Four evolutionary strata on the human X chromosome. *Science* **286**: 964–967.
- Lawson Handley, L.J., Hammond, R.L., Emaresi, G., Reber, A., and Perrin, N.** (2006). Low Y chromosome variation in Saudi-Arabian hamadryas baboons (*Papio hamadryas hamadryas*). *Heredity* **96**: 298–303.
- Liao, D.** (2000). Gene conversion drives within genic sequences: concerted evolution of ribosomal RNA genes in bacteria and archaea. *J. Mol. Evol.* **51**: 305–317.
- Liu, T.J., Liu, L., and Marzluff, W.F.** (1987). Mouse histone H2A and H2B genes: four functional genes and a pseudogene undergoing gene conversion with a closely linked functional gene. *Nucleic Acids Res.* **15**: 3023–3039.
- Lynch, M., and Force, A.G.** (2000). The origin of interspecific genomic incompatibility via gene duplication. *Am. Nat.* **156**: 590–605.
- Lysak, M.A., Berr, A., Pecinka, A., Schmidt, R., McBreen, K., and Schubert, I.** (2006). Mechanisms of chromosome number reduction in *Arabidopsis thaliana* and related Brassicaceae species. *Proc. Natl. Acad. Sci. USA* **103**: 5224–5229.
- Michelmore, R.W., and Meyers, B.C.** (1998). Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. *Genome Res.* **8**: 1113–1130.
- Ming, R., and Moore, P.H.** (2007). Genomics of sex chromosomes. *Curr. Opin. Plant Biol.* **10**: 123–130.
- Nei, M., and Gojobori, T.** (1986). Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.* **3**: 418–426.
- Paterson, A.H., et al.** (2009). The Sorghum bicolor genome and the diversification of grasses. *Nature* **457**: 551–556.
- Paterson, A.H., Bowers, J.E., and Chapman, B.A.** (2004). Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. *Proc. Natl. Acad. Sci. USA* **101**: 9903–9908.
- Paterson, A.H., Bowers, J.E., Peterson, D.G., Estill, J.C., and Chapman, B.A.** (2003). Structure and evolution of cereal genomes. *Curr. Opin. Genet. Dev.* **13**: 644–650.
- Rice Chromosomes 11 and 12 Sequencing Consortia** (2005). The sequence of rice chromosomes 11 and 12, rice in disease resistance genes and recent gene duplications. *BMC Biol.* **3**: 20.
- Salse, J., Abrouk, M., Bolot, S., Guilhot, N., Courcelle, E., Faurat, T., Waugh, R., Close, T.J., Messing, J., and Feuillet, C.** (2009). Reconstruction of monocotyledonous proto-chromosomes reveals faster evolution in plants than in animals. *Proc. Natl. Acad. Sci. USA* **106**: 14908–14913.
- Sankoff, D., Zheng, C., and Zhu, Q.** (2010). The collapse of gene complement following whole genome duplication. *BMC Genomics* **11**: 313.
- Scannell, D.R., Byrne, K.P., Gordon, J.L., Wong, S., and Wolfe, K.H.** (2006). Multiple rounds of speciation associated with reciprocal gene loss in polyploid yeasts. *Nature* **440**: 341–345.
- Schnable, P.S., et al.** (2009). The B73 maize genome: complexity, diversity, and dynamics. *Science* **326**: 1112–1115.
- Simillion, C., Vandepoele, K., Van Montagu, M.C., Zabeau, M., and Van de Peer, Y.** (2002). The hidden duplication past of *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* **99**: 13627–13632.
- Singh, N.K., et al.** (2007). Single-copy genes define a conserved order between rice and wheat for understanding differences caused by duplication, deletion, and transposition of genes. *Funct. Integr. Genomics* **7**: 17–35.
- Sugino, R.P., and Innan, H.** (2005). Estimating the time to the whole-genome duplication and the duration of concerted evolution via gene conversion in yeast. *Genetics* **171**: 63–69.
- Tang, H., Bowers, J.E., Wang, X., Ming, R., Alam, M., and Paterson, A.H.** (2008a). Synteny and colinearity in plant genomes. *Science* **320**: 486–488.
- Tang, H., Bowers, J.E., Wang, X., and Paterson, A.H.** (2010). Angiosperm genome comparisons reveal early polyploidy in the monocot lineage. *Proc. Natl. Acad. Sci. USA* **107**: 472–477.
- Tang, H.B., Wang, X.Y., Bowers, J.E., Ming, R., Alam, M., and Paterson, A.H.** (2008b). Unraveling ancient hexaploidy through multiply-aligned angiosperm gene maps. *Genome Res.* **18**: 1944–1954.
- Teshima, K.M., and Innan, H.** (2004). The effect of gene conversion on the divergence between duplicated genes. *Genetics* **166**: 1553–1560.
- Thomas, B.C., Pedersen, B., and Freeling, M.** (2006). Following tetraploidy in an *Arabidopsis* ancestor, genes were removed preferentially from one homeolog leaving clusters enriched in dose-sensitive genes. *Genome Res.* **16**: 934–946.
- Thompson, J.D., Higgins, D.G., and Gibson, T.J.** (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**: 4673–4680.
- Wang, X., Shi, X., Hao, B., Ge, S., and Luo, J.** (2005). Duplication and DNA segmental loss in the rice genome: implications for diploidization. *New Phytol.* **165**: 937–946.
- Wang, X., Shi, X., Li, Z., Zhu, Q., Kong, L., Tang, W., Ge, S., and Luo, J.** (2006). Statistical inference of chromosomal homology based on gene colinearity and applications to *Arabidopsis* and rice. *BMC Bioinformatics* **7**: 447.
- Wang, X., Tang, H., Bowers, J.E., Feltus, F.A., and Paterson, A.H.** (2007). Extensive concerted evolution of rice paralogs and the road to regaining independence. *Genetics* **177**: 1753–1763.
- Wang, X., Tang, H., Bowers, J.E., and Paterson, A.H.** (2009). Comparative inference of illegitimate recombination between rice and sorghum duplicated genes produced by polyploidization. *Genome Res.* **19**: 1026–1032.
- Werth, C.R., and Windham, M.D.** (1991). A model for divergent, allopatric speciation of polyploid pteridophytes resulting from silencing of duplicate-gene expression. *Am. Nat.* **137**: 515–526.
- Woodhouse, M.R., Schnable, J.C., Pedersen, B.S., Lyons, E., Lisch, D., Subramanian, S., and Freeling, M.** (2010). Following tetraploidy in maize, a short deletion mechanism removed genes preferentially from one of the two homologs. *PLoS Biol* **8**: e1000409.
- Yang, Z.** (1997). PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput. Appl. Biosci.* **13**: 555–556.
- Yin, T., Difazio, S.P., Gunter, L.E., Zhang, X., Sewell, M.M., Woolbright, S.A., Allan, G.J., Kelleher, C.T., Douglas, C.J., Wang, M., and Tuskan, G.A.** (2008). Genome structure and emerging evidence of an incipient sex chromosome in *Populus*. *Genome Res.* **18**: 422–430.
- Yu, J., et al.** (2005). The genomes of *Oryza sativa*: A history of duplications. *PLoS Biol.* **3**: e38.
- Zhou, T., Wang, Y., Chen, J.Q., Araki, H., Jing, Z., Jiang, K., Shen, J., and Tian, D.** (2004). Genome-wide identification of NBS genes in japonica rice reveals significant expansion of divergent non-TIR NBS-LRR genes. *Mol. Genet. Genomics* **271**: 402–415.

**Seventy Million Years of Concerted Evolution of a Homoeologous Chromosome Pair, in Parallel,  
in Major Poaceae Lineages**

Xiyin Wang, Haibao Tang and Andrew H. Paterson  
*Plant Cell* 2011;23;27-37; originally published online January 25, 2011;  
DOI 10.1105/tpc.110.080622

This information is current as of May 19, 2019

<b>Supplemental Data</b>	<a href="/content/suppl/2011/01/13/tpc.110.080622.DC1.html">/content/suppl/2011/01/13/tpc.110.080622.DC1.html</a>
<b>References</b>	This article cites 59 articles, 20 of which can be accessed free at: <a href="/content/23/1/27.full.html#ref-list-1">/content/23/1/27.full.html#ref-list-1</a>
<b>Permissions</b>	<a href="https://www.copyright.com/ccc/openurl.do?sid=pd_hw1532298X&amp;issn=1532298X&amp;WT.mc_id=pd_hw1532298X">https://www.copyright.com/ccc/openurl.do?sid=pd_hw1532298X&amp;issn=1532298X&amp;WT.mc_id=pd_hw1532298X</a>
<b>eTOCs</b>	Sign up for eTOCs at: <a href="http://www.plantcell.org/cgi/alerts/ctmain">http://www.plantcell.org/cgi/alerts/ctmain</a>
<b>CiteTrack Alerts</b>	Sign up for CiteTrack Alerts at: <a href="http://www.plantcell.org/cgi/alerts/ctmain">http://www.plantcell.org/cgi/alerts/ctmain</a>
<b>Subscription Information</b>	Subscription Information for <i>The Plant Cell</i> and <i>Plant Physiology</i> is available at: <a href="http://www.aspb.org/publications/subscriptions.cfm">http://www.aspb.org/publications/subscriptions.cfm</a>