37. Poon, L. L. M., Leung, T. N., Lau, T. K., Chow, K. C. & Lo, Y. M. D. Differential DNA methylation between fetus and mother as a strategy for detecting fetal DNA in maternal plasma. *Clin. Chem.* **48**, 35–41 (2002).
38. Grunau, C., Clark, S. J. & Rosenthal, A. Bisulfite genomic sequencing: systematic investigation of critical experimental parameters. *Nucleic Acids Res.* **29**, e65 (2001).
39. Poon, L. L. M., Leung, T. N., Lau, T. K. & Lo, Y. M. D. Presence of fetal RNA in maternal plasma. *Clin. Chem.* **46**, 1832–1834. (2000).
40. Ng, E. K. O. *et al.* Presence of filterable and nonfilterable mRNA in the plasma of cancer patients and healthy individuals. *Clin. Chem.* **48**, 1212–1217 (2002).
41. Tsui, N. B. Y., Ng, E. K. O. & Lo, Y. M. D. Stability of endogenous and added RNA in blood specimens, serum, and plasma. *Clin. Chem.* **48**, 1647–1653 (2002).
42. Ng, E. K. O. *et al.* mRNA of placental origin is readily detectable in maternal plasma. *Proc. Natl Acad. Sci. USA* **100**, 4748–4753 (2003).
43. Tsui, N. B. Y. *et al.* Systematic microarray-based identification of placental mRNA in maternal plasma: towards non-invasive prenatal gene expression profiling. *J. Med. Genet.* **41**, 461–467 (2004).
44. Oudejans, C. B. M. *et al.* Detection of chromosome 21-encoded mRNA of placental origin in maternal plasma. *Clin. Chem.* **49**, 1445–1449 (2003).
45. Go, A. T. *et al.* Detection of placental transcription factor mRNA in maternal plasma. *Clin. Chem.* **50**, 1413–1414 (2004).
46. Wong, B. C. K. *et al.* Circulating placental RNA in maternal plasma is associated with a preponderance of 5′ mRNA fragments: implications for noninvasive prenatal diagnosis and monitoring. *Clin. Chem.* **51**, 1786–1795 (2005).
47. Gupta, A. K. *et al.* Detection of fetal DNA and RNA in placenta-derived syncytiotrophoblast microparticles generated *in vitro*. *Clin. Chem.* **50**, 2187–2190 (2004).
48. Ng, E. K. O. *et al.* The concentration of circulating corticotropin-releasing hormone mRNA in maternal plasma is increased in preeclampsia. *Clin. Chem.* **49**, 727–731 (2003).
49. Masuzaki, H. *et al.* Clinical applications of plasma circulating mRNA analysis in cases of gestational trophoblastic disease. *Clin. Chem.* **51**, 1261–1263 (2005).
50. Lo, Y. M. D. *et al.* Plasma placental RNA allelic ratio permits noninvasive prenatal chromosomal aneuploidy detection. *Nature Med.* (in the press).
51. Tong, Y. K. *et al.* Noninvasive prenatal detection of fetal trisomy 18 by epigenetic allelic ratio analysis in maternal plasma: theoretical and empirical considerations. *Clin. Chem.* 13 Oct 2006 (doi:10.1373/clinchem.2006.076851).
52. Malone, F. D. *et al.* First-trimester or second-trimester screening, or both, for Down's syndrome. *N. Engl. J. Med.* **353**, 2001–2011 (2005).
53. Bianchi, D. W. At-home fetal DNA gender testing: caveat emptor. *Obstet. Gynecol.* **107**, 216–218 (2006).
54. Bianchi, D. W., Flint, A. F., Pizzimenti, M. F., Knoll, J. H. & Latt, S. A. Isolation of fetal DNA from nucleated erythrocytes in maternal blood. *Proc. Natl Acad. Sci. USA* **87**, 3279–3283 (1990).
55. Herzenberg, L. A., Bianchi, D. W., Schroder, J., Cann, H. M. & Iverson, G. M. Fetal cells in the blood of pregnant women: detection and enrichment by fluorescence-activated cell sorting. *Proc. Natl Acad. Sci. USA* **76**, 1453–1455 (1979).
56. Beroud, C. *et al.* Prenatal diagnosis of spinal muscular atrophy by genetic analysis of circulating fetal cells. *Lancet* **361**, 1013–1014 (2003).
57. Bianchi, D. W. *et al.* Fetal gender and aneuploidy detection using fetal cells in maternal blood: analysis of NIFTY I data. National Institute of Child Health and Development Fetal Cell Isolation Study. *Prenat. Diagn.* **22**, 609–615 (2002).
58. Cheung, M. C., Goldberg, J. D. & Kan, Y. W. Prenatal diagnosis of sickle cell anaemia and thalassaemia by analysis of fetal cells in maternal blood. *Nature Genet.* **14**, 264–268 (1996).
59. Bianchi, D. W. *et al.* PCR quantitation of fetal cells in maternal blood in normal and aneuploid pregnancies. *Am. J. Hum. Genet.* **61**, 822–829 (1997).
60. Chiu, R. W. K. *et al.* Noninvasive prenatal exclusion of congenital adrenal hyperplasia by maternal plasma analysis: a feasibility study. *Clin. Chem.* **48**, 778–780 (2002).
61. Lau, T. K., Lo, K. W., Chan, L. Y. S., Leung, T. Y. & Lo, Y. M. D. Cell-free fetal deoxyribonucleic acid in maternal circulation as a marker of fetal-maternal hemorrhage in patients undergoing external cephalic version near term. *Am. J. Obstet. Gynecol.* **183**, 712–716 (2000).
62. Lazar, L., Nagy, B., Ban, Z., Nagy, G. R. & Papp, Z. Presence of cell-free fetal DNA in plasma of women with ectopic pregnancies. *Clin. Chem.* **52**, 1599–1601 (2006).
63. Sekizawa, A. *et al.* Increased cell-free fetal DNA in plasma of two women with invasive placenta. *Clin. Chem.* **48**, 353–354 (2002).

Acknowledgements
The work in the authors' laboratories is supported by the Hong Kong Research Grants Council, the Innovation and Technology Fund and the Li Ka Shing Foundation.

Competing interests statement
The authors declare competing financial interests: see web version for details.

DATABASES
The following terms in this article are linked online to:
Entrez Gene: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene
CGB | *CRH* | *H19* | *HBB* | *IGF2* | *PLAC4* | *RHD* | *SERPINB5*
OMIM: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM
β-thalassaemia | Down syndrome | haemophilia A

FURTHER INFORMATION
Chinese University of Hong Kong Department of Chemical Pathology: http://www.cpy.cuhk.edu.hk
Human Epigenome Project: http://www.epigenome.org
Access to this links box is available online.

OPINION

Recombination: an underappreciated factor in the evolution of plant genomes

Brandon S. Gaut, Stephen I. Wright, Carène Rizzon, Jan Dvorak and Lorinda K. Anderson

Abstract | Our knowledge of recombination rates and patterns in plants is far from being comprehensive. However, compelling evidence indicates a central role for recombination, through its influences on mutation and selection, in the evolution of plant genomes. Furthermore, recombination seems to be generally higher and more variable in plants than in animals, which could be one of the primary reasons for differences in genome lability between these two kingdoms. Much additional study of recombination in plants is needed to investigate these ideas further.

The nuclear genomes of plants are remarkably variable in terms of characteristics such as genome size, chromosome number, gene order and gene density. The grass family (Poaceae) serves as a fitting example. Since their origin ~77 million years ago[1], the grasses have diverged to range ~55-fold in diploid genome size and at least 10-fold in diploid chromosome number[2]. Some of this genome lability can be attributed to ancient polyploid events[3]. But paleopolyploidy by itself cannot explain the structural variation that exists among the genomes of extant plants. For one thing, paleopolyploid events might be too infrequent[4]. The most recent polyploidy event in the evolutionary lineage of *Arabidopsis thaliana* occurred anywhere from 25 million years to 100 million years ago[5,6], too long ago to contribute to differences in genome size and chromosomal number between *A. thaliana* and its closest relatives.

In addition to changes in ploidy, ongoing mechanisms of gene deletion, genome rearrangement and localized gene duplications that do not involve changes in ploidy undoubtedly contribute to genomic flux. In *A. thaliana*, for example, the proportion of genes that have been duplicated by localized events is comparable to that of genes that have been duplicated by polyploidy. These mechanisms might accelerate after polyploidy events occur, but also function in their absence[7].

In recent years, plant evolutionary genomicists have focused primarily on paleopolyploid events. However, one process that we believe has not received adequate attention is recombination, which generates mutations and influences the strength of natural selection. Here we highlight the increasing body of evidence that indicates that recombination has had an important role in plant genome evolution. We begin by briefly reviewing the mutational properties of recombination, its role in natural selection and genome-wide

patterns of recombination in plants. We then explore the extent to which plant genome characteristics such as gene density and repetitive DNA are organized along recombinational gradients. Last, we discuss initial evidence that rates of recombination might be both greater and more variable in plants than some animals, thereby contributing to the increased variability of genomic characteristics of plants. We argue that these potentially crucial roles of recombination in plant genome evolution make it essential to improve our understanding of the dynamics of recombination in this kingdom.

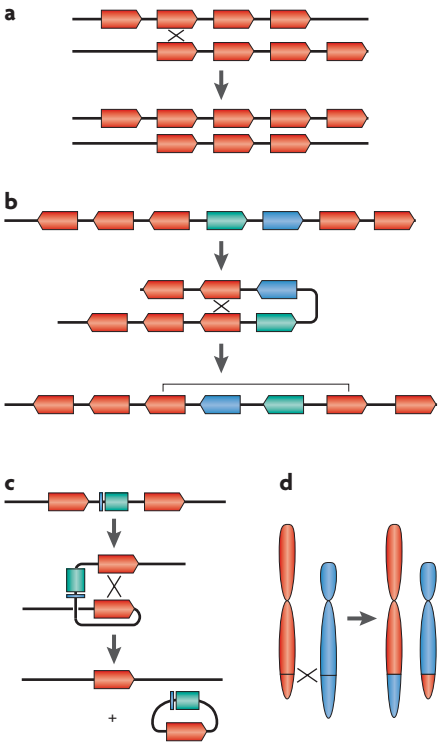


Figure 1 | **Types of mutation caused by homologous crossing over between repeated sequences.** Repeats of the same or similar sequence are shown as red boxes. Gene conversion can also occur among these related sequences. In **a**–**c**, each line represents a DNA strand from a single chromatid. **a** | Unequal crossing over between misaligned repeats on homologous chromosomes (or sister chromatids) yields an increase in repeat number for one chromosome and a decrease in repeat number for the other chromosome. **b** | Intra-strand crossing over between inverted repeats yields an inverted segment (bracket). **c** | Intra-strand crossing over between direct repeats causes a 'pop-out' of the segment of DNA between the repeats. **d** | Homologous recombination between repeated sequences on different chromosomes (represented here as unreplicated G_1 chromatids) yields a translocation.

Recombination causes mutation

Recombination is typically initiated by double-strand breaks that trigger strand exchange[8]. The resolution of this exchange can generate several types of mutation, particularly when recombination occurs between misaligned repeats. For example, unequal crossing over between sister chromatids — or between homologous chromosomes — can increase or decrease the copy number of a repetitive element (FIG. 1). Unequal crossing over can be surprisingly common. A single experimental tandem array of three genes generated copy-number variants between sister chromatids at the rate of ~10^{-6} per array per plant per F_1 meiosis in *A. thaliana*[9]. To put this number in perspective, consider that the *A. thaliana* genome contains ~1500 tandem arrays of two or more genes[10]. If the rate of ~10^{-6} is similar among arrays, at least one copy-number variant is expected to be produced in 1 out of ~700 seeds. This calculation is undoubtedly conservative, because it ignores other sources of copy-number variation, such as unequal crossing over among non-genic repetitive elements, illegitimate (non-homologous) recombination events, and meiotic recombination between homologous chromosomes, which occurs much more frequently than recombination between sister chromatids[11]. Recombination alone is therefore expected to generate substantial copy-number variation within plant populations.

Mutation by recombination is not limited to unequal crossing over between homologous chromosomes. Intrachromatid recombination between direct or inverted repeats can lead to sequence deletions and inversions and, perhaps most frequently, gene conversion[8,12] (FIG. 1). Recombination between different genomic regions (ectopic exchange) can also produce large chromosomal rearrangements. In *Drosophila melanogaster* and yeast, this mechanism has been used to engineer rearrangements, such as pericentric and paracentric inversions, and translocations[13]. To our knowledge, large-scale chromosomal events that are mediated by recombination have not been directly demonstrated in plants. Nonetheless, indirect evidence indicates that ectopic recombination contributes to rearrangements in plant genomes. For example, chromosome-specific probes have been used to 'paint' chromosome segments using fluorescence in situ hybridization (FISH) among *Brassica* species[14]. Most of the chromosomal differences between species could be explained by rearrangement of chromosome segments at or near repetitive sequences, indicating that ectopic recombination between repeat sequences leads to rearrangement. Similar evidence for ectopic recombination has been uncovered at the sequence level[15]. In fact, opportunities for ectopic exchange might be higher in plants than other eukaryotes because plants have more and larger gene families[16] that can function as templates for this process.

Although recombination is known to contribute to gross chromosomal rearrangements, changes in the copy number of repeats and even microsatellite instability[17], its role in generating single-nucleotide mutations is equivocal. Much research has been dedicated to this question in animals, primarily using molecular evolutionary approaches, which are based on the simple prediction that if recombination is mutagenic, genes in chromosomal regions with higher recombination rates will evolve more rapidly at the nucleotide level. Weak positive correlations have been documented in humans[18,19], which is consistent with the idea that recombination introduces single-nucleotide mutations, perhaps by error-prone polymerases during repair of double-strand breaks[20]. The evidence is strengthened by the observation that recombination hotspots in humans are associated with high nucleotide diversity, but these correlations could also be caused by covariates such as GC content[21]. It is even less clear whether recombination contributes to single-nucleotide mutations in *D. melanogaster*[22,23]. This question also remains unanswered in plants, for which the data are far less comprehensive than for either humans or *D. melanogaster*[24,25].

Once mutations are produced, recombination also helps to guide their evolutionary fate through its effect on natural selection. In high recombination regions, advantageous mutations rise rapidly in population frequency and deleterious mutations are purged effectively; the reverse is true in low recombination regions. To see why this is the case, imagine two new beneficial mutations that have arisen at two different loci on the same chromosome, in two different individuals of a population. In the presence of recombination, fixation of the two mutations will be more or less independent; they can recombine onto the same genetic background and can therefore be fixed simultaneously by natural selection. When there is no recombination, no single individual can contain both mutations. As a result, the two beneficial mutations compete, they cannot be simultaneously fixed in the population[26], and the efficacy of selection is low[27,28].

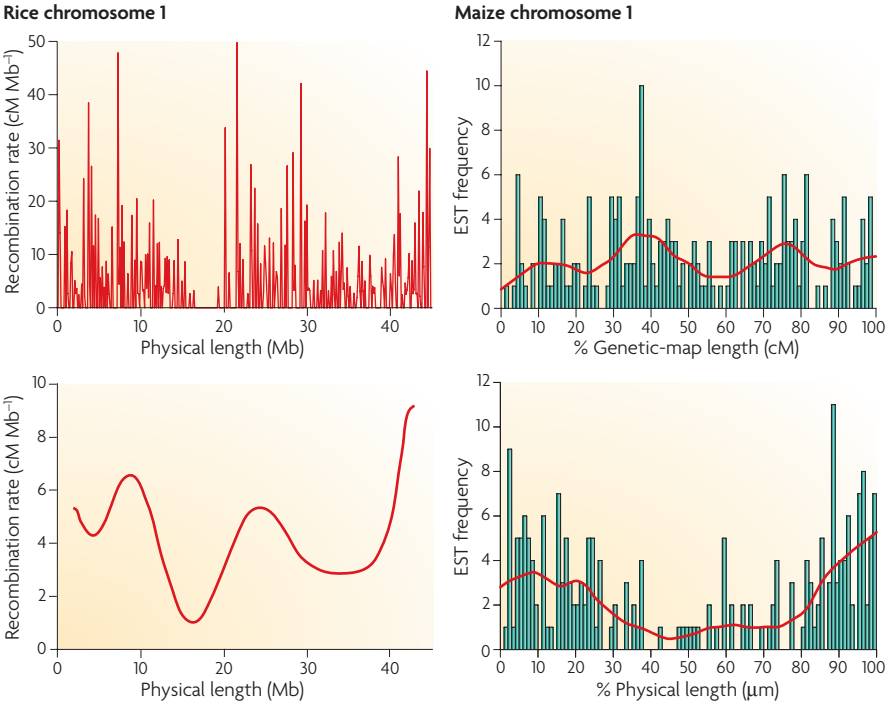
Genome-wide patterns of recombination

To understand how the effects of recombination on mutation and selection have shaped plant genome evolution, it is first essential to understand how recombination rates vary between and within genomes. Recombination is known to vary among genomic regions in both plants and animals[29,30], but genome-wide patterns of recombination are available for only a few plant species. To estimate recombination rates, both genetic (centimorgan (cM)) and physical (usually base pair) distances are required (BOX 1). Little is known about physical distances along plant chromosomes, apart from a few model species (BOX 2).

For plant species in which genome-wide recombination estimates are available, one obvious feature is the suppression of recombination in heterochromatin. The proportion of heterochromatin and the degree of suppression varies widely among species (BOX 3). For example, heterochromatin comprises 50% of the sorghum genome, with 34-fold suppression of recombination on average in heterochromatic versus euchromatic regions[31]. By contrast, pericentromeric heterochromatin comprises 75% of the tomato genome, with apparently up to 1000-fold suppression of recombination in some heterochromatic regions[32,33]. Heterochromatin is not always pericentromeric; maize (*Zea mays*) contains a substantial proportion of its genome in heterochromatic knobs, which influence recombination rates in nearby regions[34]. The mechanisms that govern heterochromatic suppression remain unclear, but might be related to epigenetic modification[35].

In euchromatin, recombination tends to be highest in distal chromosome regions, such as subtelomeres and telomeres[10,36–38] (BOX 2). However, this observation is not universal. In *Allium fistulosum*, recombination is highest in proximal regions and the recombination-rate gradient along chromosomes is reversed[39]. These broad-scale chromosomal patterns might be heterogenous on smaller scales, which could depend, in part, on a mechanistic relationship between genes and recombination events[40]. For example, recombination within the maize *bronze* gene can be up to 100 times higher than the genome average (but see BOX 3 for a discussion of caveats relating to recombination-rate estimates), and can differ by two orders of magnitude between distal and proximal regions of the gene[41]. With a few notable exceptions[36,42], the extent of fine-scale variation along plant chromosomes is not well characterized. Nonetheless, it seems that levels of

Box 1 | **Methods for estimating physical recombination rates on a genomic scale**



Genetic maps provide the frequency of meiotic crossovers between markers. To investigate genome-wide recombination rates, one also needs to know the physical distances between markers. The comparison of genetic to physical distances has been achieved by four methods.

Genome sequence

Three steps are required to compare genome sequences to genetic maps. First, the molecular markers on the genetic map must also be mapped onto the genome sequence by a computational homology search. Second, the ratio of genetic (cM) and physical (Mb) distances is calculated for every pair of contiguous markers. These ratios can be plotted along the length of the chromosome, as shown in the figure for chromosome 1 of rice (left panel). Each bar in the top left panel is the estimated cM Mb^{-1} rate between two markers on the genetic map. Third, the cM Mb^{-1} plot is smoothened, usually by fitting the data to a polynomial function (bottom left panel). This smoothening step is necessary because many marker intervals contain small Mb or cM distances, making individual cM Mb^{-1} estimates inaccurate. However, smoothening can mask the existence of recombination hotspots and coldspots, and underemphasize recombination-rate heterogeneity on fine scales. Gaps in the genome sequence also contribute to inaccuracies in cM Mb^{-1} estimates.

Physical maps

Whole-genome physical maps have been composed using BAC fingerprinting and similar methods. If molecular genetic markers have been hybridized to BACs, these maps can form the basis for estimates of physical recombination rates. At present, however, most BAC-based maps have too many gaps to be useful for whole-genome analysis of recombination rates.

Recombination nodule maps

Recombination nodules (RNs) are protein complexes that are associated with meiotic (pachytene) chromosomes, which can be observed by electron microscopy[57]. Because each RN marks a crossover site, the cumulative distribution of RNs on each chromosome produces a cytological recombination map that inherently relates genetic and physical distances. The RN map can be used to translate marker positions from a genetic map to a physical location, as shown for maize (panels on the right). The top right panel shows the distribution of mapped ESTs on the genetic map. The lower right panel shows the distribution of mapped ESTs, translated to a physical scale at 0.2 μm resolution. The curves that are superimposed on the two maize graphs show general trends of the distributions. The resolution of this approach is relatively crude, at a scale of about 1.3 Mb for each 0.2 μm of euchromatic chromosome length.

Deletion maps

Deletion maps are constructed on the basis of Southern hybridization of cloned markers with DNAs of a large number of stocks that are homozygous for terminal deletions. Because homozygosity for microscopic deletions is tolerated only in polyploids, genome-wide mapping strategies using overlapping terminal deletions is limited to polyploid plants such as wheat[77]. The maize panels are reproduced with permission from REF. 50 © (2006) Cold Spring Harbor Laboratory Press.

Box 2 | **Plant taxa with genome-wide estimates of physical recombination**

Arabidopsis thaliana
The genome sequence has been compared to genetic maps. Graphs of recombination along chromosomes have been published[10,36].

Rice
Both the genome sequence and BAC physical maps have been compared to genetic maps[37,51].

Maize
A recombination nodule map has been published for maize[78]. Molecular markers from genetic maps have been placed on chromosomes to infer genome organization[50]. Higher resolution cM Mb^{-1} estimates should be a by-product of ongoing genome-sequencing efforts.

Tomato
A recombination nodule map has been published for tomato[79], and has been used to assess diversity across the genome[80]. Additional physical maps and genome-sequence data are forthcoming.

Wheat
16,099 markers have been mapped into 156 deletion bins[77]. The average recombination rate across each bin was estimated and expressed as a function of the distance of the bin from the centromeres[59].

Other taxa
Several taxa have genetic maps and are also in the process of full genome sequencing, such as *Sorghum bicolour* (sorghum), *Arabidopsis lyrata*, *Capsella rubella*, *Populus trichocarpa* (poplar), *Mimulus guttatus*, *Medicago truncatula* and *Lotus japonicus*. Estimates of cM Mb^{-1} recombination rates across chromosomes should follow.

recombination vary not only on chromosomal scales, but also on kilobase scales.

Recombination and genome structure

Given the variation in recombination levels across plant genomes, and the effects of recombination on mutation and selection, is it possible that plant genomes are organized along recombinational gradients? If this were the case, three simple predictions should be met. First, deleterious mutations should accumulate in heterochromatin and other low-recombination regions, because natural selection is inefficient in those regions; of course, the opposite should be true in high-recombination regions. Second, rates of localized gene duplication and deletion should be elevated in high-recombination regions because of the mutagenic properties of recombination. Finally, the rate of gene-order evolution within and between species should scale with recombination rate.

Deleterious mutations in low-recombination regions. There is evidence that duplications and deletions occur in low-recombination genomic regions[43], but at present the rate of these events, the proportion that are deleterious and the rates at which they persist are unknown. The accumulation of both transposable elements (TEs) and organellar DNA in recombination-poor heterochromatic regions[44–46] indicates that they are slowly eliminated by selection. However, it is also possible that the accumulation of TEs and organellar DNA reflects biased insertion of these sequences in heterochromatin[47,48].

Currently, we do not have a complete picture of the many forces that govern the accumulation of TEs[7], but there is reason to think that TE-accumulation patterns are influenced by selection pressures that vary with recombination rate. For example, the TEs that accumulate in gene-poor, low-recombination regions in rice (*Oryza sativa*) tend to be large, whereas those that accumulate in gene-rich, high-recombination regions tend to be small (FIG. 2). Large TEs might be more likely than small ones to undergo ectopic exchange; such exchanges in gene-rich, high-recombination regions would have a high chance of gene disruption and efficient removal by natural selection. By contrast, ectopic exchange in gene-poor heterochromatin might be less likely to disrupt genome function, so that large TEs would be freer to accumulate in these regions. To better understand the interplay among insertion bias, natural selection and recombination in plants, studies that compare the distribution of *de novo* insertions with accumulated insertions, as have been carried out in *Caenorhabditis elegans*[49], will be important. Other information about rates of unequal and ectopic crossing over among TEs of different sizes will also be useful.

The converse prediction is that deleterious mutations should not accumulate in high-recombination regions, so these regions should be enriched for functional genomic elements. Consistent with this idea, gene density and recombination are positively correlated in maize, rice, wheat (*Triticum aestivum*) and *A. thaliana*[48,50–52]. Much of this correlation is attributable to a dearth of genes in heterochromatin[51,53]. Maize and tomato (*Solanum lycopersicum*), for example, have fourfold to tenfold lower gene density in heterochromatin than euchromatin[32,50]. For some species, like maize[50] and rice (FIG. 2), the correlation between recombination and gene density also holds true in euchromatic regions alone. Maize telomeric regions are an interesting exception to this trend, in that they have high recombination rates and apparently low gene density[50]. It is unclear whether these regions are exceptional because of their biological properties, or whether it is owing to the difficulty of mapping and measuring recombination rates at chromosomal ends. By contrast, fully sequenced rice telomeric regions have high gene density[54].

However, there is no detectable correlation between recombination and gene density in euchromatic regions of *A. thaliana*[48]. This observation is puzzling when one considers the potential feedback mechanism that is produced between genes and recombination. On the basis of studies that largely involve maize, plant meiotic recombination seems to to occur primarily within genes[40], perhaps because epigenetic phenomena, chromatin structure and insertion-deletion polymorphisms suppress recombination in intergenic regions[55]. If genes facilitate recombination, it follows logically that recombination should be high in regions of high gene density, as is seen in many plant species. Why is this correlation not evident in *A. thaliana* euchromatin? One possibility is that recombination in *A. thaliana* differs from maize, in that recombination might be predominantly intergenic (that is, non-genic)[36,56]. Another is that structural features of chromosomes could supersede the effects of gene density on recombination in small genomes. For example, the initiation of synapsis between homologous chromosomes during meiosis is distal in most plants, and chromosomal regions that synapse first seem to be more likely to undergo a crossover[57]. Such patterns, in combination with interference (which reduces the probability of another crossover nearby), could decrease proximal crossing over, independent of gene density. Although speculative, these factors would be more prominent in plants with short chromosomes (like those of *A. thaliana*), because interference is strongest near crossovers and decreases with distance. A final possibility is that the underlying correlation between gene density and recombination rates might be too weak to detect with inaccurate recombination-rate estimates. More conclusive evidence

for a direct relationship between gene density and recombination will come from plants with reversed patterns of recombination rates along the centromere–telomere axis, such as *Allium fistulosum*[39]. Gene density has not been measured directly in this species, but markers derived from hypomethylated regions, which are expected to correlate with active genes, do tend to cluster in high-recombination proximal regions[39].

Localized duplications in high-recombination regions. If localized duplications are caused by unequal crossing over, duplicated genes should cluster in high recombination regions. This is true for tandemly arrayed genes (TAGs), which are more abundant in high-recombination regions of *A. thaliana*[10], rice[58] (FIG. 2) and wheat[59] than in low-recombination regions, even after correction for underlying variation in gene density. TAGs in *C. elegans* are organized similarly[60], indicating that the enrichment of TAGs in high-recombination regions might be a general feature of eukaryotic genomes. However, unequal crossing over alone does not fully explain TAG distributions. If the TAG distribution were solely a property of unequal crossing over, then TAGs in high-recombination regions should have high turnover rates, owing to higher rates of both duplication and deletion. In fact, the age of TAGs is not correlated with recombination rates in *A. thaliana*[10] or rice[58]. TAG distribution is therefore likely to be a product of increased production in high-recombination regions, along with complex patterns of retention and loss.

Dispersed gene duplicates — that is, single-gene duplications that are not on the same chromosome — are also enriched in high-recombination regions[59]. In hexaploid wheat and its diploid relatives, dispersed duplication events occur at a rate of at least $\sim 2 \times 10^{-3}$ duplications at each locus every million years[61]. Duplications are countered by a tenfold higher deletion rate, indicating that most new duplicates are eventually removed from populations, except when a gene duplication confers a strong selective advantage. Both duplication and deletion occur more often in high-recombination regions in wheat and its relatives; for example, duplication rates are three times higher in distal high-recombination chromosomal regions than in proximal low-recombination regions[61]. The fact that this pattern is seen in both diploids and polyploids, coupled with the fact that purifying selection is ineffective at high ploidy levels, argues that recombination has a mechanistic role in both duplication and deletion.

Recombination and gene order. Genomic features like localized duplication and deletion — both of which are governed, and probably produced, by recombination — shape chromosomal organization and ultimately gene order (synteny). Inversions also change gene order[14], which might be caused by ectopic recombination (FIG. 1). But predictions vary about the effect of recombination on gene-order evolution. On the one hand, duplications and deletions are generated preferentially in high-recombination regions; therefore, gene rearrangement could evolve more rapidly in these regions. However, selection is also more efficient in high-recombination regions. If gene rearrangement is maladaptive, then gene order will be maintained in high-recombination regions.

There is little literature addressing this question directly, but a comparison of rice and sorghum physical maps indicates that synteny is markedly less conserved in recombination-suppressed heterochromatic regions[46]. Other observations indicate that low-recombination pericentromeric regions rearrange rapidly[14,38,62]. Superficially, this information seems inconsistent with the idea that recombination drives genome rearrangement, because low-recombination regions should have low mutation rates. It is clear that the effect of recombination on natural selection is important here. Natural selection is not only expected to be inefficient in these regions, but heterochromatic rearrangements might not be strongly deleterious because gene density is low.

In wheat and its relatives, telomeric and other high-recombination euchromatic regions rapidly evolve in gene order, owing in large part to the high incidence of duplication and deletion[61]. The net result is that

Box 3 | What do average cM Mb^{-1} estimates really mean?

With the advent of BAC genomic sequencing, investigators have expressed relative recombination rates for specific genic and intergenic sequences by comparing them with a total genomic cM Mb^{-1} average. Recombination 'hotspots' and 'coldspots' are inferred on the basis of cM Mb^{-1} rate estimates above and below the genomic average, respectively. But how useful is this comparison? Plant genomes contain a significant proportion of recombinationally suppressed heterochromatin, ranging from about 12% in *Arabidopsis thaliana*[81] to 75% in tomato[32]. Because of heterochromatin, the average recombination rate, estimated on the basis of the entire genome, is always lower than estimates that are made on the basis of euchromatin alone.

The practice of including heterochromatic portions in average cM Mb^{-1} rates has two implications. First, comparisons of average cM Mb^{-1} rates among taxa are relatively meaningless; the comparisons probably provide as much insight into the proportion of heterochromatin as genic variation in recombination rates. Second, the inferred hotspots and coldspots of recombination might not have recombination rates that are much different from the average genic rate that is measured in euchromatin. A more informative method would be to assess average recombination rates in euchromatin and use this average as the basis for comparisons among taxa and genic regions.

How can one estimate recombination rates in the euchromatic fraction? One way is to specifically stain DNA within meiotic chromosomes and measure the intensity of staining and chromosome length in heterochromatic versus euchromatic regions (for example, REF. 50). Such estimates will still be approximate, but will help to adjust for known biases.

Glossary

Fluorescence *in situ* hybridization
A technique that is used to label specific sequences on chromosomes with fluorescent molecules.

Heterochromatic knobs
Cytologically visible regions of highly condensed chromatin that are distinct from pericentromeric regions.

Microsatellite instability
A change in the number of repeats of microsatellites.

Paracentric inversion
A structural chromosome alteration that results from breakage, inversion and reinsertion of a fragment of a chromosomal arm.

Pericentric inversion
A structural alteration to a chromosome that results from breakage, inversion and reinsertion of a fragment that spans the centromere.

Polyploid
Having three or more sets of homologous chromosomes (for example, tetraploid organisms have four sets of chromosomes).

Synapsis
Formation of a synaptonemal complex between homologous chromosomes during prophase I of meiosis. Pairing is a more general term that refers to homologous associations in somatic as well as meiotic nuclei.

gene content and gene order might be best conserved in medium- to low-recombination euchromatic regions[38,63], in which mutation rates are relatively low but recombination is still adequate to allow natural selection to work efficiently. This conjecture is supported by observations made in yeast and other systems[64], indicating that this pattern of gene-order evolution might be a general feature of eukaryotic genomes. More observations, particularly those made on the basis of plant data, are needed to further investigate this theory.

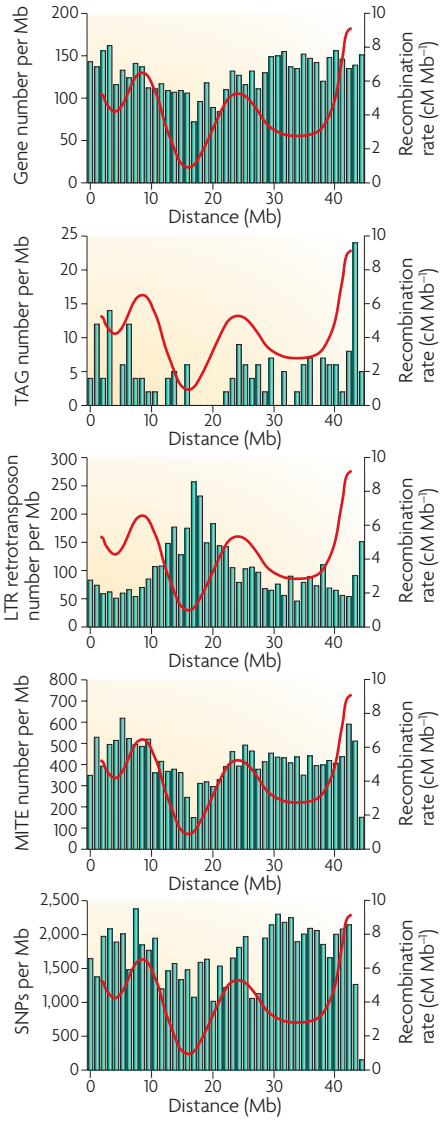


Figure 2 | **The correlation of genomic features with recombination rates.** For all panels, genomic features are binned in 100 kb windows. MITEs are small DNA transposons. Data on gene number and tandemly arrayed genes (TAGs) are from REF. 57; data on transposable elements are from REF. 51; data on SNPs are from REF. 82. The red curve shown in each panel is a smoothed cM Mb^{-1} rate curve, as explained in BOX 1.

There are too few data at this time to unequivocally conclude that the rate of gene-order evolution in plants varies within euchromatin as a function of recombination rates. However, this is an issue of practical importance. Researchers often contend that gene-order information that is gleaned from species with small, tractable genomes can be applied directly to species with larger, less-tractable genomes. The opposing view is that synteny breaks down rapidly and might not be of great practical benefit[1,65,66]. These two extremes miss the important point that the rate and pattern of gene-order evolution probably varies by species, population history and — perhaps most importantly — genomic location, as a function of recombination rates.

Plants compared with animals

The preceding arguments indicate that recombination shapes plant genome structure, but plants are not unique in this respect. Genomic features like TAGs[60] and gene order[64] are also organized along recombinational gradients in other eukaryotes. However, despite conspicuous exceptions such as amphibians[67], animal genomes are generally less labile than plant genomes. For example, as a group, mammals are older than grasses, but vary only 5- to 8-fold in genome size and chromosome number[68,69], compared with 55- and 10-fold, respectively, for grasses. Is it possible that recombination contributes to observed differences in genome variability between plants and animals?

One important difference between plants and animals is the lack of a dedicated germ line in plants. So, mitotic as well as meiotic mutations can be passed to progeny, and somatic recombination could be a potent source of genetic and genomic variation in plants. Interestingly, somatic recombination in *A. thaliana* increases with DNA damage and other stress conditions[12], implying a commensurate increase in mutation rate. Further, these environmentally induced elevations in somatic recombination rates are heritable, and can be retained for several generations[70]. The ability to detect such events using molecular genetic tools is currently limited to *A. thaliana*, but there is no obvious reason to think that this species will be unique in this respect. These observations indicate that relative amounts of genomic mutation caused by recombination could be much higher in plants than animals.

Comparisons of genome-wide recombination rates in plants and animals also indicate potential differences. Genome-wide recombination rates have been recently characterized in rats, mice and humans by comparing genomic sequences to genetic maps (BOX 1). Average recombination rates are 0.60 cM Mb^{-1}, 0.56 cM Mb^{-1} and 1.25 cM Mb^{-1} for rats, mice and humans, respectively[29]. The highest recombination-rate estimate for any region of these three genomes is 4.26 cM Mb^{-1}. A distant relative, *D. melanogaster*, has higher average rates of recombination (2.9 cM Mb^{-1}), but a similar range among genomic regions (0.0–3.98 cM Mb^{-1}) (REF. 71).

All of these rates are substantially lower, both in average value and in range, than those of *A. thaliana* and rice, the two plant genomes for which recombination rates can be estimated by similar methods (BOX 2). The *A. thaliana* and rice genomic averages are 4.80 cM Mb^{-1} and 4.12 cM Mb^{-1}, respectively[10,58]. Genomic averages are misleading (BOX 3), so it is perhaps more telling that the ranges of recombination rates are much greater in these two plant species — varying from 0.0 cM Mb^{-1} to >10 cM Mb^{-1} — than in the mammalian genomes. Although not definitive, these numbers indicate that recombination rates in plant genomes might have a greater range than those in other species. If mutation rates scale commensurately, plants could harbour large genomic regions in which mutation rates greatly exceed those of animal genomes.

Admittedly, these observations must be interpreted with caution, for three reasons. First, both *A. thaliana* and rice predominantly self-fertilize; there are theoretical reasons to indicate that selfing organisms have elevated crossover rates[72]. Indeed, hermaphroditic *C. elegans* also seems to have a high range in genomic recombination rate[73]. Second, both *A. thaliana* and rice have small genomes, such that their patterns of recombination could be atypical, as small chromosomes might harbour higher cM Mb^{-1} recombination rates. For comparison, reasonable genome-wide cM Mb^{-1} estimates are available for hexaploid wheat, which is an outcrosser with a large genome (BOX 2). Average cM Mb^{-1} rates are modest (<1.0 cM Mb^{-1}) in wheat[59], but the range is higher than that of rats, mice and humans, because distal chromosomal regions have recombination rates exceeding 6.0 cM Mb^{-1} (REF. 38). Third, all recombination-rate estimates are imprecise, are dependent on the methods used to estimate rates and are usually made on the basis of only a single genetic cross. In summary, more data are needed to draw firm conclusions, but the

high variance in recombination rates in plants that has been documented so far could signal an important role for recombination in the comparative lability of plant genomes.

Conclusions and perspectives

We argue that recombination is a strong force driving some aspects of plant genome variability. This argument is not intended to downplay the significance of paleopolyploidy, but rather to emphasize that other mechanisms also have an important role in plant genome evolution, and merit increased attention.

Our argument relies on the observations that recombination is mutagenic and that plant genomes are organized along recombinational gradients. The former is undoubtedly true, but there are relatively few estimates of the rates of specific types of mutation that are caused by recombination. Some important features of plant genomes are consistent with the second observation, including the accumulation of putatively deleterious features in heterochromatin and high duplication and deletion rates in distal chromosomal regions. However, some predictions — such as the correlation between gene density and recombination rate in euchromatin — do not hold true in all species that have been examined to date. A corollary claim — that disruption of synteny between species should be a function of recombination rates — is not yet firmly established. Nonetheless, given the limitations of current recombination-rate estimates, it is remarkable that any genomic features exhibit a discernible relationship with recombination. The fact that some patterns emerge attests to the strength of recombination as an evolutionary force.

Recombination rates seem to vary more widely among regions of plant genomes than mammalian genomes, providing a potential mechanism that might contribute to greater variability in genome size, chromosome number and genome organization among plants. This part of our argument is at best only a suggestion, formed on the basis of limited data in both plants and animals. To date, it is for only a handful of animal genomes, and only two plant genomes, that recombination-rate estimates have been derived by the comparison of sequence and genetic-map data. Even these estimates might be inaccurate, owing to poor mapping resolution and the fact that genetic maps are usually made on the basis of a single cross. As argued by Coop and Przeworski for humans[74], further studies of plant recombination should consider variation among genotypes[55], recombination patterns along whole chromosomes[36] and patterns within genes[55]. Powerful new population-genetic tools can also be used to assess both species-wide recombination rates[75] and the relative frequency of crossing over to gene conversion[76]. More careful characterization of recombination patterns and rates will help to investigate what we believe is an important mechanism in plant genome evolution.

Brandon S. Gaut is at the Department of Ecology and Evolutionary Biology, University of California Irvine, Irvine, California, USA.

Stephen I. Wright is at the Department of Biology, York University, Toronto, Ontario, Canada.

Carène Rizzon is at the Department of Statistics and Genome, University of Evry Val d'Essonne, Evry, 91034, France.

Jan Dvorak is at the Department of Plant Sciences, University of California Davis, Davis, California, USA.

Lorinda K. Anderson is at the Department of Biology, Colorado State University, Fort Collins, Colorado, USA.

Correspondence to B.S.G.
e-mail: bgaut@uci.edu

doi:10.1038/nrg1970

1. Gaut, B. S. Evolutionary dynamics of grass genomes. *New Phytol.* **154**, 15–28 (2002).
2. Bennett, M. D. & Smith, J. B. Nuclear DNA amounts in angiosperms. *Phil. Trans. R. Soc. Lond. B Biol. Sci.* **334**, 309–345 (1991).
3. Adams, K. L. & Wendel, J. F. Polyploidy and genome evolution in plants. *Curr. Opin. Plant Biol.* **8**, 135–141 (2005).
4. Cui, L. *et al*. Widespread genome duplications throughout the history of flowering plants. *Genome Res.* **16**, 738–749 (2006).
5. Blanc, G., Hokamp, K. & Wolfe, K. H. A recent polyploidy superimposed on older large-scale duplications in the *Arabidopsis* genome. *Genome Res.* **13**, 137–144 (2003).
6. Vision, T. J., Brown, D. G. & Tanksley, S. D. The origins of genomic duplications in *Arabidopsis*. *Science* **290**, 2114–2117 (2000).
7. Bennetzen, J. L., Ma, J. & Devos, K. M. Mechanisms of recent genome size variation in flowering plants. *Ann. Bot.* **95**, 127–132 (2005).
8. Schuermann, D., Molinier, J., Fritsch, O. & Hohn, B. The dual nature of homologous recombination in plants. *Trends Genet.* **21**, 172–181 (2005).
9. Jelesko, J. G., Carter, K., Thompson, W., Kinoshita, Y. & Gruissem, W. Meiotic recombination between paralogous *RBCSB* genes on sister chromatids of *Arabidopsis thaliana*. *Genetics* **166**, 947–957 (2004).
10. Zhang, L. & Gaut, B. S. Does recombination shape the distribution and evolution of tandemly arrayed genes (TAGs) in the *Arabidopsis thaliana* genome? *Genome Res.* **13**, 2533–2540 (2003).
11. Yandeau-Nelson, M. D., Xia, Y., Li, J., Neuffer, M. G. & Schnable, P. Unequal sister chromatid and homolog recombination at a tandem duplication of the *a1* locus in maize. *Genetics* **173**, 2211–2226 (2006).
12. Molinier, J., Stamm, M. E. & Hohn, B. SNM-dependent recombinational repair of oxidatively induced DNA damage in *Arabidopsis thaliana*. *EMBO Rep.* **5**, 994–999 (2004).
13. Egli, D., Hafen, E. & Schaffner, W. An efficient method to generate chromosomal rearrangements by targeted DNA double-strand breaks in *Drosophila melanogaster*. *Genome Res.* **14**, 1382–1393 (2004).
14. Lysak, M. A. *et al*. Mechanisms of chromosome number reduction in *Arabidopsis thaliana* and related *Brassicaceae* species. *Proc. Natl Acad. Sci. USA* **103**, 5224–5229 (2006).
15. Ziolkowski, P. A., Blanc, G. & Sadowski, J. Structural divergence of chromosomal segments that arose from successive duplication events in the *Arabidopsis* genome. *Nucleic Acids Res.* **31**, 1339–1350 (2003).
16. Lockton, S. & Gaut, B. S. Plant conserved non-coding sequences and paralogue evolution. *Trends Genet.* **21**, 60–65 (2005).
17. Pearson, C. E., Edamura, K. N. & Cleary, J. D. Repeat instability: mechanisms of dynamic mutations. *Nature Rev. Genet.* **6**, 729–742 (2005).
18. Lercher, M. J. & Hurst, L. D. Can mutation or fixation biases explain the allele frequency distribution of human single nucleotide polymorphisms (SNPs)? *Gene* **300**, 53–58 (2002).
19. Hellmann, I., Ebersberger, I., Ptak, S. E., Paabo, S. & Przeworski, M. A neutral explanation for the correlation of diversity with recombination rates in humans. *Am. J. Hum. Genet.* **72**, 1527–1535 (2003).
20. Rattray, A. J. & Strathern, J. N. Error-prone DNA polymerases: when making a mistake is the only way to get ahead. *Annu. Rev. Genet.* **37**, 31–66 (2003).
21. Spencer, C. C. A. *et al*. The influence of recombination on human genetic diversity. *PLoS Genet.* **2**, e148 (2006).
22. Betancourt, A. J. & Presgraves, D. C. Linkage limits the power of natural selection in *Drosophila*. *Proc. Natl Acad. Sci. USA* **99**, 13616–13620 (2002).
23. Ometto, L., Glinka, S., De Lorenzo, D. & Stephan, W. Inferring the effects of demography and selection on *Drosophila melanogaster* populations from a chromosome-wide scan of DNA variation. *Mol. Biol. Evol.* **22**, 2119–2130 (2005).
24. Tenaillon, M. I. *et al*. Patterns of diversity and recombination along chromosome 1 of maize (*Zea mays* ssp. mays L.). *Genetics* **162**, 1401–1413 (2002).
25. Wright, S. I. *et al*. Testing for effects of recombination on nucleotide diversity in natural populations of *Arabidopsis lyrata*. *Genetics* **174**, 1421–1430 (2006).
26. Fisher, R. A. *The Genetical Theory of Natural Selection* (Oxford Univ. Press, Oxford, 1930).
27. Hill, W. G. & Robertson, A. The effect of linkage on limits to artificial selection. *Genet. Res.* **8**, 269–294. (1966).
28. Felsenstein, J. The evolutionary advantage of recombination. *Genetics* **78**, 737–756 (1974).
29. Jensen-Seaman, M. I. *et al*. Comparative recombination rates in the rat, mouse, and human genomes. *Genome Res.* **14**, 528–538 (2004).
30. Anderson, L. K., Hooker, K. D. & Stack, S. M. The distribution of early recombination nodules on zygotene bivalents from plants. *Genetics* **159**, 1259–1269 (2001).
31. Kim, J. S. *et al*. Comprehensive molecular cytogenetic analysis of sorghum genome architecture: distribution of euchromatin, heterochromatin, genes and recombination in comparison to rice. *Genetics* **171**, 1963–1976 (2005).
32. Wang, Y. *et al*. euchromatin and pericentromeric heterochromatin: comparative composition in the tomato genome. *Genetics* **172**, 22529–22540 (2006).
33. Tanksley, S. D. *et al*. High density molecular linkage maps of the tomato and potato genomes. *Genetics* **132**, 1141–1160 (1992).
34. Buckler, E. S. T. *et al*. Meiotic drive of chromosomal knobs reshaped the maize genome. *Genetics* **153**, 415–426 (1999).
35. Yan, H. *et al*. Transcription and histone modifications in the recombination-free region spanning a rice centromere. *Plant Cell* **17**, 3227–3238 (2005).
36. Drouaud, J. *et al*. Variation in crossing-over rates across chromosome 4 of *Arabidopsis thaliana* reveals the presence of meiotic recombination 'hot spots'. *Genome Res.* **16**, 106–114 (2006).
37. Wu, J. Z. *et al*. Physical maps and recombination frequency of six rice chromosomes. *Plant J.* **36**, 720–730 (2003).
38. See, D. R. *et al*. Gene evolution at the ends of wheat chromosomes. *Proc. Natl Acad. Sci. USA* **103**, 4162–4167 (2006).
39. Khrustaleva, L. I., de Melo, P. E., van Heusden, A. W. & Kik, C. The integration of recombination and physical maps in a large-genome monocot using haploid genome analysis in a trihybrid allium population. *Genetics* **169**, 1673–1685 (2005).
40. Schnable, P. S., Hsia, A. P. & Nikolau, B. J. Genetic recombination in plants. *Curr. Opin. Plant Biol.* **1**, 123–129 (1998).
41. Fu, H., Zheng, Z. & Dooner, H. K. Recombination rates between adjacent genic and retrotransposon regions in maize vary by 2 orders of magnitude. *Proc. Natl Acad. Sci. USA* **99**, 1082–1087 (2002).
42. Yao, H. *et al*. Molecular characterization of meiotic recombination across the 140-kb multigenic *a1–sh2* interval of maize. *Proc. Natl Acad. Sci. USA* **99**, 6157–6162 (2002).

43. Ma, J. & Bennetzen, J. L. Recombination, rearrangement, reshuffling, and divergence in a centromeric region of rice. *Proc. Natl Acad. Sci. USA* **103**, 383–388 (2006).
44. Matsuo, M., Ito, Y., Yamauchi, R. & Obokata, J. The rice nuclear genome continuously integrates, shuffles, and eliminates the chloroplast genome to cause chloroplast-nuclear DNA flux. *Plant Cell* **17**, 665–675 (2005).
45. Lin, X. *et al*. Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*. *Nature* **402**, 761–768 (1999).
46. Bowers, J. E. *et al*. Comparative physical mapping links conservation of microsynteny to chromosome structure and recombination in grasses. *Proc. Natl Acad. Sci. USA* **102**, 13206–13211 (2005).
47. Peterson-Burch, B. D., Nettleton, D. & Voytas, D. F. Genomic neighborhoods for *Arabidopsis* retrotransposons: a role for targeted integration in the distribution of the Metaviridae. *Genome Biol.* **5**, R78 (2004).
48. Wright, S. I., Agrawal, N. & Bureau, T. E. Effects of recombination rate and gene density on transposable element distributions in *Arabidopsis thaliana*. *Genome Res.* **13**, 1897–1903 (2003).
49. Rizzon, C. *et al*. Patterns of selection against transposons inferred from the distribution of Tc1, Tc3 and Tc5 insertions in the mut-7 line of the nematode *Caenorhabditis elegans*. *Genetics* **165**, 1127–1135 (2003).
50. Anderson, L. K., Lai, A., Stack, S. M., Rizzon, C. & Gaut, B. S. Uneven distribution of expressed sequence tag loci on maize pachytene chromosomes. *Genome Res.* **16**, 115–122 (2006).
51. International Rice Genome Sequencing Project. The map-based sequence of the rice genome. *Nature* **436**, 793–800 (2005).
52. Dvorak, J., Yang, Z. L., You, F. M. & Luo, M. C. Deletion polymorphism in wheat chromosome regions with contrasting recombination rates. *Genetics* **168**, 1665–1675 (2004).
53. Initiative, A. G. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**, 796–815 (2000).
54. Mizuno, H. *et al*. Sequencing and characterization of telomere and subtelomere regions on rice chromosomes 1S, 2S, 2L, 6L, 7S, 7L and 8S. *Plant J.* **46**, 206–217 (2006).
55. Yao, H. & Schnable, P. S. *Cis*-effects on meiotic recombination across distinct *a1–sh2* intervals in a common zea genetic background. *Genetics* **170**, 1929–1944 (2005).
56. Mezard, C. Meiotic recombination hotspots in plants. *Biochem. Soc. Trans.* **34**, 531–534 (2006).
57. Anderson, L. K. & Stack, S. M. Recombination nodules in plants. *Cytogenet. Genome Res.* **109**, 198–204 (2005).
58. Rizzon, C., Ponger, L. & Gaut, B. S. Striking similarities in the genomic distribution of tandemly arrayed genes in *Arabidopsis* and rice. *PLoS Comp. Bio.* 1 Sep 2006 (doi:10.1371/journal.pcbi.0020115).
59. Akhunov, E. D. *et al*. The organization and rate of evolution of wheat genomes are correlated with recombination rates along chromosome arms. *Genome Res.* **13**, 753–763 (2003).
60. Thomas, J. H. Analysis of homologous gene clusters in *Caenorhabditis elegans* reveals striking regional cluster domains. *Genetics* **172**, 127–143 (2006).
61. Dvorak, J. & Akhunov, E. D. Tempos of gene locus deletions and duplications and their relationship to recombination rate during diploid and polyploid evolution in the aegilops-triticum alliance. *Genetics* **171**, 323–332 (2005).
62. Hall, A. E., Kettler, G. C. & Preuss, D. Dynamic evolution at pericentromeres. *Genome Research* **16**, 355–364 (2006).
63. Akhunov, E. D. *et al*. Synteny perturbations between wheat homoeologous chromosomes caused by locus duplications and deletions correlate with recombination rates. *Proc. Natl Acad. Sci. USA* **100**, 10836–10841 (2003).
64. Hurst, L. D., Pal, C. & Lercher, M. J. The evolutionary dynamics of eukaryotic gene order. *Nature Rev. Genet.* **5**, 299–310 (2004).
65. Sorrells, M. E. *et al*. Comparative DNA sequence analysis of wheat and rice genomes. *Genome Res.* **13**, 1818–1827 (2003).
66. Rice Chromosomes 11 and 12 Sequencing Consortia. The sequence of rice chromosomes 11 and 12, rich in disease resistance genes and recent gene duplications. *BMC Biol.* **3**, 20 (2005).
67. Olmo, E. Quantitative variations in the nuclear DNA and phylogenesis of the Amphibia *Caryologia*. **26**, 43–68 (1973).
68. Gregory, T. R. *Animal Genome Size Database* [onlline], < http://www.genomesize.com > (2005).
69. Murphy, W. J., Stanyon, R. & O'Brien, S. J. Evolution of mammalian genome organization inferred from comparative gene mapping. *Genome Biol.* **2**, R1–R8 (2001).
70. Molinier, J., Ries, G., Zipfel, C. & Hohn, B. Transgeneration memory of stress in plants. *Nature* **442**, 1046–1049 (2006).
71. Singh, N. D., Arndt, P. F. & Petrov, D. A. Genomic heterogeneity of background substitutional patterns in *Drosophila melanogaster*. *Genetics* **169**, 709–722 (2005).
72. Charlesworth, D. & Wright, S. I. Breeding systems and genome evolution. *Curr. Opin. Genet. Dev.* **11**, 685–690 (2001).
73. Cutter, A. D. & Payseur, B. A. Selection at linked sites in the partial selfer *Caenorhabditis elegans*. *Mol. Biol. Evol.* **20**, 665–673 (2003).
74. Coop, G. & Przeworski, M. An evolutionary view of human recombination. *Nature Rev. Genet.* 5 Dec 2006 (doi:10.1038/nrg1947).
75. Myers, S., Bottolo, L., Freeman, C., McVean, G. & Donnelly, P. A fine-scale map of recombination rates and hotspots across the human genome. *Science* **310**, 321–324 (2005).
76. Plagnol, V., Padhukasahasram, B., Wall, J. D., Marjoram, P. & Nordborg, M. relative influences of crossing over and gene conversion on the pattern of linkage disequilibrium in *Arabidopsis thaliana*. *Genetics* **172**, 2441–2448 (2006).
77. Qi, L. L. *et al*. A chromosome bin map of 16,000 expressed sequence tag loci and distribution of genes among the three genomes of polyploid wheat. *Genetics* **168**, 701–712 (2004).
78. Anderson, L. K. *et al*. High-resolution crossover maps for each bivalent of *Zea mays* using recombination nodules. *Genetics* **165**, 849–865 (2003).
79. Sherman, J. D. & Stack, S. M. Two-dimensional spreads of synaptonemal complexes from Solanaceous plants: high-resolution recombination nodule map for tomato (*Lycopersicon esculentum*). *Genetics* **141**, 683–708 (1995).
80. Stephan, W. & Langley, C. H. DNA polymorphism in lycopersicon and crossing-over per physical length. *Genetics* **150**, 1585–1593 (1998).
81. Fransz, P. F. *et al*. Integrated cytogenetic map of chromosome arm 4S of *A. thaliana*: structural organization of heterochromatic knob and centromere region. *Cell* **100**, 367–376 (2000).
82. Feltus, F. A. *et al*. An SNP resource for rice genetics and breeding based on subspecies indica and japonica genome alignments. *Genome Res.* **14**, 1812–1819 (2004).

Acknowledgements

The work on recombination and genome evolution in the Gaut, Anderson and Dvorak laboratories are funded by the US National Science Foundation. S.I.W. is funded by the Alfred P. Sloan Foundation and the Canadian National Science and Engineering Research Council.

Competing interests statement

The authors declare no competing financial interests.