Recent advances in plant recombination
Jin Li¹, An-Ping Hsia² and Patrick S Schnable²

Recombination is an essential cellular process and a source of genetic diversity. Recent studies have demonstrated the effects of various factors (e.g. DNA sequence similarity and activation of transposons) on rates of recombination and the distribution of recombination breakpoints in plants. These studies have also provided detailed characterizations of interchromatid and interhomolog recombination events. New approaches offer the promise of achieving the long-awaited goal of gene targeting in plants.

Introduction
Recombination plays an important role in the repair of DNA damage and in chromosome segregation, and can create novel haplotypes (i.e. genetic diversity). To a large extent, the mechanisms and regulation of recombination pathways in plants have been inferred from extensive studies conducted in yeast and mammalian cells.

Homologous recombination events are initiated by double-stranded breaks (DSBs) [1] and can be classified as crossovers (COs) and noncrossovers (NCOs). Although it has been hypothesized that both types of events are resolved from double Holliday junctions (DHJs) [2], evidence from yeast shows that COs and NCOs are differentially regulated during meiosis [3,4]. The kinetics of COs and NCOs also differ in mitotic recombination, and mutations can affect the proportion of the two types of recombination events [5]. It has been suggested that most NCOs arise from synthesis-dependent strand annealing rather than from the resolution of DHJs [3–5].

There are two major pathways to repair DSBs: homologous recombination (HR) and non-homologous end-joining (NHEJ). The Rad51 protein plays a central role in HR, whereas the Ku70/Ku80 heterodimer functions in NHEJ. In mammalian cells, defects in proteins that are involved in NHEJ (e.g. Ku70) can increase HR, indicating that the two pathways compete for the repair of DSBs [6,7]. In Arabidopsis, rates of HR decrease as plants age, whereas rates of point mutations increase. Coincidently, the expression of rad51 decreases while the expression of ku70 increases as plants age. These data indicate that these two major recombination pathways are developmentally regulated [8]. The understanding that these two major pathways can compete with each other has been used to increase the efficiency of gene targeting in fungi [9,10].

This review summarizes recent progress in understanding genetic modifiers that affect recombination, interchromatid and interhomolog recombination, and genome-wide patterns of recombination in plant genomes. The resulting contributions to our mechanistic understanding have the potential to increase the efficiency with which HR-based tools can be applied to plants.

Cis-acting genetic modifier
Recombination does not occur randomly in a genome. Regions where recombination rates are significantly higher or lower than the genome average are termed recombination hot- and cold-spots, respectively [11–13]. A growing body of research has sought to understand the underlying mechanisms that are responsible for the differential distribution of recombination events. Two general classes of genetic modifiers, cis and trans, contribute to variation in the distribution of recombination events.

Cis genetic modifiers are closely linked to the intervals in which they modify rates of recombination. Although it has been shown previously that genic sequence polymorphisms in plants can influence recombination in cis [13], it is not clear how sequence divergence regulates either the rate of recombination or its effects on multi-genic intervals. In Arabidopsis, divergence levels of 0.16% [14] and 1.9% [15] can lead to three- and ten-fold reductions in rates of somatic recombination, respectively. In nearly isogenic maize lines, DNA sequence and structural polymorphisms from three haplotypes of a chromosomal interval, defined by the a1 and sh2 genes, have cis-effects on both meiotic recombination rates and the distribution of recombination breakpoints [16]. Inactivation of AtMSH2 can increase rates of somatic homeologous recombination up to nine-fold, indicating that the mismatch repair system contributes to the decrease in recombination rates that is associated with DNA sequence polymorphisms [17].
Trans-acting genetic modifiers

In contrast to cis elements, trans-acting genetic modifiers are not closely linked to the interval in which they modify recombination. Trans genetic modifiers include chromatin remodelers, autonomous transposons, recombination machinery proteins, DNA-binding proteins, and other functional elements.

For example, the maize autonomous MuDR transposase can increase the rate of meiotic COs near a Mu insertion four-fold as compared with the rate of meiotic COs in plants that do not express MuDR transposase [18]. It has been hypothesized that the additional meiotic DSBs produced by the MuDR transposase are responsible for the increased rate of COs. To study the effect of trans genetic modifiers on meiotic recombination, a sequence identical a1–sh2 interval was introgressed into three unique maize lines [19]. Molecular characterization of the recombinants demonstrated that trans modifiers affected both the rates of meiotic recombination and the distribution of breakpoints across the a1–sh2 interval. It appears that these trans modifiers exert their effects in a region-specific manner because rates of recombination across a chromosomal interval in Chromosome 1S were not affected by these modifiers [19].

Interchromatid and interhomolog recombination

Recombination can occur both between sister chromatids (interchromatid recombination) and between homologous chromosomes (interhomolog recombination). In yeast, interchromatid recombination is the major pathway during mitosis, whereas interhomolog recombination is the major pathway during meiosis [20,21].

In plants, tandemly arrayed duplicates have been utilized to study relationships between interchromatid and interhomolog recombination. Using synthetic direct repeats in Arabidopsis [22], frequencies of somatic recombination in homozygous plants exceeded those in hemizygous plants by a factor of two, suggesting that both interchromatid and interhomolog recombination occur during mitosis. Molecular characterization showed that most of these somatic events had resulted from gene conversion events as opposed to unequal COs.

Although meiotic unequal COs occur naturally in plants, detailed analyses of the distribution of the resulting recombination breakpoints were not previously available. Mapping of 25 independent meiotic unequal COs in Arabidopsis plants that were hemizygous for a synthetic RBCSB gene cluster established that recombination breakpoints clustered in regions that share high sequence homology between RBCSB3B and ΔRBCSB1 [23]. The A1-b allele of maize is a naturally occurring tandem duplication that shares the same genomic location as the single-copy a1 gene. Meiotic unequal COs that involve A1-b occur preferentially between homologous chromosomes rather than between sister chromatids. Pairing of the duplicated A1-b segments with the segments in the homolog does not occur randomly. The different pairing configurations lead to different outcomes: gene loss or duplication [24**]. Because of the high frequency of tandem genetic duplications in maize [25,26] and other plant species, such as rice [27] and Arabidopsis [28], unequal COs might have greatly influenced genome structure throughout evolution [24**].

Genome-wide studies of recombination

Many earlier molecular genetic studies of recombination were focused on the characterization of recombination events in genes or small multi-genic intervals [12,13]. The availability of the Arabidopsis genome sequence (thanks to The Arabidopsis Genome Initiative 2000) has made it possible to study recombination on a genome-wide scale. Using 71 single nucleotide polymorphisms (SNPs) covering Arabidopsis chromosome 4, 1171 COs were detected in 702 F2 plants. Further analyses showed that CO rates were significantly negatively correlated with G+C content and were not significantly correlated with the presence of genes or transposons [29]. Additionally, Singer et al. [30**] used a whole-genome exon array to map 16 000 single-feature polymorphisms in a collection of Arabidopsis recombinant inbred lines. Highly variable rates of recombination per physical distance were observed on all five chromosomes. Recombination hot spots had recombination rates as much as 70-fold greater than the genome average. The Arabidopsis quartet mutant has been used to show that the fractions of interference-insensitive COs on Chromosomes 2 and 4 are significantly smaller than those on the other three chromosomes. Lam et al. [31] hypothesized that the presence of nucleolus organizing regions (NORs) on these two chromosomes reduces the need for non-interfering COs.

Although the maize genome has not yet been fully sequenced, various methods have been used to study recombination in this species on a genome-wide scale. Single-copy in situ hybridization (FISH) analysis of maize chromosome 9, in combination with an analysis of genetic maps, showed a dramatic reduction in recombination rates in pericentromeric regions [32]. Consistently, cytogenetic mapping of recombination nodules (RNs) across the ten maize chromosomes showed that distal portions of chromosome arms have the highest average number of RNs and that the frequencies of RNs decrease towards the kinetochore [33]. The RNCm map [34] was used to position 1195 expressed sequence tags (ESTs) onto the ten pachytene chromosomes. Strong correlations were observed between relative EST and RN frequencies for 2-μm intervals, suggesting that recombination is closely associated with genes in maize [35]. Similarly, in wheat, gene densities are positively correlated with recombination rates along
chromosomes [36] and recombination rates are dramatically reduced in centromeric regions [37]. These results differ from the previously discussed findings from chromosome 4 of Arabidopsis [29]. One possible explanation for this difference is that the self-fertilizing mating behavior of Arabidopsis might decrease the influence of recombination on genome organization. By contrast, maize is a strongly cross-fertilizing species that exhibits high levels of DNA sequence polymorphism. Given the dense genetic map that is available for maize [38], the completion of the on-going maize genome-sequencing project will enhance our understanding of the distribution of recombination along maize chromosomes.

Gene targeting and homologous recombination

Altering genes in their native environment is a powerful tool with which to study gene function and to modify organisms genetically. Although gene targeting has been widely applied in yeast and mice, its efficiency in plants is still not high enough for routine applications [39–41,42]. Various methods have been tested to increase the efficiency of gene targeting in plants. DSBs that are generated by a rare cutting restriction enzyme, I-SceI, can enhance homologous integration frequency at the target site [43]. However, this strategy involves transgenic target sites that are inserted into the genome at random, and thus it is not likely to target endogenous genes [44].

Recently, two groups reported the use of zinc-finger nucleases (ZFN) to target specific genome sequences for mutagenesis [45] and gene modification [46]. A ZFN consists of a set of (usually three) zinc-finger motifs and a non-specific endonuclease, such as FokI [47]. Each zinc-finger motif comprises about 30 amino acids and recognizes a DNA triplet. Dimerization of ZFNs is required for the efficient cleavage of double-strand DNA; two recognition sites in inverted orientation flanking a 6-base spacer promote cleavage [48,49]. A DSB is introduced into the target sequence when ZFNs bind to recognition sites, dimerize, and activate the endonuclease. Lloyd et al. [45] separately transformed into Arabidopsis a construct containing a three-finger ZFN under the control of a heat-shock promoter and a synthetic recognition site for ZFN. Single-locus transformant lines were selected and subjected to heat shock to activate the ZFN and the resulting mutations analyzed. The mutation frequency at the ZFN recognition site was estimated to be as high as 0.2 (average 0.08) mutations per ZFN recognition site assayed by the disruption of an EcoRI site in the targeted sequence. These mutations range from simple deletions of 1–52 bp (78%), simple insertions of 1–4 bp (13%) and deletions/insertions (8%). Using tobacco, Wright et al. [46] measured the rate of homologous recombination (HR) induced by a three-finger ZFN through the repair of a defective β-glucuronidase (GUS):NPTII site transgene with a ZFN recognition site. Protoplasts from 10 transgenic lines were electroporated with DNA that encoded ZFN and repair template donor DNA, and then screened for HR products. One of 50 transformants carried HR products without additional insertions or deletions. The high rate of targeting from both studies is encouraging. Because two inverted recognition sites are required, a three-finger ZFN actually has a recognition site of 18 bases. With a library of three-finger ZFN, it should be possible to target naturally occurring ZFN recognition sites in most of the genes in a plant genome [45**,46**].

Another approach for enhancing gene targeting efficiency relies on overexpressing endogenous plant recombination genes or heterologous HR pathway genes, but this has not typically been sufficient to alter recombination rates [50,51]. Recently, Shaked et al. [52] demonstrated that overexpressing the yeast Rad54 protein, a member of the SWI2/SNF2 chromatin remodeling gene family, greatly increased the rate of HR and at the same time seems to reduce NHEJ events. Expression of RAD54 increased the rate of gene targeting by an order of magnitude. In fungi, the disruption of Ku70 greatly increases the rate of gene targeting [9,10,53**,54]. Arabidopsis homologs of Ku70 and Ku80 [55] have been identified [56]. It is thus possible that inhibiting the NHEJ pathway could be used to increase rates of HR (and thus gene targeting) in plants.

Conclusions

The availability of additional genome sequences will enable global analyses of DSBs and recombination breakpoints, the identification of correlations between recombination breakpoints and genomic features such as DNA sequences or expression status, and the identification of homologous or novel genes that contribute to or regulate recombination pathways. We also anticipate that the combination of ZFN and the manipulation of endogenous recombination pathways will yield robust and efficient protocols for targeted gene replacement and mutagenesis in plants.

Acknowledgements

Research on recombination in PSS’s laboratory is supported by competitive grants from the US National Science Foundation (NSF) Plant Genome Initiative (award DBI-0321711) and by the National Research Initiative (NRI) of the US Department of Agriculture Cooperative State Research, Education and Extension Service, grant number 2005-35301-15715. Support was also provided by Hatch Act and State of Iowa funds.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- ** of outstanding interest

Differential timing and control


44. Puchta H, Dujon B, Hohn B: Two different but related mechanisms are used in plants for the repair of genomic double-strand breaks by homologous recombination. *Proc Natl Acad Sci USA* 1996, 93:5055-5060.


