Meiotic Recombination Between Paralogous *RBCSB* Genes on Sister Chromatids of *Arabidopsis thaliana*

John G. Jelesko,*[,1] Kristy Carter,* Whitney Thompson,* Yuki Kinoshita[†] and Wilhelm Gruissem[‡]

**Plant Pathology, Physiology, and Weed Science Department, Fralin Biotechnology Center, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061-0346, [†]National Institute of Genetics, Agricultural Genetics, Mishima, Shizuoka, 411-8540, Japan and [‡]Institute of Plant Sciences, Swiss Federal Institute of Technology, ETH Zentrum, CH-8092, Zurich, Switzerland*

Manuscript received October 10, 2003
Accepted for publication October 20, 2003

ABSTRACT

Paralogous genes organized as a gene cluster can rapidly evolve by recombination between misaligned paralogs during meiosis, leading to duplications, deletions, and novel chimeric genes. To model unequal recombination within a specific gene cluster, we utilized a synthetic *RBCSB* gene cluster to isolate recombinant chimeric genes resulting from meiotic recombination between paralogous genes on sister chromatids. Several F_1 populations hemizygous for the *synthRBCSB1* gene cluster gave rise to Luc^+ F_2 plants at frequencies ranging from 1 to 3×10^{-6}. A nonuniform distribution of recombination resolution sites resulted in the biased formation of recombinant *RBCS3B/1B::LUC* genes with nonchimeric exons. The positioning of approximately half of the mapped resolution sites was effectively modeled by the fractional length of identical DNA sequences. In contrast, the other mapped resolution sites fit an alternative model in which recombination resolution was stimulated by an abrupt transition from a region of relatively high sequence similarity to a region of low sequence similarity. Thus, unequal recombination between paralogous *RBCSB* genes on sister chromatids created an allelic series of novel chimeric genes that effectively resulted in the diversification rather than the homogenization of the *synthRBCSB1* gene cluster.

MOST genes in plants are members of small multigene families. When each member is located at a different position in the genome it will tend to evolve independently of the other unlinked members. However, when a multigene family is organized as a gene cluster (*i.e.*, closely linked paralogous genes) two paralogous genes can misalign during meiosis and recombine to alter the gene cluster in four ways: create a deletion on one chromosome, create a duplication on the other, and create two reciprocal chimeric genes (one on each chromosome). This effect of a single recombination event within a gene cluster can have important implications for the evolution of the gene cluster. Comparisons of a particular gene cluster in distantly related species often show significant patterns of gene duplication and deletion. For example, the *HOX* gene cluster in Drosophila can be correlated with four duplications of the cluster in humans. Within each of the duplicated human *HOX* clusters, additional apparent gene duplications and deletions occurred since divergence from a common ancestor (Ruddle *et al.* 1994). Likewise, evidence for recombination within gene clusters is found in several human genetic disorders. For example, many cases of β-thalassemia (Collins and Weissman 1984) and red-green color blindness (Nathans *et al.* 1986a,b) in humans result from deletions within gene clusters caused by an apparent unequal crossover event between two paralogous members of the gene cluster. Newly formed alleles of the Drosophila *bobbed* locus are either deletions or duplications of the rDNA gene cluster caused by unequal crossing over (Hawley and Marcus 1989). The maize *R-r* gene complex is another example of linked paralogous genes (Robbins *et al.* 1991; Walker *et al.* 1995) that undergo misalignment and unequal crossing over (Dooner and Kermicle 1971) at frequencies ranging from 0.25 to 3.9×10^{-4}.

Using artificial gene duplications in *Saccharomyces cerevisiae* it was shown that unequal crossing over is most frequent between homologous chromosomes (Jackson and Fink 1985; Maloney and Fogel 1987). Unequal crossing over either within a chromosome (*i.e.*, intrachromatid exchanges) or between sister chromatids occur ~5- to 10-fold less frequently than interhomolog exchanges (Jackson and Fink 1985; Maloney and Fogel 1987). Interallelic (*i.e.*, between different alleles of the same gene on homologous chromosomes) gene conversion events are frequently associated with crossing over of flanking genetic markers (Hurst *et al.* 1972; Symington and Petes 1988) and thus gene conversion and crossing over may be mechanistically coupled processes. The double-strand break (DSB) model (Szostak *et al.* 1983) was proposed, in part, to explain how crossing over and gene conversion could be mechanistically

[1]*Corresponding author:* Fralin Biotechnology Center, Virginia Polytechnic Institute and State University, West Campus Dr., Blacksburg, VA 24061-0346. E-mail: jelesko@vt.edu

related. In this model, recombination begins with a double-strand break on one of the chromosomes that is subsequently processed to form two Holliday junctions (HJs; Holliday 1964). Depending upon which of the opposing strands are cut, resolution of each HJ will result in either a crossover or a noncrossover product.

Research on human minisatellites provides additional insights into germ-line unequal crossing over and gene conversion processes (Bois and Jeffreys 1999). Although minisatellites differ from one to another in specific details, several common features are exemplified by the well-characterized MS32 minisatellite. This minisatellite is a 29-bp core sequence that is tandemly repeated up to several hundred times. Each core repeat unit can show four different polymorphic states and the number of core repeats is highly variable from one allele to another. The MS32 minisatellite shows a high frequency of germ-line instability, reaching up to 0.008 recombinants per sperm (Jeffreys *et al.* 1994). Minisatellite instability appears to be a meiotic gene conversion-like process, whereby sequences from one allele are added to the other allele, without an exchange of flanking markers (Jeffreys *et al.* 1994). *Bona fide* unequal and equal crossing over between different minisatellite alleles occurs ~270 times less frequently than the gene conversion-like events (Jeffreys *et al.* 1998). In contrast, intraallelic recombinants (either intrachromosomal looping out or unequal sister chromatid exchange) show several important differences from the interallelic recombinants. For example, using minisatellite variant repeat analysis (Jeffreys *et al.* 1990, 1991), intraallelic recombinants frequently result in simple deletions of repeat units, without alteration in the remaining individual repeats (Jeffreys *et al.* 1990). Also, the positioning of the intraallelic deletions in the MS32 minisatellite do not show the 5′ polarity observed with interallelic gene conversion events, but are more evenly distributed over the entire repeat (Jeffreys *et al.* 1990). Intraallelic recombinants (*i.e.*, intrachromosomal recombinants) occur less frequently than the germ-line gene conversion-like events (Jeffreys *et al.* 1990, 1994). Given that minisatellites are generally located in noncoding subtelomeric regions of chromosomes (Bois and Jeffreys 1999), the degree to which minisatellite recombination/gene conversion events accurately model recombination within protein coding regions remains to be determined.

Similar to the artificial gene duplication experiments in yeast, nonallelic mutant reporter gene duplications or nonfunctional overlapping fragments of a reporter gene were used in plants to determine the frequency and character of unequal crossing over and/or gene conversion events (reviewed in Lichtenstein *et al.* 1994; Puchta *et al.* 1994). While most of these approaches provide an estimate of the frequency of unequal crossing over and gene conversion events in somatic plant tissues, only two reports determined the frequency of germinal (*i.e.*, meiotic) events leading to activation of a reporter gene (Assaad and Signer 1992; Tovar and Lichtenstein 1992). Both germinal unequal crossing over and gene conversion events were identified, but the results did not distinguish between inter- or intrachromosomal recombination.

We developed a synthetic gene cluster technology for *Arabidopsis thaliana* to identify recombination between authentic paralogous eukaryotic genes by coupling the formation of a chimeric gene to the activation of a firefly luciferase (*LUC*) gene, thereby producing a bioluminescent phenotype in the plant (Jelesko *et al.* 1999). Using a transgenic synthetic *RBCSB* gene cluster (*synthRBCSB1*) in hemizygous plants, we identified three luciferase positive (Luc^+) F_2 seedlings that contained a chimeric *RBCS3B/1B::LUC* gene that had formed by meiotic unequal crossing over between sister chromatids. Using polymorphic sequences between the parental *RBCS1B* and *RBCS3B* genes, the recombination resolution sites were mapped to a region spanning the intron 1/exon 2 boundary (Jelesko *et al.* 1999). While our approach demonstrated the feasibility of synthetic gene cluster technology, our initial sample of recombinant plants was insufficient to allow conclusions about either the frequency or the character of meiotic unequal crossing over between sister chromatids. Here we report results from a significantly expanded analysis of recombinant plants that provides new insights into meiotic recombination between paralogous genes located on sister chromatids.

MATERIALS AND METHODS

Plasmid construction: The *cre-nos* gene in pED23 (Dale and Ow 1990) was amplified using oligos oJGJ151 (CCATCGATG TCCAATTTACTGACCG) and oJGJ152 (GAAGATCTAATCG CCATCTTCC) with *Pfu* thermostable DNA polymerase (Promega, Madison, WI); manufacturer's buffers; and the following thermocycling conditions: 25 cycles of 1-min denaturation at 94°, 2-min annealing at 50°, and 2-min extension at 72°, followed by 1 cycle composed of a 7-min extension at 72°. The resulting PCR fragment was blunt-end cloned into pBlueScript SK+ (Stratagene, La Jolla, CA) cut with *Eco*RV, yielding pJGJ154 containing the modified 5′ *Cla*I-*cre2-nos-Bgl*II allele. The *cre2-nos* insert was sequenced on both strands, which revealed a single-base-pair mutation consisting of an A → G change in the first position of codon 207, resulting in a threonine-to-alanine amino acid change in the predicted protein. Therefore, the mutant gene was renamed *cre2*. Two pollen-specific promoters (Albani *et al.* 1990) were amplified from 100 ng *Brassica napus* cv. Wistar genomic DNA using either oJGJ153 (AAAGAATTCAATTAAATACTTACTATTTTTTTCAT ATAATC) or oJGJ155 (GCCGAATTCTAAAAATAGCAATAAC) oligonucleotide primer in conjunction with the oJGJ154 (TATCT CGAGAAAACATAAGAAGTATTCTTAC) primer and *Taq* DNA polymerase (Promega). The denaturation and extension conditions were 94° for 1 min and 72° for 2 min, respectively, for all cycles, whereas the annealing temperature was 49° for the first cycle and decreased 3° in each of the next 2 cycles. From cycle 4 to 35 the annealing temperature was 55°. The resulting PCR fragments were TA cloned into pCRII (Invitrogen, Carls-

bad, CA), yielding pJGJ139 containing the *Bp4D* promoter and pJGJ141 containing the *Bp4F* promoter. The *Bp4D* and *Bp4F* promoters were subcloned upstream of the TMV Ω-fragment-*gus-nos* gene in pSLJ4K1 (Jones *et al.* 1992) as an *Eco*RI-*Xho*I fragment, resulting in pJGJ147 and pJGJ144, respectively. pJGJ158 (*Bp4D-cre2-nos*) and pJGJ155 (*Bp4F-cre2-nos*) were generated by replacing the *Cla*I-*gus-nos-Bam*HI fragment from pJGJ147 and pJGJ144 with the *Cla*I-*cre2-nos-Bgl*II fragment from pJGJ154, respectively. pJGJ159 containing the CaMV 35S promoter-*cre2-nos* construct was generated by similarly replacing the *gus-nos* fragment from pSLJ4K1 with the *Cla*I-*cre2-nos-Bgl*II fragment from pJGJ154. The CaMV 35S promoter-*cre2-nos*, *Bp4D* promoter-*cre2-nos*, and *Bp4F* promoter-*cre2-nos* genes were cut out of pJGJ59, pJGJ158, and pJGJ155, respectively, as *Eco*RI-*Hin*dIII fragments and ligated into the binary vector pSLJ7292 (similarly cut) to yield pJGJ164, pJGJ163, and pJGJ160, respectively.

Generating transgenic plant lines and subsequent crosses: The creation of the transgenic AtJGJ203.10 (*synthRBCSB1-10* allele) and AtJGJ203.15 (*synthRBCSB1-15* allele) was previously described (Jelesko *et al.* 1999). Twelve micrograms of genomic DNA from both lines were restricted with *Xho*I, which cuts only once at the most 3′ end of the construct, and probed with a 1.2-kb *Eco*RI fragment derived from the *LUC* gene in pJGJ184 (Jelesko *et al.* 1999) that was randomly labeled with [^{32}P]dCTP using the PrimeItII random labeling kit (Stratagene). The blot was washed under high-stringency conditions and imaged on an ABI phosphorimager. Both lines showed a single primary hybridizing band (data not shown), indicating that the *synthRBCSB1* gene cluster was present as a single-copy insert. The Km-resistance trait segregated as a single genetic trait in both lines during the T_2 generation (*i.e.*, AtJGJ203.10 showed 64 Km^r to 20 Km^s seedlings and AtJGJ203.15 showed 62 Km^r to 19 Km^s seedlings). Transgenic Arabidopsis AtJGJ164.19 (*CaMV35S-cre2-nos-19*), AtJGJ163.1 (*Bp4D-cre2-nos-1*), and AtJGJ 160.1 (*Bp4F-cre2-nos-1*) were generated by the floral dip method (Bechtold *et al.* 1993) using *Agrobacterium tumefaciens* strain GV3101 harboring plasmids pJGJ164, pJGJ163, and pJGJ160, respectively. The Col-0 line used in these transformations was derived from a single seed isolate (143.11) obtained from the Staskawicz laboratory at the University of California, Berkeley. Transgenic lines showing an ~3:1 segregation ratio of Km^r to Km^s seedlings in the F_2 generation were used to develop homozygous F_3 lines showing 100% Km^r segregation. Crosses were performed using F_3 homozygous transgenic lines on three-day-old emasculated flowers that were covered with pollination bags both prior to and immediately after pollination. At least 2000 F_1 seeds from each cross were planted and grown to maturity to yield large F_2 seed bulks. The F_2 seed bulks were collected into four to five independent lots to ensure that Luc$^+$ plants isolated from independent F_2 seed lots would represent truly independent recombination events.

Isolation of Luc$^+$ seedlings and characterization of *RBCS3B/1B::LUC* chimeric gene sequences: Approximately 150–200 mg of F_2 seed (7500–10,000 F_2 seeds) were distributed on a 20 × 20-cm piece of Whatman 3MM chromatography paper resting on felt pads saturated with 50 ml of 1× Hoagland's solution in Jiffy 232 trays fitted with 246 domes (Jiffy Products). Trays were incubated for 2 nights at 4° and then put under fluorescent lights on a 16-hr light/8-hr dark cycle for 5–7 days at room temperature. F_2 seedlings were assayed for *in vivo* luciferase activity and single Luc$^+$ seedlings were isolated. Genomic DNA was extracted from Luc$^+$ F_2 seedlings and recombinant chimeric *RBCS(2* or *3)B/1B::LUC* fragments were amplified by PCR and subcloned as previously described (Jelesko *et al.* 1999). The cloned PCR fragments were sequenced using the Perkin-Elmer (Norwalk, CT) Big Dye cycle sequencing kit and the following oligonucleotide primers: M13 forward and reverse primers, oJGJ13 and oJGJ14 (Jelesko *et al.* 1999); oJGJ50 (CTATCTTACCTCCCTGAC); oJGJ52 (TAATAATGA TTAGTAGAC); oJGJ53 (GAATGGAGCGACCATGGT); oJGJ54 (CTTCTCCGCAACAAATGGATTC); and oJGJ55 (TTGTCCAG TACCGTCCATCGTAG). The cycle sequencing reactions were run on an ABI 377 DNA automated sequencing machine and the overlapping contigs were assembled and analyzed using the Lasergene DNA Star software package. A two-proportion test of significance for recombination frequencies was performed using the MINITAB program, Windows version 13.1 (Minitab, State College, PA). Because of the small population sizes, we utilized a two-tailed exact binomial test to determine the significance of our observed events against the expected, using the following function in Microsoft Excel: *P* value is equal to "=IF(p0*N>X,2*BINOMDIST(X,N,p0,TRUE),2*(1-BINOMDIST(X-1,N,p0,TRUE)))", where *p*0 is the expected probability, *N* is the total events, and *X* is the events in the sample. The expected probability for any given interval was calculated as the length of the given interval divided by the total length of homology available for recombination.

RESULTS

Frequency of meiotic recombination between sister chromatids: The frequency of unequal meiotic recombination between paralogous *RBCSB* genes on sister chromatids was determined by measuring the rate at which a transgenic synthetic *RBCSB* (*synthRBCSB1*) gene cluster yielded Luc$^+$ F_2 seedlings. Figure 1A shows the organization of the *synthRBCSB1* gene cluster that was used in these experiments. Two independent transformed lines homozygous for the *synthRBCSB1* gene cluster (*synthRBCSB1-10 and synthRBCSB1-15*) were crossed to several different lines to create hemizygous F_1 populations of at least 2000 plants (see materials and methods). Within cells undergoing meiosis the *synthRBCSB1* gene was present in two copies, each on identical sister chromatids. The duplicated *synthRBCSB1* gene clusters could then potentially misalign and undergo unequal crossing over between the inactive Δ*RBCS1B::LUC* gene and either the *RBCS2B* or *RBCS3B* genes. As illustrated in Figure 1B, recombination between misaligned Δ*RBCS1B::LUC* and *RBCS3B* genes would create a recombinant chimeric *RBCS3B/1B::LUC* gene and a duplicated *RBCS2B* gene. The chimeric gene imparts a Luc$^+$ phenotype to a plant because it contains a functional *RBCS3B* promoter and reconstituted exon 1 coding region upstream of the Δ*RBCS1B::LUC* gene (Jelesko *et al.* 1999). Given the relative order of the paralogous *RBCSB* genes within the *synthRBCSB1* gene cluster and the hemizygous nature in the F_1 generation, the majority of expected Luc$^+$ recombinants should be generated by recombination between paralogous *RBCSB* genes located on sister chromatids. This screen should not detect intrachromosomal looping-out events because the resulting activated luciferase gene would reside on a circular extrachromosomal fragment lacking a centromere and thus be mitotically unstable and lost in a manner previously shown (Peterhans *et al.* 1990).

Large F_2 populations (1–4 million seedlings) were

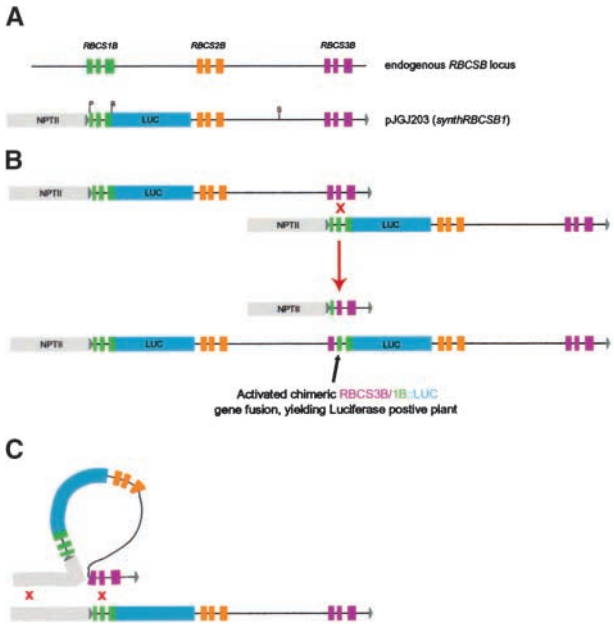

FIGURE 1.—Unequal crossing over of the *synthRBCSB1* gene cluster. (A) Endogenous *RBCSB* gene cluster on chromosome 5 of *A. thaliana*. Colored boxes indicate *RBCSB* exons (green, *RBCS1B*; orange, *RBCS2B*; and purple, *RBCS3B*). Triangles indicate lox sites. Plasmid pJGJ203 contained the *synthRBCSB1* gene cluster used to create transgenic lines AtJGJ203.10 and AtJGJ203.15. (B) Formation of chimeric *RBCS3B/1B::LUC* genes by a single unequal crossover between sister chromatids. The red X's indicates a selected crossover with or without an associated gene conversion event. (C) Formation of chimeric *RBCS3B/1B::LUC* genes either by a double unequal crossover between sister chromatids or by a gene conversion event of the upper sister chromatid replacing sequences on the lower chromatid between the two recombination initiation sites (*i.e.*, red crosses).

screened for *in vivo* luciferase activity. Table 1 shows the number of F_2 Luc^+ seedlings from eight independent crosses. The *synthRBCSB1-10* allele was crossed with two different ecotypes of Arabidopsis (Col-0 and L*er*-0) and accession CS3219 (a supposed *Cardaminopsis petraea* accession, Arabidopsis Biological Resource Center, Columbus, OH). The CS3219 × *synthRBCSB1* F_1 population was highly fertile and resembled other intraspecific *A. thaliana* crosses, suggesting that the CS3219 accession was not *C. petraea*, but rather a misidentified *A. thaliana* accession. The frequency of Luc^+ F_2 seedlings ranged from 1.2 to 3.0 × 10^{-6} recombinants per F_2 seedling.

To obtain additional measures of unequal crossing over within the *synthRBCSB1-10* allele, F_2 seedlings derived from crosses with three other transgenic Arabidopsis lines containing a mutant *cre2-nos* transgene [under the control of either the constitutive *CaMV35S* promoter or one of two Bp4 pollen-specific promoters (*Bp4D* or *Bp4F*)] were assayed for *in vivo* luciferase activity. These crosses were part of an unsuccessful gene-targeting strategy, but provided additional F_1 and F_2 populations for measuring unequal crossing over within the *synthRBCSB1-10* locus residing on sister chromatids. The observed frequencies of Luc^+ F_2 seedlings in these populations were quite similar to those observed with nontransgenic Col-0 and L*er* lines (two-proportion test, $P = 0.524$), indicating that the *cre2-nos* gene had no significant effect in these experiments. Thus, six separate crosses with the *synthRBCSB1-10* transgenic locus yielded a similar frequency of Luc^+ F_2 seedlings (1.0–3.6 × 10^{-6}).

Two crosses using an independent transgenic line (*synthRBCSB1-15*), in which the cluster was inserted at a different locus, were also examined (Table 1). These crosses yielded F_2 Luc^+ seedlings at frequencies of 2.9–3.0 × 10^{-6}, which were not statistically different from the above six *synthRBCSB1-10* crosses (two-proportion test, $P = 0.42$). These experiments indicated that the *synthRBCSB1* gene cluster yielded consistent rates of meiotic unequal crossing over between sister chromatids with two independent transgenic lines. Therefore, the observed frequencies were averaged to give a final estimate of the meiotic unequal exchange rate of 2.2 ± 1.0 × 10^{-6} per F_1 meiosis. The frequency of Luc^+ F_2 seedlings observed in these experiments was similar to a previous cross between *synthRBCSB1-10* and Col-0 (JELESKO *et al.* 1999). The F_2 Luc^+ seedlings were approximately evenly distributed over several independent

TABLE 1

Frequency of Luc^+ seedlings in several independent F_2 populations

Cross	Pollen genotype	Ovule genotype	Luciferase positive F_2 seedlings	Total F_2 seedlings screened	Meiotic recombination frequency
XJGJ317	Col-0 (CS1092)	*synthRBCSB1-10*	6	3,132,500	1.9 × 10^{-6}
XJGJ318	L*er*-0 (CS20)	*synthRBCSB1-10*	5	4,082,500	1.2 × 10^{-6}
XJGJ380	*C. petraea* (CS3219)	*synthRBCSB1-10*	5	1,640,000	3.0 × 10^{-6}
XJGJ198	*synthRBCSB1-10*	*Bp4Dprom-cre2-nos-1*	4	1,125,000	3.6 × 10^{-6}
XJGJ191	*synthRBCSB1-10*	*Bp4Fprom-cre2-nos-1*	1	1,050,000	1.0 × 10^{-6}
XJGJ189	*synthRBCSB1-10*	*CaMV35Sprom-cre2-nos-1*	3	1,125,000	2.7 × 10^{-6}
XJGJ197	*synthRBCSB1-15*	*Bp4Dprom-cre2-nos-1*	3	1,050,000	2.9 × 10^{-6}
XJGJ201	*synthRBCSB1-15*	*Bp4Fprom-cre2-nos-1*	4	1,350,000	3.0 × 10^{-6}

F_2 seed lots, confirming that each recombinant was an independent meiotic event and not a somatic event in one F_1 plant that subsequently produced many F_2 progeny with the same recombinant allele.

Mapping of the meiotic recombination resolution sites: To confirm that the Luc^+ plants contained chimeric *RBCSB* genes, chimeric *RBCSB::LUC* alleles were sequenced from 25 independent Luc^+ lines representing all eight crosses. All of the sampled Luc^+ lines contained the *RBCS3B* promoter and *RBCS3B* exon 1 sequences at the 5′ end of the chimeric gene. There was no addition or deletion of nucleotides within the recombinant chimeric *RBCS3B/1B* intervals, indicating that all of the sequences present in the chimeric genes originated from parental *RBCS3B* and *RBCS1B* sequences. Several Luc^+ recombinants from this report were shown by Southern blot analysis to contain a novel 6.5-kb *Sph*I *LUC* hybridizing fragment indicative of an *RBCS1B::LUC-RBCS2B* gene duplication within the *synthRBCSB1* gene cluster (data not shown). No chimeric *RBCS2B/1B::LUC* genes were identified. The absence of *RBCS2B/1B::LUC* chimeras was likely due to the fact that the *Cla*I site used to position the *RBCS2B* gene within the *synthRBCSB1* gene cluster is located only 133 bp upstream of the mapped *RBCS2B* transcription start site (Krebbers *et al.* 1988) and thus may not contain sufficient *RBCS2B* promoter domains to efficiently initiate transcription of an *RBCS2B/1B::LUC* recombinant. This would result in insufficient expression of luciferase activity and thus such recombinants would not be detected in this genetic screen.

The recombination resolution sites were localized for each chimeric *RBCS3B/1B::LUC* gene on the basis of the distribution of single nucleotide polymorphisms that distinguish the *RBCS1B* and *RBCS3B* genes. The red bars in Figure 2 illustrate the intervals where recombination resolution occurred for each chimeric *RBCS3B/1B::LUC* gene. Most of the recombination resolution sites mapped to either the intron 1/exon 2 boundary or the intron 2/exon 3 boundary. No resolution sites were identified within exon 1, most of exon 2, and the 5′ regions of both introns 1 and 2. The majority of the characterized chimeric *RBCS3B/1B::LUC* genes showed a polymorphic base-pair distribution consistent with a single resolution site responsible for the formation of the chimeric gene. Therefore, these were referred to as simple recombinants.

The distribution of recombination resolution breakpoints appeared to be nonuniformly distributed over the region of shared *RBCS3B* and *RBCS1B* sequence homology. A χ^2 test was performed to test the hypothesis that the overall distribution of resolution sites was uniform. The region of shared *RBCSB* homology was divided into four approximately equal intervals (rounding to the nearest polymorphic site) and the chi-square test was used to determine if the number of observed recombination resolution sites in each of the four intervals

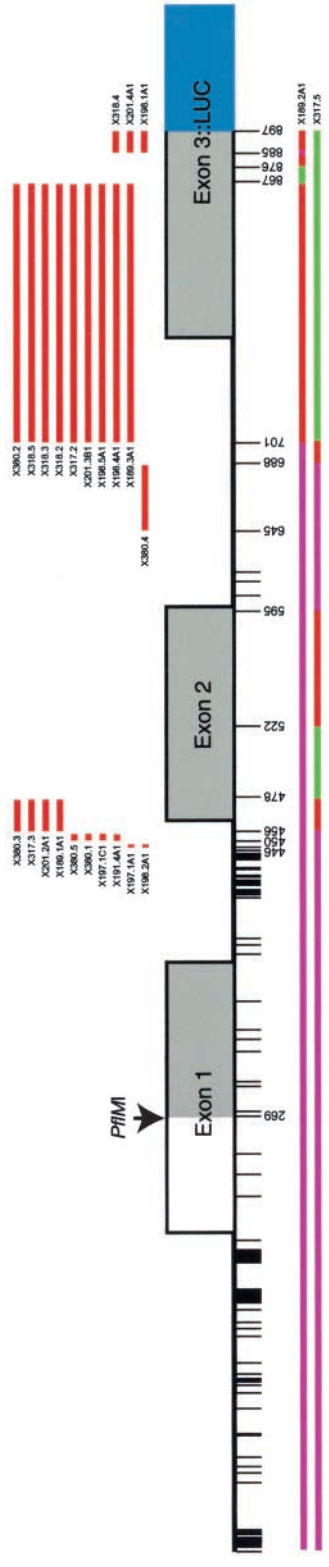

Figure 2.—Mapping of recombination resolution sites. Gray regions indicate DNA sequences available for homologous recombination between the *RBCS3B* and Δ*RBCS1B* genes. The black arrow indicates the *Pfl*MI site used to generate the 5′ end of the Δ*RBCS1B::LUC* gene. The blue box indicates the *LUC* reporter gene (not to scale). Vertical black lines indicate single polymorphic bases between the *RBCS3B* and Δ*RBCS1B* genes used to map the resolution sites. Red horizontal lines are the intervals where recombination resolution occurred for each labeled chimeric gene. Purple and green lines indicate *RBCS3B*- and *RBCS1B*-specific sequences, respectively, for the XJGJ317.5 and XJGJ189.2A1 genes, while the interspersed red lines indicate either crossover or gene conversion tract.

TABLE 2

Statistical analysis of observed simple resolution sites

Interval	Interval length (bp)	No. of recombinants	*P* value
701–868	167	9	0.2610
522–595	73	0	0.1171
478–522	44	0	0.3773
645–688	43	1	0.0700
456–478	22	4	0.0156
885–897	12	3	0.0185
456–478	6	4	0.0001
446–450	4	2	0.0187

The *P* values were calculated using the exact binomial test to determine the statistical significance (two-tail test) of the observed compared to the expected number of resolution sites within each interval.

was similar to the expected number of recombination resolution sites, estimated by the fractional proportion of each interval relative to the entire homologous region available for recombination. The calculated chi-square value of 13.54 (with 3 d.f., $P = 0.004$) confirmed that the resolution sites were not uniformly distributed over the region of homology.

The frequency of plant somatic extrachromosomal recombination increases with increasing length of DNA sequence identity (Lyznik *et al.* 1991; Puchta and Hohn 1991). Therefore, one could model the position and frequency of recombination resolution sites as a probability function based upon the fractional length of a given interval between polymorphic bases relative to the total length of *RBCS3B* and *RBCS1B* sequences available for recombination. In this model, relatively long regions of sequence identity would have proportionally more recombination resolution sites and very short intervals would have proportionally fewer resolution sites. Table 2 shows that 10 of 23 simple crossovers were in good agreement ($P \geq 0.05$) with this model. For example, the longest interval of sequence identity (positions 701–867) showed the highest frequency of resolution sites. Furthermore, the three next longest intervals of sequence identity showed frequencies of recombination resolution that were in reasonable agreement with this model (Table 2). Upon visual inspection of Figure 2, however, there was a conspicuous absence of recombination resolution sites within most of exon 2. As noted above, intervals 478–522 and 522–595 alone did not show a significant underrepresentation of resolution sites. However, when the region from 478 to 595 was evaluated as a signal uninterrupted interval (*i.e.*, ignoring the polymorphism at position 522), there was a significant absence of expected resolution sites (two-tailed exact binomial test, $P = 0.018$) within this combined region.

On the other hand, several intervals showed significantly more resolution sites than would be predicted from their proportional length of sequence identity. Ten of the 23 simple recombinants mapped to quite short intervals (4, 6, or 22 bp; between positions 446, 450, 456, and 478, respectively) near the intron 1/exon 2 boundary. The frequency of recombination resolution within each of these three intervals was significantly higher than that predicted on the basis of their fractional interval length alone ($P \leq 0.05$; Table 2). Similarly, three simple recombinants that mapped within a relatively short interval (885–897) adjacent to the exon-3*::LUC* boundary were also statistically overrepresented on the basis of the fractional interval length ($P = 0.0187$). Thus, more than half of the observed simple recombinants did not fit a simple model of recombination frequency and position being determined by the fractional length of sequence identity.

The DSB model of meiotic recombination predicts that a gene conversion patch could form by either the repaired gap or the unrepaired mismatched bases formed during HJ branch migration. Only 2 of the 25 mapped recombinants showed *RBCS3B* polymorphisms interspersed within *RBCS1B* sequences (or vice versa) that would be consistent with gene conversion patches. For example, Figure 2 illustrates that up to position 456, the XJGJ317.5 chimeric gene sequence showed a clear pattern of contiguous *RBCS3B* polymorphisms. However, between positions 478 and 701 there is an interspersion of *RBCS1B* and *RBCS3B* polymorphisms that are consistent with a resolution site and a closely associated gene conversion patch. Likewise, the chimeric XJGJ189.2A1 gene also showed an intermingling of *RBCS1B* and *RBCS3B* polymorphisms between positions 867 and 897. In both cases, it was not possible to distinguish the location of the recombination resolution site from the associated conversion patch. The deduced chimeric gene sequence from these two Luc^+ recombinants was not due to artifacts caused by template switching during the PCR amplification of genomic DNA because independent PCR clones derived from independent PCR reactions of genomic DNA were sequenced and yielded the same polymorphic base pattern. Alternatively, it is formally possible that the chimeric genes in XJGJ317.5 and XJGJ189.2A1 originated by three closely positioned crossover events. In either case, the frequency of this type of more complex recombinant was low relative to the simple recombinants.

Preferential shuffling of intact exons during chimeric *RBCS3B/1B::LUC* gene formation: The nonuniform distribution of simple recombination resolution sites in the chimeric *RBCS3B/1B::LUC* genes had a marked affect on the composition of the resulting chimeric genes. All of the simple recombinants contained chimeric *RBCS3B/1B* genes composed of a different array of intact parental exons. These ranged from chimeric genes containing *RBCS3B* exon 1 with *RBCS1B* exon 2 and

exon 3 domains (*e.g.*, XJGJ380.3) to those that were composed of only *RBCS3B*-specific sequences up to the *LUC* gene fusion junction (*e.g.*, XJGJ318.4). Thus, the nonuniform distribution of single resolution sites resulted in chimeric *RBCS3B/1B::LUC* genes in which the *RBCSB* exons were in effect shuffled as intact gene-specific modules.

Interestingly, this pattern was not observed for the two complex recombinants. Recombinant XJGJ317.5 showed a chimeric exon 2 containing polymorphic bases from both parental *RBCSB* genes (Figure 2). Similarly, recombinant XJGJ189.2A1 contained both *RBCS1B* and *RBCS3B* polymorphic bases within exon 3. However, because the parental polymorphic bases were at conservative wobble positions, the chimeric exons in recombinants XJGJ317.5 and XJGJ189.2A1 did not result in chimeric polypeptide sequences.

DISCUSSION

Formation of chimeric *RBCS3B/1B::LUC* genes: Reports of stably integrated reporter gene-based recombination substrates show somatic recombination events during most stages of plant development (Peterhans *et al.* 1990; Assaad and Signer 1992; Swoboda *et al.* 1994; Puchta *et al.* 1995; Lucht *et al.* 2002; Kovalchuk *et al.* 2003). These are observed as small, medium, and large recombinant sectors on cotyledons and leaves. These chimeric sectors generally do not result in transmission of the recombinant gene to the next generation because they do not occur in tissues giving rise to the inflorescence meristem. However, occasionally whole seedlings showing the recombinant phenotype arise and this trait is stably transmitted to the next generation (Assaad and Signer 1992; Tovar and Lichtenstein 1992; Ries *et al.* 2000; Kovalchuk *et al.* 2003) and may have arisen from somatic recombination events just prior to or just after meiosis. For this reason, it is worth considering whether the chimeric *RBCS3B/1B::LUC* genes may have formed by one of three different recombination events: (i) a somatic recombination event prior to the formation of the gametes, (ii) a somatic recombination event occurring very early during embryogenesis in cells that give rise to the shoot apical meristem, or (iii) a meiotic recombination event. Premeiotic somatic recombination events can be estimated by two different approaches. For example, frequent somatic recombination within the *synthRBCSB1* gene cluster should show Luc^+ sectors in fully expanded vegetative leaves, yet these were not observed (data not shown). Alternatively, if a somatic recombination event occurred as late as carpel initiation, all resulting ovules should contain the recombinant gene. Thus within that F_2 seed lot, there would be at least 20 Luc^+ seedlings, all of which would contain the same recombinant allele (*i.e.*, showing the same recombination resolution site). An analogous situation could occur during anther development, also resulting in a disproportionate number of Luc^+ seedlings in a particular F_2 seed lot. In contrast to this prediction, none of the F_2 seed lots in this report showed an aberrant distribution of Luc^+ seedlings and most of the chimeric genes from any given cross had different resolution sites. Thus, neither physical nor genetic methods indicated a significant frequency of premeiotic somatic recombination that was transmitted germinally. Additionally, if somatic recombination was frequent during embryo development, there should be a class of chimeric embryos in which somatic recombination occurred in only one cotyledon. This could give rise to a large Luc^+ sector on the cotyledon of the germinating seedling sufficient to allow its isolation. However, all of the F_3 progeny from such an isolate should be Luc^- because the chimeric gene would not be present in the shoot apical meristem. In contrast to this prediction, all of the F_2 Luc^+ lines showed reasonable Mendelian segregation ratios of the Luc^+ trait in the F_3 generation. In summary, several predicted classes of somatic recombinants (*e.g.*, in both premeiotic and late embryonic tissue) were not observed. Therefore, given the absence of both physical and genetic evidence for several predicted classes of somatic recombinants, together with the stochastic nature of somatic recombination, it was not likely that somatic recombination during early embryogenesis significantly contributed to the formation of Luc^+ lines in this report. The more likely conclusion was that meiotic recombination was the primary cause for the formation of the chimeric *RBCS3B/1B::LUC* genes. Consistent with this assertion, Assaad and Signer (1992) showed that meiotic recombination rates were on the order of 10-fold higher than somatic recombination rates.

The chimeric *RBCS3B/1B::LUC* genes described in this report could have formed by three types of meiotic unequal exchanges between misaligned paralogous genes on sister chromatids. The simplest type was a potential single crossover event between misaligned *RBCS3B* and Δ*RBCS1B::LUC* genes as illustrated in Figure 1B. Alternatively, alignment of *RBCS3B* and Δ*RBCS1B::LUC* genes at the 3′ end and the concomitant alignment of the *nptII* genes at the 5′ end could be resolved in either of two ways: (i) a double crossover between the aligned *nptII* and misaligned *RBCSB* genes or (ii) a contiguous gene conversion event in which the Δ*RBCS1B::LUC-RBCS2B-RBCS3B* sequences from one chromatid converted the intervening region on the other chromatid (see Figure 1C). We could not differentiate between these three possibilities because, as products of sister chromatid exchange, there were no polymorphic flanking markers to definitively evaluate whether a crossover had resolved. Moreover, the genetic screen was not designed to isolate the other three gametes formed during the meiotic recombination event that gave rise to the activated *RBC3B/1B::LUC* chimeric gene and therefore we could not analyze these other meiotic products for the various predicted changes. This type of tetrad-like

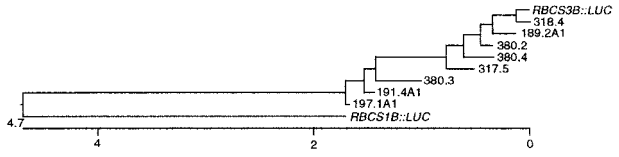

FIGURE 3.—Phylogenetic analysis of chimeric *RBCS3B/1B::LUC* genes. Chimeric *RBCS3B/1B::LUC* genes were aligned using the Clustal algorithm.

analysis is possible in Arabidopsis, using the quartet (*qrt*) mutation (PREUSS *et al.* 1994). However, given the very low frequency of unequal recombination at the *synthRBCSB1* locus (2.2×10^{-6}), it would be impractical to perform the millions of crosses with separate *qrt* pollen grains that would be required to identify the equally rare reciprocal recombinant alleles.

Regardless of the exact mechanism of recombination, our results clearly showed that novel chimeric *RBCS3B/1B::LUC* genes formed by meiotic recombination between misaligned paralogous *RBCSB* genes located on sister chromatids. Figure 3 illustrates how this resulted in an allelic series of chimeric *RBCS3B/1B::LUC* genes that were intermediate to either of the parental genes. Our results are in contrast to models of concerted evolution (ELDER and TURNER 1995), which propose that meiotic unequal intrachromosomal recombination within a gene cluster leads to homogenization of the gene cluster (LIAO 1999). Most reported analyses of concerted evolution were based upon sequence comparisons between extant gene clusters in which recombination events were inferred and the contribution of selection is not well characterized. In contrast, our analysis revealed that selected contemporary unequal recombination events between paralogous genes on sister chromatids resulted in the diversification of the gene cluster (*i.e.*, an *RBCS2B* gene duplication and the formation of nonequivalent chimeric *RBCS3B/1B::LUC* genes) rather than in its homogenization. Thus, intrachromosomal recombination within a gene cluster does not *a priori* need to result in homogenization of the gene cluster.

Low frequency of meiotic recombination between sister chromatids: Previously, we reported that the *synthRBCSB1-10* gene cluster yielded *RBCS3B/1B::LUC* recombinants at a frequency of $\sim 3 \times 10^{-6}$ per F_1 meiosis; however, this was based upon only three recombinants per 1×10^6 F_2 seedlings (JELESKO *et al.* 1999). Here we report eight additional hemizygous F_1 populations that showed statistically similar rates of meiotic unequal crossing over, with a combined average frequency of $2.2 \pm 1.0 \times 10^{-6}$. Two previous reports of meiotic recombination rates using artificial gene duplications of nonallelic inactive *nptII* genes (ASSAAD and SIGNER 1992; TOVAR and LICHTENSTEIN 1992) provide a useful comparison. In the case of an inverted mutant *nptII* gene duplication, meiotic recombinants appeared at 1.3×10^{-5} and these were formed by gene conversion events (TOVAR and LICHTENSTEIN 1992). Nonallelic *nptII* duplicated genes configured in a direct repeat also yielded gene conversion events; however, unambiguous unequal crossovers in one transgenic line (*i.e.*, DIRA1) occurred at 5.4×10^{-6} (ASSAAD and SIGNER 1992). From those experiments it was difficult to determine, however, if the gene conversion resulted from recombination between sister chromatids or homologous chromosomes. Nevertheless, the frequency of unequal sister chromatid exchange for the *synthRBCSB1* alleles was comparable to the frequency of recombination between the direct *nptII* repeats reported by ASSAAD and SIGNER (1992). Moreover, we did not observe a significant difference in the frequency of unequal exchanges between two independent *synthRBCSB1* transgenic lines. These results were in contrast to that from ASSAAD and SIGNER (1992) who observed very different rates of meiotic unequal crossing over in a second transgenic line of the same construct (*i.e.*, DIRA2). However, more independent *synthRBCSB1* transgenic lines need to be assayed to fully understand how chromosomal position contributes to the frequency of meiotic unequal sister chromatid exchange at the *synthRBCSB1* locus.

There are marked differences in the frequency of meiotic unequal exchange between homologous/paralogous genes in eukaryotes. The above estimates of meiotic unequal exchanges in Arabidopsis were at least 1000 times less frequent than estimates of meiotic unequal crossing over reported in yeast rDNA (1.2×10^{-1}; PETES 1980) or between artificially duplicated genes in yeast (2.5×10^{-3} to 1.2×10^{-1}; KLEIN and PETES 1981; KLEIN 1984; JACKSON and FINK 1985; LICHTEN *et al.* 1987; MALONEY and FOGEL 1987). The maize *R-r* complex undergoes meiotic unequal exchange ranging from 0.25 to 3.93×10^{-4} (DOONER and KERMICLE 1971). In contrast, the human MS32 minisatellite shows intrachromosomal unequal crossing over at 1.4×10^{-3}/sperm and interchromosomal unequal crossing over at 4×10^{-5}/sperm (JEFFREYS *et al.* 1990). Taken together, there does not seem to be a "typical" frequency of meiotic unequal exchange (for either inter- or intrachromosomal events). The observed frequency of meiotic unequal exchange for the *synthRBCSB1* gene cluster was at the lower end of the spectrum. Two factors may contribute to the observed low frequency of unequal exchange. The first factor may be the relatively short region of *RBCSB* homology available for recombination (*i.e.*, 629 bp). The second factor may be an effect of mismatched bases between the paralogous *RBCS3B* and *RBCS1B* genes (46 mismatched bp of a total of 629 bp). Increasing the number of mismatched bases lowers the frequency of recombination during bacterial conjugation (RAYSSIGUIER *et al.* 1989) and meiotic recombination in yeast (BORTS and HABER 1987; CHEN and JINKS-ROBERTSON 1999).

DNA sequence similarity and dissimilarity both affect the positioning of resolution sites: The observed recom-

bination resolution sites were not uniformly distributed over the 629-bp region of homology between the Δ*RBCS1B* and *RBCS3B* genes. This result was similar to the nonuniform distribution of interhomolog resolution sites between the *LEU2* and *CENIII* regions of yeast chromosome III (Symington and Petes 1988). However, the observed nonuniformity of resolution sites between the *RBCS3B* and *RBCS1B* sequences was not biased toward the 5′ end of the region of shared homology, as is the case for the human M32 minisatellite (Jeffreys *et al.* 1998). Instead, the nonuniformity of resolution sites between the homologous *RBCS3B* and *RBCS1B* sequences clustered to three distinct regions. Since homologous recombination requires regions of DNA sequence identity, longer intervals of perfect sequence identity should show proportionally more crossover events than shorter intervals of perfect identity. Approximately half of the simple recombinants fit this model (see Table 2). For example, the longest interval of sequence identity (*i.e.*, 167 bp defined by positions 701–867) was also the most frequent target for recombination resolution. The absence of mapped resolution sites between the next two longest intervals of sequence identity (478–522 and 522–595) were not statistically significant and thus generally supported this model. The region between 478 and 595 was not incapable of recombination because a complex resolution site mapped to this combined interval in recombinant XJGJ317.5. Thus, approximately half of the simple recombinants fit a model in which the frequency and position of recombination resolution was determined by the relative length of uninterrupted DNA sequence identity. These results were similar to results from transgenic yeast containing inverted repeats of paralogous chicken β-tubulin cDNAs in which the recombination resolution sites were positioned according to the length of perfect sequence identity (Chen and Jinks-Robertson 1999). Likewise, using somatic intrachromosomal recombination between duplicated cauliflower mosaic virus (CaMV) sequences to generate a functional viral infection, the most frequent interval of apparent recombination resolution was also the longest interval of DNA sequence identity (Gal *et al.* 1991). Despite the overall 92% DNA sequence identity between the *RBCS3B* and *RBCS1B* genes, it is noteworthy that they are relatively small paralogous genes with many relatively short intervals of sequence identity (Krebbers *et al.* 1988). The importance of this observation was underscored by the distribution of the other recombination resolution sites.

There was a statistically significant overrepresentation of resolution sites that mapped to two distinct regions in the chimeric *RBCS3B/1B::LUC* gene. One of these overrepresented regions was composed of two intervals 5′ to and one interval spanning the intron 1/exon 2 boundary (see Figure 2). This clustering was reminiscent of the cluster of resolution sites spanning the intron 2/exon 3 boundary. While this correlation is intriguing, it is not likely that intron/exon boundaries *per se* caused the clustering of recombination resolution sites, because the 3′ exon/5′ intron boundaries did not show a similar clustering of resolution sites. The second significantly overrepresented region was the interval within exon 3 (885–897) immediately adjacent to the *LUC* gene fusion. Both of the overrepresented regions (446–478 and 885–897) shared a similar general characteristic. Both regions were rather abrupt transitional zones between a region of relatively high sequence similarity and a region of quite low sequence similarity. Specifically, positions 446–478 are located between the highly conserved exon 2 coding region and a dense clustering of polymorphisms within intron 1. Similarly, positions 885–897 are a transition zone between the highly conserved exon 3 coding region and the *Bsm*I cloning site, after which the *RBCS3B* and Δ*RBCS1B::LUC* sequences diverge completely. These results were similar to an exceptional overrepresented interval in the yeast experiments that measured recombination between inverted repeats of paralogous chicken β-tubulin genes (Chen and Jinks-Robertson 1999) and in which the most 3′ end of the chicken β-tubulin genes adjacent to the cloning vector showed a disproportionate number of meiotic resolution sites compared to that predicted by the interval length alone. Therefore, in some regions of the paralogous *RBCSB* genes, the length of DNA sequence identity *per se* may not be as important as a perceived transition from a region of high DNA sequence similarity to a region of lower DNA sequence similarity. The "perceived" transition might be manifested as the recognition of mismatched base pairs, within heteroduplex DNA formed by HJ branch migration, in such a way as to affect the positioning and character of HJ resolution.

The mismatch repair heteroduplex rejection model first proposed by Reenan and Kolodner (1992) and later expanded by Alani *et al.* (1994) could account for the positioning of the overrepresented resolution sites between misaligned *RBCSB* genes during meiotic unequal crossing over. In the mismatch repair heteroduplex rejection model, mismatched base pairs formed by an HJ, undergoing branch migration, are recognized by highly conserved *mutS* orthologs (*i.e.*, *MSH*) and *MutL* orthologs (*i.e.*, *MLH* and *PMS*) causing the HJ to stall and/or reverse the direction of branch migration. This proposed model for scanning and sensing mismatched base pairs within heteroduplex DNA is consistent with the *in vitro* activities of several *Escherichia coli* proteins (specifically RecA, RuvA,B branch migration motor, RuvC resolvase, and the MutS,L complex) involved in DNA recombination (Bianchi and Radding 1983; Fabisiewicz and Worth 2001) and the recombination behavior of yeast *MSH2* and *MLH/PMS1* mismatch repair mutants (Borts *et al.* 1990; Reenan and Kolodner 1992; Alani *et al.* 1994; Chen and Jinks-Robertson 1999). Consistent with this proposed model, pre-

dicted *recA*, *MSH*, and *MLH/PMS1* orthologous genes are present in the Arabidopsis genome (Arabidopsis Genome Initiative 2000). Assuming the heteroduplex rejection model was operating in Arabidopsis, the results from this report suggest that relatively few mismatched bases were required to initiate HJ resolution. For example, assuming that an HJ initiated between positions 701 and 867, ~70% (*i.e.*, 9 of 13) of recombinants resolved prior to the first encountered mismatched base pair, whereas the remaining four recombinants traversed two and three mismatched bases before resolving (*i.e.*, intervals 645–688 and 885–897, respectively). Likewise, assuming that HJs formed within exon 2 and then resolved near the intron 1/exon 2 boundary, 40% resolved within one mismatch, 40% within two mismatches, and the remaining 20% within three mismatched bases. Alternatively, a similar result would occur if gap repair spanned the same regions and then immediately resolved into a crossover. Therefore, these data suggest that neither gap repair nor HJ branch migration proceeded further than three closely positioned mismatched base pairs during the formation of simple crossovers between *RBCS3B* and *RBCS1B* genes located on sister chromatids.

A potent mismatch repair mechanism acting to abrogate the processivity of HJ branch migration could also explain the observed preponderance of simple recombinants and the relative low frequency of complex recombinants isolated in this screen. Specifically, if a potent mismatch repair mechanism caused HJs to efficiently pause (or possibly reverse direction) at single base mismatches and then resolve (as either crossover or noncrossover), there would be little opportunity to form extensive tracts of heteroduplex DNA. Assuming a DSB mechanism had formed two HJ intermediates during Arabidopsis meiotic recombination, the apparent simple crossovers would be generated by one HJ resolving into a crossover and the second posited HJ resolving into a noncrossover before it could traverse a mismatched base pair. Consistent with this hypothesis, a majority of recombinants were chimeric *RBCS3B/1B* genes with a single resolution site. Extending this hypothesis to account for the formation of complex recombinants with gene conversion patches, the probability of forming a given gene conversion patch would be the product of the probabilities for the HJ to traverse each subsequent mismatched base pair during branch migration. In other words, the likelihood of forming multiple mismatched base pairs (*i.e.*, heteroduplex DNA) during HJ branch migration would become increasingly improbable. Consistent with a heteroduplex rejection model, only two complex meiotic recombinants were isolated and the number of polymorphic bases incorporated into the heteroduplex DNA in these was no more than two interspersed polymorphic base pairs. These results were similar to somatic intrachromosomal recombination within duplicated polymorphic CaMV sequences that show few complex recombinants and those that form have relatively short gene conversion tracts (Gal *et al.* 1991). Taken together, the relative paucity of complex recombinants and the short length of apparent heteroduplex sequences in those that did form were also in general agreement with a heteroduplex rejection model.

Sister chromatid exchange and *RBCSB* intergenic exon shuffling: The nonuniform distribution of mostly simple crossovers during sister chromatid exchange introduced an interesting bias to the resulting chimeric *RBCSB* genes. Specifically, the chimeric *RBCS3B/1B* genes formed by single resolution sites were composed of different combinations of intact parental exon modules. This bias resulted in a form of "exon shuffling" between paralogous *RBCSB* genes, such that intact parental exons were shuffled into new assortments. In contrast, the two complex chimeric *RBCS3B/1B* genes contained a chimeric exon, *i.e.*, polymorphic DNA sequences derived from both parental exons. These results suggest that underlying molecular processes that affect both the nonuniform positioning and the character of HJ branch migration/resolution can introduce a significant bias to the evolution of the *RBCSB* gene cluster. For example, excessive unequal sister chromatid exchanges would tend to create relatively simple chimeric *RBCSB* genes, composed of different assortments of intact parental exons.

The authors are grateful to the following persons and agencies for making this work possible: Emily Dale for supplying plasmid pED23, Valentin Parvu and Xin Zhong for statistical consulting, and James Keddy for providing *B. napus* cv. Wistar genomic DNA. We are especially grateful to Masaki Furuya who provided generous use of his single-photo video imaging equipment during the initial stages of this study and John M. McDowell for critical review of the manuscript. This work was supported by grants from the Monsanto Company and the National Science Foundation (IBN9727044) to W.G. J.G.J. was supported by a National Science Foundation Postdoctoral Fellowship in Plant Biology (DBI9404014), a Center for Global Partnership travel grant (INT9622319), and a National Institutes of Health grant (R01GM62352).

LITERATURE CITED

Alani, E., R. A. G. Reenan and R. D. Kolodner, 1994 Interactions between mismatch repair and genetic recombination in *Saccharomyces cerevisiae*. Genetics **137:** 19–39.

Albani, D., L. S. Robert, P. A. Donaldson, I. Altosaar, P. G. Arnison *et al.*, 1990 Characterization of a pollen-specific gene family from *Brassica napus* which is activated during early microspore development. Plant Mol. Biol. **15:** 605–622.

Arabidopsis Genome Initiative, 2000 Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature **408:** 796–815.

Assaad, F. F., and E. R. Signer, 1992 Somatic and germinal recombination of a direct repeat in Arabidopsis. Genetics **132:** 553–566.

Bechtold, N., J. Ellis and G. Pelletier, 1993 In planta Agrobacterium mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. C. R. Acad. Sci. **316:** 1194–1199.

Bianchi, M. E., and C. M. Radding, 1983 Insertions, deletions, and mismatches in heteroduplex DNA made by RecA protein. Cell **35:** 511–520.

Bois, P., and A. J. Jeffreys, 1999 Minisatellite instability and germline mutation. Cell. Mol. Life Sci. **55:** 1636–1648.

Borts, R. H., and J. E. Haber, 1987 Meiotic recombination in yeast alteration by multiple heterozygosities. Science **237:** 1459–1465.

Borts, R. H., W. Y. Leung, W. Kramer, B. Kramer, M. Williamson *et al.*, 1990 Mismatch repair-induced meiotic recombination requires the Pms1 gene product. Genetics **124:** 573–584.

Chen, W., and S. Jinks-Robertson, 1999 The role of the mismatch repair machinery in regulating mitotic and meiotic recombination between diverged sequences in yeast. Genetics **151:** 1299–1313.

Collins, F. S., and S. M. Weissman, 1984 The molecular genetics of human hemoglobin. Prog. Nucleic Acid Res. Mol. Biol. **31:** 315–462.

Dale, E. C., and D. W. Ow, 1990 Intra- and intermolecular site-specific recombination in plant cells mediated by bacteriophage P1 recombinase. Gene **91:** 79–85.

Dooner, H. K., and J. L. Kermicle, 1971 Structure of the *Rr* tandem duplication in maize. Genetics **67:** 427–436.

Elder, J. F. J., and B. J. Turner, 1995 Concerted evolution of repetitive DNA sequences in eukaryotes. Q. Rev. Biol. **70:** 297–320.

Fabisiewicz, A., and L. Worth, Jr., 2001 *Escherichia coli* MutS,L modulate RuvAB-dependent branch migration between diverged DNA. J. Biol. Chem. **276:** 9413–9420.

Gal, S., B. Pisan, T. Hohn, N. Grimsley and B. Hohn, 1991 Genomic homologous recombination *in planta*. EMBO J. **10:** 1571–1578.

Hawley, R. S., and C. H. Marcus, 1989 Recombinational controls of rDNA redundancy in Drosophila. Annu. Rev. Genet. **23:** 87–120

Holliday, R., 1964 A mechanism for gene conversion in fungi. Genet. Res. **5:** 282–304.

Hurst, D. D., S. Fogel and R. K. Mortimer, 1972 Conversion associated recombination in yeast. Proc. Natl. Acad. Sci. USA **69:** 101–105.

Jackson, J. A., and G. R. Fink, 1985 Meiotic recombination between duplicated genetic elements in *Saccharomyces cerevisiae*. Genetics **109:** 303–332.

Jeffreys, A. J., R. Neumann and V. Wilson, 1990 Repeat unit sequence variation in minisatellites: a novel source of DNA polymorphism for studying variation and mutation by single molecule analysis. Cell **60:** 473–486.

Jeffreys, A. J., A. Macleod, K. Tamaki, D. L. Neil and D. G. Monckton, 1991 Minisatellite repeat coding as a digital approach to DNA typing. Nature **354:** 204–209.

Jeffreys, A. J., K. Tamaki, A. Macleod, D. G. Monckton, D. L. Neil *et al.*, 1994 Complex gene conversion events in germline mutation at human minisatellites. Nat. Genet. **6:** 136–145.

Jeffreys, A. J., D. L. Neil and R. Neumann, 1998 Repeat instability at human minisatellites arising from meiotic recombination. EMBO J. **17:** 4147–4157.

Jelesko, J. G., R. Harper, M. Furuya and W. Gruissem, 1999 Rare germinal unequal crossing-over leading to recombinant gene formation and gene duplication in *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. USA **96:** 10302–10307.

Jones, J. D. G., L. Shlumukov, F. Carland, J. English, S. R. Scofield *et al.*, 1992 Effective vectors for transformation, expression of heterologous genes, and assaying transposon excision in transgenic plants. Transgenic Res. **1:** 285–297.

Klein, H. L., 1984 Lack of association between intrachromosomal gene conversion and reciprocal exchange. Nature **310:** 748–753.

Klein, H. L., and T. D. Petes, 1981 Intrachromosomal gene conversion in yeast. Nature **289:** 144–148.

Kovalchuk, I., O. Kovalchuk, V. Kalck, V. Boyko, J. Filowski *et al.*, 2003 Pathogen-induced systemic plant signal triggers DNA rearrangements. Nature **423:** 760–762.

Krebbers, E., J. Seurinck, L. Herdies, A. R. Cashmore and M. P. Timko, 1988 Four genes in two diverged subfamilies encode the ribulose-1,5-bisphosphate carboxylase small subunit polypeptides of *Arabidopsis thaliana*. Plant Mol. Biol. **11:** 745–759.

Liao, D., 1999 Concerted evolution: molecular mechanism and biological implications. Am. J. Hum. Genet. **64:** 24–30.

Lichten, M., R. H. Borts and J. E. Haber, 1987 Meiotic gene conversion and crossing over between dispersed homologous sequences occurs frequently in *Saccharomyces cerevisiae*. Genetics **115:** 233–246.

Lichtenstein, C. P., J. Paszkowski and B. Hohn, 1994 Intrachromosomal recombination between genomic repeats, pp. 95–122 in *Homologous Recombination and Gene Silencing in Plants*, edited by J. Paszkowski. Kluwer Academic Publishers, Dordecht, The Netherlands.

Lucht, J. M., B. Mauch-Mani, H. Y. Steiner, J. P. Metraux, J. Ryals *et al.*, 2002 Pathogen stress increases somatic recombination frequency in Arabidopsis. Nat. Genet. **30:** 311–314.

Lyznik, L. A., J. D. McGee, P. Y. Tung, J. L. Bennetzen and T. K. Hodges, 1991 Homologous recombination between plasmid DNA molecules in maize protoplasts. Mol. Gen. Genet. **230:** 209–218.

Maloney, D. H., and S. Fogel, 1987 Gene conversion, unequal crossing-over and mispairing at a non-tandem duplication during meiosis of *Saccharomyces cerevisiae*. Curr. Genet. **12:** 1–8.

Nathans, J., T. P. Piantanida, R. L. Eddy, T. B. Shows and D. S. Hogness, 1986a Molecular genetics of inherited variation in human color vision. Science **232:** 203–210.

Nathans, J., D. Thomas and D. S. Hogness, 1986b Molecular genetics of human color vision: the genes encoding blue, green, and red pigments. Science **232:** 193–202.

Peterhans, A., H. Schlüpman, C. Basse and J. Paszkowski, 1990 Intrachromosomal recombination in plants. EMBO J. **9:** 3437–3445.

Petes, T. D., 1980 Unequal meiotic recombination within tandem arrays of yeast ribosomal DNA genes. Cell **19:** 765–774.

Preuss, D., S. Y. Rhee and R. W. Davis, 1994 Tetrad analysis possible in Arabidopsis with mutation of the QUARTET (QRT) genes. Science **264:** 1458–1460.

Puchta, H., and B. Hohn, 1991 The mechanism of extrachromosomal homologous DNA recombination in plant cells. Mol. Gen. Genet. **230:** 1–7.

Puchta, H., P. Swoboda and B. Hohn, 1994 Homologous recombination in plants. Experientia **50:** 277–284.

Puchta, H., P. Swoboda, S. Gal, M. Blot and B. Hohn, 1995 Somatic intrachromosomal homologous recombination events in populations of plant seedlings. Plant Mol. Biol. **28:** 281–292.

Rayssiguier, C., D. S. Thaler and M. Radman, 1989 The barrier to recombination between *Escherichia coli* and *Salmonella typhimurium* is disrupted in mismatch-repair mutants. Nature **342:** 396–401.

Reenan, R. A. G., and R. D. Kolodner, 1992 Characterization of insertion mutations in the *Saccharomyces cerevisiae* MSH1 and MSH2 genes: evidence for separate mitochondrial and nuclear functions. Genetics **132:** 975–985.

Ries, G., W. Heller, H. Puchta, H. Sandermann, H. K. Seidlitz *et al.*, 2000 Elevated UV-B radiation reduces genome stability in plants. Nature **406:** 98–101.

Robbins, T. P., E. L. Walker, J. L. Kermicle, M. Alleman and S. L. Dellaporta, 1991 Meiotic instability of the *R-r* complex arising from displaced intragenic exchange and intrachromosomal rearrangement. Genetics **129:** 271–284.

Ruddle, F. H., J. L. Bartels, K. L. Bentley, C. Kappen, M. T. Murtha *et al.*, 1994 Evolution of *HOX* genes. Annu. Rev. Genet. **28:** 423–442.

Swoboda, P., S. Gal, B. Hohn and H. Puchta, 1994 Intrachromosomal homologous recombination in whole plants. EMBO J. **13:** 484–489.

Symington, L. S., and T. D. Petes, 1988 Expansions and contractions of the genetic map relative to the physical map of yeast chromosome III. Mol. Cell. Biol. **8:** 595–604.

Szostak, J. W., T. L. Orr-Weaver, R. J. Rothstein and R. W. Stahl, 1983 The double-strand-break repair model for recombination. Cell **33:** 25–35.

Tovar, J., and C. Lichtenstein, 1992 Somatic and meiotic chromosomal recombination between inverted duplications in transgenic tobacco plants. Plant Cell **4:** 319–332.

Walker, E. L., T. P. Robbins, T. E. Bureau, J. Kermicle and S. L. Dellaporta, 1995 Transposon-mediated chromosomal rearrangements and gene duplications in the formation of the maize R-r complex. EMBO J. **14:** 2350–2363.

Communicating editor: B. Bartel