Review TRENDS in Genetics Vol.21 No.12 December 2005

Chromosome evolution in eukaryotes: a multi-kingdom perspective

Avril Coghlan[1,2,*], Evan E. Eichler[3,*], Stephen G. Oliver[4,*], Andrew H. Paterson[5,*] and Lincoln Stein[6,*]

[1]Conway Institute, University College Dublin, Belfield, Dublin 4, Ireland
[2]Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, UK, CB10 1SA
[3]Box 357730, Department of Genome Sciences, University Washington School of Medicine, Seattle, WA 98145, USA
[4]Faculty of Life Sciences, The University of Manchester, Michael Smith Building, Oxford Road, Manchester, UK, M13 9PT
[5]Plant Genome Mapping Laboratory, University of Georgia, 111 Riverbend Road Rm 228, Athens, GA 30602, USA
[6]Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724, USA

In eukaryotes, chromosomal rearrangements, such as inversions, translocations and duplications, are common and range from part of a gene to hundreds of genes. Lineage-specific patterns are also seen: translocations are rare in dipteran flies, and angiosperm genomes seem prone to polyploidization. In most eukaryotes, there is a strong association between rearrangement breakpoints and repeat sequences. Current data suggest that some repeats promoted rearrangements via non-allelic homologous recombination, for others the association might not be causal but reflects the instability of particular genomic regions. Rearrangement polymorphisms in eukaryotes are correlated with phenotypic differences, so are thought to confer varying fitness in different habitats. Some seem to be under positive selection because they either trap favorable allele combinations together or alter the expression of nearby genes. There is little evidence that chromosomal rearrangements cause speciation, but they probably intensify reproductive isolation between species that have formed by another route.

Darwin's theory of natural selection [1,2], required that there must be naturally occurring variation among individuals of a species, and that this variation was stably inherited from parent to child. However, the nature of this genetic variation was a mystery until the discovery [3] that the chromosomes of *Drosophila melanogaster* and *D. simulans* differ by a large inversion in chromosome III. Then Dobzhansky reported that flies in wild populations of *D. pseudoobscura* have numerous gross rearrangements in their chromosomes that are polymorphic between strains. He realized that these rearrangements provided crucial evidence for Darwin's theory because chromosomal changes can 'supply the raw materials for evolution', thereby enabling populations to evolve rapidly as environmental conditions change [4].

Common types of 'gross' chromosomal rearrangement (usually several megabases long) can be detected at the microscope level and include deletions, inversions, duplications and translocations (Box 1). Rearrangements often occur in complex combinations of these basic types.

Advances in molecular biology enabled the detection of fine-scale changes in chromosomal structure, shifting researchers' interest from gross rearrangements to smaller changes, such as single base pair substitutions and indels. Recently, with the sequencing of many eukaryotic genomes, the pendulum has swung back, bringing a renewed appreciation of how frequently chromosomal rearrangements occur, and how rearrangement rates vary between species. The rearrangements can reshuffle genes relative to regulatory elements, delete several genes or part of a single gene or even duplicate the entire genome (i.e. polyploidization – associated with changes in organism size and fitness [5]).

Patterns of chromosomal rearrangement

Fungi

The whole-genome sequences (WGS) of ~20 species of yeasts and filamentous fungi either are, or soon will be, available. These organisms cover a wide evolutionary range: the hemiascomycetes are more evolutionarily diverged than the chordate phylum of the animal kingdom [6]. Within the fungi, the yeasts (e.g. *Saccharomyces* clade) represent excellent subjects for chromosome evolutionary studies. *Saccharomyces*, *Kluyveromyces* and *Candida* are all members of the hemiascomycete fungi that adopt, in at least part of their life cycle, a unicellular or yeast-like form. *Kluyveromyces* has a smaller genome than the other two genera and has 50% fewer chromosomes in its haploid complement than *Saccharomyces* [7]. This difference reflects the whole-genome duplication that occurred in an ancestor of *Saccharomyces* some point after it split with the lineage leading to *Kluyveromyces*. *Candida* has a genome size ~30% larger than that of *Saccharomyces* but, similar to *Kluyveromyces*, has only eight chromosomes. (Although a large-scale genome duplication event might have occurred in the evolution of *Candida*, this was a separate and distinct event to that in the *Saccharomyces* lineage [7–10]).

It is unclear whether the *Saccharomyces* duplication was the result of autopolyploidization or allopolyploidization

Corresponding author: Stein, L. (lstein@cshl.edu).
* All authors contributed equally to this article.
Available online 19 October 2005

 0168-9525/$ - see front matter doi:10.1016/j.tig.2005.09.009

Box 1. Gross rearrangements – a primer

Gross rearrangements are large-scale changes in chromosome structure between species or between individuals in a species. Large rearrangements can be seen in light microscope examinations of chromosomal spreads. Smaller rearrangements can be identified by genetic mapping or genome sequencing.

Deletions: are the loss of a piece of chromosome, and can range in size from microdeletions of a few base pairs to the deletion of a chromosome arm or even an entire chromosome.

Duplications: are the creation of an extra copy of a piece of chromosome. These are further subclassified into tandem duplications, in which the copy is adjacent to the original segment and has the same orientation, and inverted duplications in which the copy is reversed in orientation relative to the original.

Inversions: occur when a region of the chromosome is flipped 'head to tail' but its position is unchanged.

Translocations: occur when a region of the chromosome moves from one location to another. The two most common type of translocation are reciprocal translocations, when two non-homologous chromosomes exchange chunks of DNA, and transpositions, when a chromosomal region simply moves from one location to another without a reciprocal change.

Chromosomal fusions and fissions: occur when two chromosomes fuse together, forming a single chromosome, or, conversely, when a single chromosome splits, forming two new chromosomes.

Polyploidization: occurs when the entire genome is duplicated, doubling or trebling the number of chromosomes. It can be followed by an evolutionary period in which the genome is gradually reduced in size, sometimes back to its pre-polyploidization size and chromosome count, a process known as diploidization. However, the traces of ancient polyploidization are preserved in the form of multiple copies of genes. Polyploidization can occur by either autopolyploidization, in which the genome is duplicated within a single species, or allopolyploidization, in which the genome is duplicated when two closely related species cross-hybridize [7–11].

(Box 1; [8,11]). However, several forces have remodeled the genomes of *Saccharomyces* yeasts since this ancient duplication event and, in other yeast clades, extensive genome expansions have occurred by segmental rather than whole-genome duplication [10]. Smaller scale inversions, translocations and duplications, involving segments containing only a few genes, are also common in fungal genomes [9]. For further information on the organisms discussed in this article, see Table 1.

Invertebrates

The genomes of five invertebrate species have been sequenced: two nematodes (*Caenorhabditis elegans* and *C. briggsae*) and three arthropods (*D. melanogaster*, *Anopheles gambiae* and *Bombyx mori*). Most is known about gross chromosomal rearrangement in *Drosophila* because rearrangements are easily detected in the chromosomes of their giant salivary glands. The most common type of gross chromosomal rearrangement are paracentric inversions, which do not span the centromere. Paracentric inversions are common polymorphisms within the fly species: different populations of *D. melanogaster* alone harbor >500 inversion polymorphisms [12]. Most other *Drosophila* species also have many known paracentric inversion polymorphisms [13]. Paracentric inversions are also common as fixed differences between fly species [14]. Comparison of the WGS of *D. melanogaster* and *A. gambiae* has revealed that, although their chromosomal arms have retained similar gene content since they diverged, numerous gross and fine-scale paracentric inversions have occurred within each arm [15]. Thus, the same pattern of frequent paracentric inversions is seen within *Drosophila* species and between *Drosophila* and other dipterans, such as *Anopheles*.

Little was known about the evolution of nematode chromosomes until recently, when the genome of *Caenorhabditis briggsae* was sequenced and compared with that of *C. elegans* [16]. As in fly genomes, fine-scale inversions are common in these nematodes' genomes (although their inversions cannot be classified as paracentric because nematode chromosomes lack localized centromeres). Conversely, translocations of whole chromosome arms or small segments of arms are mysteriously rare in fly genomes, but have occurred frequently in the *Caenorhabditis* genomes – there has been almost one translocation for every five inversions [16].

Polyploidy is relatively rare in animals (compared with plants); however hundreds of polyploid invertebrates are known, including arthropods, nematodes, molluscs and flatworms [5]. Indeed, two diploid species that have been chosen for sequencing – the root knot nematode

Table 1. Genomes sizes and karyotypes of model organisms[a]

Species	Genome size	Number of chromosomes	Number of genes
Drosophila melanogaster [124]	180 Mb (including heterochromatin)	3 A, XY	~13 600
Anopheles gambiae [125]	278 Mb	2 A, XY	~13 600
Bombyx mori [126]	429 Mb	28 A, ZW	~18 500
Caenorhabditis elegans [127]	100 Mb	5 A, X	~19 100
Caenorhabditis briggsae [16]	104 Mb	5 A, X	~19 500
Homo sapiens [59,128]	3100 Mb	22 A, XY	~20 000–25 000
Mus musculus [56]	2500 Mb	19 A, XY	~22 000
Rattus norvegicus [60]	2750 Mb	20 A, XY	~21 000
Gallus gallus [61]	1100 Mb	38 A, ZW	~20 000–23 000
Takifugu rubripes [129]	365 Mb	22 A[b]	~31 100
Arabidopsis thaliana [37]	125 Mb	5 A	~25 500
Oryza sativa [130,131]	420–470 Mb	12 A	~32 000–55 600
Saccharomyces cerevisiae [73]	12.1 Mb	16 A	5538 (genes of ≥100 codons) [132] 5773 [133]

[a]The karyotype is given in terms of the haploid number of autosomes (A) and sex chromosomes (X,Y,Z or W)
[b]The sex chromosome has not yet been identified for *Takifugu rubripes*.

Meloidogyne hapla and the water flea *Daphnia pulex* – have both diploid and polyploid populations in the wild ([17,18]; D. Bird and J. Colbourne, personal communication).

Vertebrates

With the exception of teleost fish and some amphibian species [19,20] whole-genome duplications are not a common occurrence in the recent evolution of most vertebrate lineages. Most vertebrate genome sequence analyses, however, support at least one ancient whole-genome duplication occurring ~500 million years ago (Mya).

Instead, chromosomal rearrangements (including fissions, fusions and translocations) and the differential expansion of repetitive sequences have been the primary forces of chromosomal change among the vertebrates [21]. Several studies have revealed numerous short intrachromosomal rearrangements in mammals [22]. For example, comparisons of the human, mouse and rat genome sequences uncovered thousands of tiny 'microrearrangements', which range from part of a gene or intergenic region to several genes long. Although a small proportion of these might be artefacts stemming from errors in the draft genome assemblies of mouse and rat [22], the number of intrachromosomal rearrangements observed is consistent with studies showing that inversions are common in invertebrate animals [14,16]. Even if we disregard putative microrearrangements, mammalian genomes contain far more short syntenic blocks than we would expect to see if rearrangement breakpoints occurred at random sites throughout the genome [23]. This has led to the controversial suggestion that chromosomal breakages tend to reoccur at 'fragile sites' or 'hotspots' in mammalian chromosomes [22,24]. This hypothesis is also supported by evidence that several other types of rearrangement tend to recur at particular sites (e.g. movements of centromeres [25,26] and segmental duplications [27–29]).

Plants (angiosperms)

Polyploidization and subsequent diploidization seem to have a greater role in the evolution of angiosperms than in other eukaryotes [30]. It has long been suspected that many angiosperms have undergone polyploidization events during their evolutionary history [31], and recent findings suggest that virtually all angiosperms are ancient polyploids. Early evidence for ancient segmental duplications in *Arabidopsis* [32,33], *Oryza* [34,35] and *Sorghum* [36] has been confirmed by analysis of the nearly completed *Arabidopsis* [30,37–41] and *Oryza* [42,43] genome sequences. In *Arabidopsis*, there seem to have been three different episodes of duplication [30,44], but only 30% of the genes duplicated by the most recent event (<83 Mya) still retain a duplicate copy. Similarly, in *Oryza*, only ~21% of genes have retained duplicate copies since the polyploidization event that occurred ~70 Mya [43,45].

Polyploidization can obscure other types of rearrangement in plant genomes because the loss of different copies of a duplicated gene in different lineages will lead to incongruities in the comparative maps of plant taxa [45]. However, computer programs can reconstruct the ancestral gene order that probably existed before polyploidization, thereby enabling the identification of regions of synteny and colinearity between distantly related plants [30]. After reconstruction of the ancestral state, we see striking parallels in the gene order among diverse plants, which are peppered by many exceptions. These changes in gene order have arisen because of transpositions of single genes (generally to positions nearby on the chromosome), movements of transposable elements (which in some cases have carried unrelated sequences with them) and diploidization of ancient segmental duplications [46]. Comparisons of the sequences of BACs from genomes that have not yet been fully sequenced have provided further insights into recent plant chromosomal evolution [46–53]. For example, analysis of sequenced BACs from a pair of *Zea mays* inbred strains that diverged <1 Mya showed the loss of colinearity in 27/72 (30%) of protein-coding genes, representing a dramatic rate of genomic change [54].

Dramatic expansion in the genome sizes of plants has occurred over short evolutionary times as a result of amplification of rapidly evolving repeat families, without destroying the underlying synteny [48]. Plant repetitive DNA seems to be different from that of many other taxa, in that individual repeat families only have a life span of a few million years [55]. This is in sharp contrast with mammalian repeat families: for example, 70% of L1 elements have been retained since humans diverged from mice 75 Mya [56].

Rearrangement rates

Although difficult to estimate reliably (Box 2), the rates of genomic rearrangement appear to vary between different species and can vary significantly throughout the evolutionary history of a species.

Invertebrates

The rate of chromosomal rearrangement in invertebrates is almost twice that in vertebrates: *Drosophila* has ~0.05–0.07 breakpoints per Mb per million years (Myr) [14], compared with only ~0.03 breakpoints per Mb per Myr in rodents (including both gross and fine-scale rearrangements [22]). This difference is probably largely because of life history traits. *Drosophila* species have a larger effective population size than rodents (~10^6 versus ~10^5), and more generations per year than rodents (~ten generations per year versus ~two generations per year [57]). There are also large differences in rate among other invertebrates. *Caenorhabditis* has a surprisingly high rate of ~0.5–0.7 breakpoints per Mb per Myr [16], which could be due to an even larger effective population size and shorter generation time in *Caenorhabditis* than in *Drosophila*. However, chromosomal rearrangements can be less deleterious in nematode genomes because nematodes have holocentric chromosomes [58].

Vertebrates

The availability of WGS from several vertebrate species [56,59–61] has provided unprecedented resolution in

Box 2. Estimating the rate of chromosomal rearrangement

To estimate the rate of chromosomal rearrangement, we can divide the number of rearrangements that have occurred since two species diverged by their divergence date. However, genome size and the type of chromosomal rearrangement vary between lineages, so Ranz *et al.* [14] estimated the rate of rearrangement as the number of chromosomal breakages caused by rearrangements (rather than the number of rearrangements itself) per Mb per Myr. Each reciprocal translocation causes two breakpoints, each inversion two breakpoints and each transposition three breakpoints [121]. Using this approach, Ranz *et al.* [14] estimated that fly genomes evolve two orders of magnitude faster than those of mammals, and at least fivefold faster than the most dynamic plant genomes, in the *Arabidopsis–Brassica* clade.

The calculation of rearrangement rates is fraught with problems. The scarcity of fossils means that dates must often be estimated using the unreliable molecular clock assumption, leading to errors in divergence date estimates (discussed in detail by Graur and Martin [122]). The estimate of the number of rearrangements can also have large margins of error, sometimes because of low data quality, and sometimes because of ascertainment issues. For example, rearrangement rates estimated using genetic maps will miss rearrangements that are smaller than the spacing between markers. Conversely, partially sequenced genomes might miss large rearrangements that are larger than the average contig. When two finished genomes are compared, assembly errors can incorrectly inflate the apparent number of rearrangements.

Even when comparing two finished, high-quality genomes, it can be difficult to accurately estimate the number of rearrangements owing to inaccuracy in the methods used to detect syntenic regions. A standard approach is to make nucleotide-level alignments, and to merge contiguous alignments into larger syntenic regions, but the number of syntenic regions found depends on the alignment and 'merging' algorithms used. Furthermore, estimating the number of rearrangements from a set of syntenic regions is not trivial, because it can be impossible to distinguish a transposition from overlapping inversions [123] or a translocation from a fusion followed by a fission event. Improved methods for finding syntenic regions, and inferring rearrangements from them, can reduce these sources of error.

studies of vertebrate chromosomal evolution. For example, a four-way comparison of the human, mouse, rat and chicken genomes revealed that rearrangements in the chicken lineage have been relatively infrequent and mainly intrachromosomal in nature (Figure 1). A similarly slow rate of chromosomal evolution was observed in fish, by comparing gene order in *Tetraodon* and *Fugu* to that in human [61,62]. In contrast to chicken and fish, the rate of interchromosomal rearrangements in mammals has accelerated [61]. Rodent lineages, in particular, have undergone far more rapid evolutionary change than primates since the two lineages diverged (3.2–3.5 rearrangements per Myr in rodent versus 1.6 rearrangements per Myr in humans). Consequently, the gene order is more highly conserved between the human and chicken genomes than between humans and rodents. It is possible to reconstruct the gene order along the chromosomes of the ancestral placental mammal with considerable confidence [22,61]. For example, from sequence and comparative mapping data, we know that human chromosomes 3 plus 21; 4 plus 8; 12 plus 22a; and 12 plus 22b were originally part of larger chromosomes that later underwent fissions in several mammalian lineages.

Comparative gene maps indicate that the rate of rearrangement in vertebrates has varied by more than an order of magnitude during the past 500 Myr [61,63]. Using chicken as an outgroup, Burt *et al.* [63] distinguished three different phases in vertebrates' chromosome evolution. During the initial phase (100–300 Mya) chromosome rearrangement was slow (0.2 fixed rearrangements per Myr). This was followed by an episode of elevated chromosomal rearrangement (65–100 Mya), during which the rate increased to >1.1 rearrangements per Myr. Since the mammalian radiation, rates have varied radically in a lineage-specific fashion. In some lineages (e.g. the human branch), chromosomal rearrangement has slowed to a crawl ($\sim$0.1 events per Myr), whereas in others (e.g. the New World Monkeys and hylobatid primate branches), rates of chromosomal rearrangement have suddenly surged (1.5–2.3 events per Myr).

Plants

The extensive variation in angiosperm genome sizes makes it challenging to devise rate estimates that are directly comparable to those of taxa with smaller size variation. For example, rice and wheat differ by at least 20 major rearrangements; they last shared common ancestry $\sim$50 Mya and have genome sizes of 420 Mb and $\sim$15 000 Mb, respectively. The 40-fold difference in genome size means that the comparison of rice with wheat can be used to make two rate estimates that are 40-fold apart (0.0008 breakpoints per Mb per Myr, or 0.00002 breakpoints per Mb per Myr). This is an extreme case, but by no means the only one. For example, rate estimates for the more rapidly rearranging plant genomes, *Arabidopsis*

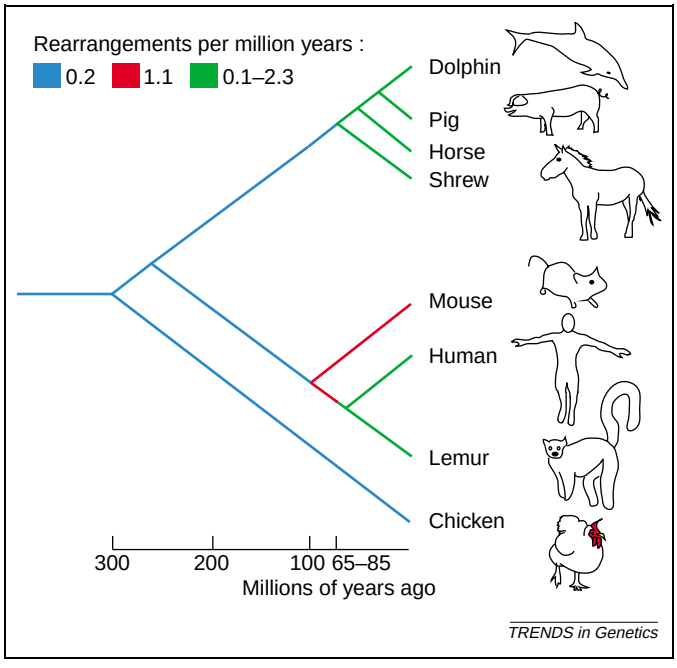

Figure 1. Using chicken as an outgroup, Burt *et al.* [63] distinguished three phases in vertebrate chromosomal evolution. During the initial phase (100–300 Mya) chromosome rearrangement was slow. This was followed by an episode of elevated rearrangement (65–100 Mya). Since the mammalian radiation, rates have varied radically in a lineage-specific fashion.

and *Brassica*, vary by approximately fivefold (0.002 – 0.009 breakpoints per Mb per Myr).

Mechanisms of chromosomal rearrangement

Fungi

Even in the relatively small yeast genomes, there are several repeated sequences: rRNA genes, tRNA genes, telomeres, telomere-associated genes and transposons (Ty elements). These repeated sequences can act as regions of homology that can promote gene duplication or genome rearrangement events. For example, recombination between the long terminal repeats (LTRs or delta sequences) of two Ty elements on the same chromosome arm will result in the excision of a chromosomal segment, which might be subsequently reinserted into another chromosome through delta×delta recombination. Mating events could then result in the duplication of the segment (a so-called chromosome homology region or CHR), with the gene order frequently being circularly permuted (see Refs [64,65]). Unequal meiotic crossing-over between Ty elements located at different sites within a chromosome arm will produce a duplication-deletion event, with the chromosome bearing the tandem duplication usually surviving [66–68].

Repeated sequences can also promote major karyotypic changes through ectopic homologous recombination [66,69]. Rearrangements in *Saccharomyces cerevisiae* are frequently associated with Ty elements or their LTRs [70–72]; a survey of interspecific chromosomal translocations in species of the *Saccharomyces* 'sensu stricto' clade [70] revealed that the breakpoints were always associated with repeated sequences, such as Tys, their LTRs or tRNA genes in the paradigmatic *S. cerevisiae* genome sequence [73]. Studies on filamentous fungi also point to a key role for transposable elements in promoting chromosomal rearrangements [74,75].

Invertebrates

Several lines of evidence suggest that repetitive elements also trigger chromosomal rearrangements in invertebrates. Cáceres *et al.* [76] found compelling evidence that repetitive elements are responsible for *Drosophila* species' rich inversion polymorphisms, when they sequenced the breakpoints of the *2j* polymorphic inversion from *D. buzzattii*. They found that chromosomes carrying the *2j* inversion contained transposon insertions at both breakpoints, and hypothesized that the inversion arose by unequal meiotic crossing-over between the two copies of the transposon. They subsequently sequenced the *2j* region in 39 different flies, and found that the inversion breakpoints are genetically unstable hotspots that are rapidly accumulating transposons and other microrearrangements [77]. Inversion polymorphisms in *D. melanogaster* and in *Anopheles* species probably originate by a similar mechanism [78,79]. There is evidence that repetitive elements also generate chromosomal rearrangements in nematodes: the breakpoints of translocations that have occurred since the divergence of *C. elegans* and *C. briggsae* are associated with repetitive elements in the *C. elegans* genome (although it is not known which species they occurred in), and most of these rearrangements occur in the repeat-rich chromosome ends [16,80].

Vertebrates

Similar to fungi and invertebrates, mammalian breakpoint regions also tend to be enriched for various classes of repeats [27,81,82]. For example, breakpoints in conserved synteny between chicken and mammalian chromosomes map to repeat-dense centromeric locations [61]. Similarly, an analysis of conserved synteny between human and mouse for human chromosome 19 revealed that ten out of fifteen breakpoint regions examined mapped to clusters of tandem gene duplications [81].

There is also a striking association between synteny breakpoints and large segmental duplications. Global analyses of human and mouse breakpoints have found that 25–53% of all breakpoints map to regions of segmental duplication within either species' genome [29,82]. This association is equally prevalent among interchromosomal and intrachromosomal rearrangements. A similar trend is seen among great-ape and human chromosomes, where sites of chromosome fusion (human chromosome 2), reciprocal translocation [gorilla chromosome t(4:19)] and pericentric inversion (human chromosomes 12, 15 and 18) are enriched for segmental duplications [83–87].

It is tempting to explain the association of large-scale rearrangements and segmental duplications by speculating that such regions promote such events by non-allelic homologous recombination. However, comparisons among the mouse, human and rat genomes do not support such a model [29]. When only rodent or primate lineage-specific breakpoints are considered, the association between rearrangement breakpoints and segmental duplications disappears. In other words, chromosomal rearrangements and segmental duplications are associated with each other between different lineages but not within the same lineage. This suggests that segmental duplications do not themselves cause chromosomal rearrangements, but that both are manifestations of the instability of particular chromosomal regions.

Plants

Inversions of entire chromosome arms are found even in closely related angiosperm taxa, such as *Lycopersicon* (tomato family), *Solanum* (potato) [88–90] and recently diverged diploids of the *Gossypium* (cotton) genus [91,92]. A variety of mechanisms have been proposed to explain gene gain, loss, inversions and tandem duplications in plants, including transposition and illegitimate recombination [46,47,50,51]. Local regions of repetitive DNA, such as centromeres, have also been implicated in many large-scale rearrangements. Indeed, the difficulty of detecting any traces of ancient duplication in the centromeric regions of both *Arabidopsis* and *Oryza* might be due to the instability of these regions.

Functional consequences of rearrangement

Are chromosomal rearrangements merely a nuisance for the genome, or do they have functional significance in the short term (e.g. by enabling a species to adapt to changing

environmental conditions) or in the long term (e.g. by facilitating speciation)?

Fungi

In fungi, chromosomal rearrangements might be expected to promote speciation by imposing post-zygotic isolation between strains (each translocation event results in a 50% drop in spore viability in hybrids of strains with wild-type and rearranged genomes). However, there is no correlation between the incidence of chromosomal translocations and the sequence-based phylogeny of *Saccharomyces* 'sensu stricto' species. Instead, it appears that species arise through another mechanism, although chromosomal translocations can intensify the reproductive isolation between species that has already formed by another route [70]. The problem with such retrospective surveys is that it is not possible to distinguish the impact on fertility of genetic differences caused by sequence divergence, from that of chromosome rearrangements. However, in the genetically malleable *S. cerevisiae*, it has proved possible to separate the contribution of these two factors, using an interventionist strategy. Delneri *et al.* [11] inserted *loxP* sites into specific regions of *S. cerevisiae* chromosomes and mediated chromosome breakage by inducing the *Cre* nuclease [93]. Next, they screened for survivors in which the same chromosomal translocations found in *S. mikatae* [70] had been produced in *S. cerevisiae*. They found, using inter-specific crosses, that the rearrangement of the *S. cerevisiae* genome so that it was collinear with that of *S. mikatae* resulted in the elevation of spore viability from < 3% to 20–30% in some zygote clones. Thus chromosomal rearrangements have a significant, although not dominant, role in the reproductive isolation between yeast species. Moreover, analysis of the engineered *S. cerevisiae* strains [94], and evidence from natural isolates [95], shows that such chromosomal rearrangements can have adaptive significance, thus accelerating the speciation process.

Invertebrates

Coluzzi *et al.* [96] observed three non-interbreeding populations of *A. gambiae* that live in the same region of Mali (and differ by chromosomal inversions) and speculated that chromosomal rearrangements might be contributing to speciation in the *A. gambiae* species complex (Figure 2). This is difficult to prove: even a highly significant coincidence in time between chromosomal rearrangements and speciation does not prove a causal relationship. However, some evidence suggests that inversions have contributed to speciation between the close relatives *D. pseudoobscura* and *D. persimilis*, because inversions are found within the genomic regions associated with hybrid sterility [97]. There are also some hints that chromosomal inversions might be contributing to speciation in the apple maggot fly (*Rhagoletis pomonella*), because genic differences between two reproductively isolated races of fly seem to be disproportionately located within inversions [98]. These observations suggest that chromosomal rearrangements reduce recombination between the genomes of incipient species, thereby enabling genetic differences to accumulate within the rearranged regions [99].

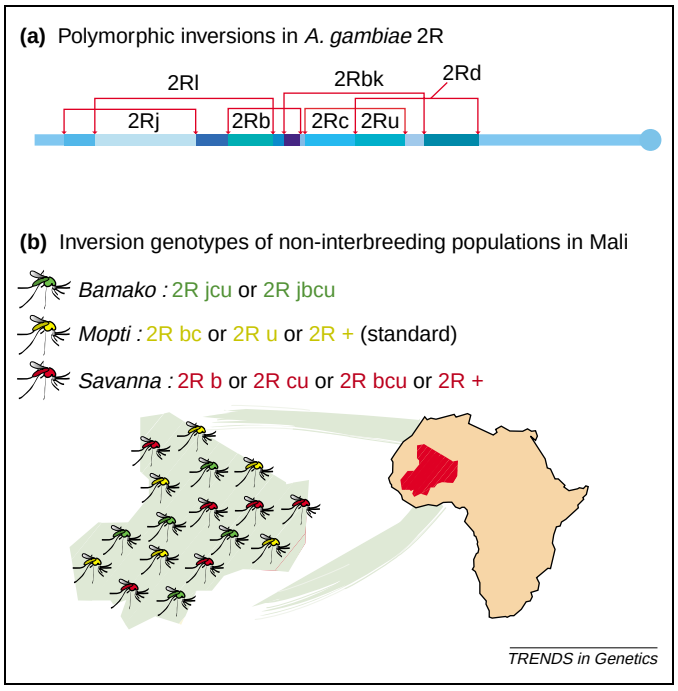

Figure 2. Do chromosomal rearrangements contribute to speciation in *Anopheles gambiae*? **(a)** The main polymorphic paracentric inversions in *A. gambiae* chromosome arm 2R [96]. All inversions are based on the standard 2R genotype, except for inversion 2Rk, which is based on a pre-existing 2Rb inversion. For example, a chromosome with arrangement 2Rjcu has inversions j, c and u. **(b)** Coluzzi *et al.* [96] observed three non-interbreeding populations of *A. gambiae* (named *Bamako*, *Savanna* and *Mopti*) that live in the same region of Mali (shown in red). The three populations differ by chromosomal inversions that might be contributing to speciation in *A. gambiae*.

Some inversion polymorphisms in *Drosophila* and *Anopheles* species are correlated with seasonal changes and altitude, and so are thought to confer varying fitness in different habitats [100,101]. Other *Drosophila* inversions are associated with variations in body size and shape [102], whereas *A. gambiae* has polymorphic inversions associated with differences in *Plasmodium* infection rates [103]. In the Australian grasshopper *Keyacris scurra*, two inversion polymorphisms located on different chromosomes are associated with differences in body size and viability (Figure 3; [13]).

At the molecular level, Schaeffer *et al.* [104] have shown that some *Drosophila* inversions are under positive selection because they keep favorable allele combinations together. However, at least one *Drosophila* inversion seems to be adaptive because it alters the expression level of a nearby gene [105].

Vertebrates

Among vertebrates, the relationship between chromosomal rearrangement and speciation has been the subject of considerable speculation [106]. The observation that closely related vertebrate species frequently differ in chromosome number or morphology (as shown by a change in the centromere position) might be viewed as weak, indirect evidence of a cause-and-effect relationship. Direct evidence, particularly experimental data, is sparse for vertebrates. Most of the advances in this area have

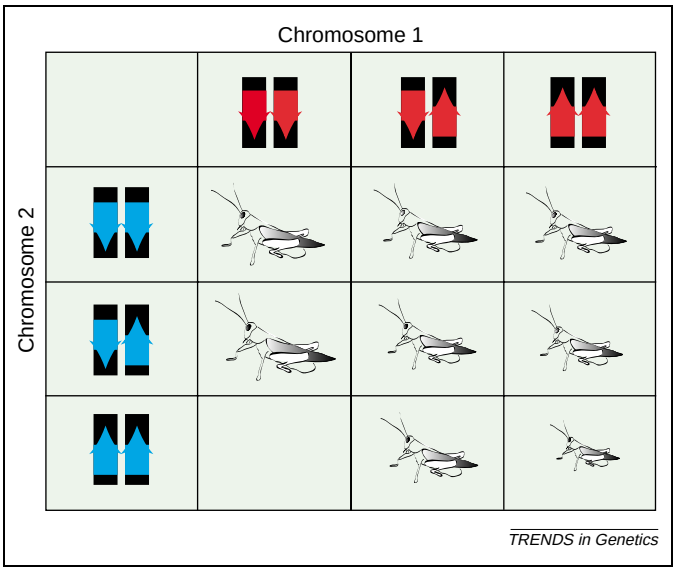

Figure 3. The Australian grasshopper *Keyacris scurra* carries two inversion polymorphisms. Individuals carrying the nine possible genotypes of these polymorphisms differ in body size and viability [13]. This is one of the most striking examples known of an inversion polymorphism with a visible phenotype.

been theoretical, with a particular emphasis on modeling the reduced evolutionary fitness of chromosomal hybrids thought to form as a result of such events. In one model by Pardo-Manuel de Villena and Sapienza [107], a bias in the segregation of rearranged chromosomes during maternal meiosis helped to explain the rapidity with which such events become fixed. Navarro and Barton [108] argued for another model (parapatric speciation), in which the suppression of recombination that results from chromosomal rearrangement serves as a genetic barrier that leads to the accumulation of post-zygotic isolation genes – genes that have alleles that are compatible within species but incompatible across species – as suggested for *Rhagoletis*. The finding of an increased number of positively selected genes on rearranged as opposed to collinear chromosomes between chimpanzee and human [109] provides some controversial [110,111] support for this new model of chromosomal speciation. Studies of natural large-scale structural variation (50 kb–1000 kb) within the human genome [112,113] suggest that rearrangements (deletions and duplications) are common in the human population. It is possible that the deleterious effects of chromosomal rearrangements and their role in speciation have been over-estimated and that such rearrangements might confer a selective advantage on evolving populations [114,115].

Plants

The exact relationship between chromosomal rearrangement and speciation remains unclear in plants; but, in at least one example, recombination between two divergent diploid species appears to have provoked a speciation event in sunflowers [116]. Although there is little direct evidence for the impact of gene or chromosomal rearrangements on phenotype in angiosperms, the occurrence (even in large genomes) of 'gene-rich' regions that lack repetitive DNA [117] might reflect selection against rearrangement events (e.g. gain of repetitive DNA) near certain genes. Alternatively, repeat-rich regions might arise because the invasion of a chromosomal region by repetitive DNA tends to lead to a cascade, as recently activated transposable elements continually jump into older transposable elements [55]. That is, the formation of non-random concentrations of repetitive DNA might leave other genomic regions 'gene-rich' by default rather than by selection.

A related observation in plants is that quantitative trait loci (QTLs) are frequently found concentrated in common genomic regions that affect different traits [118]. Computer simulations have suggested that domestication can confer a selective advantage to the *de novo* evolution of tightly linked combinations of genes or 'supergenes' [119]. Rearrangement by mechanisms such as transposition and illegitimate recombination might increase the opportunity for such supergenes to form. However, there are several plausible alternatives to the hypothesis of *de novo* evolution of tightly linked combinations of domestication-related supergenes, such as ancient evolution of QTL clusters and non-random distribution of allelic variation [118].

Concluding remarks

Despite 80 years of research, many gaps remain in our understanding of the mechanisms of both small-scale and large-scale chromosomal rearrangements and the role that they have in evolution.

Although there are many similarities among the eukaryotic kingdoms with respect to the characteristics of chromosomal rearrangements and their tendency to co-localize with repetitive DNA and duplicated regions, there are also significant differences. Polyploidization, a dominant force in the evolution of plants and fungi, occurs far less frequently in invertebrates, and is a rare event in most vertebrate lineages. Chromosomal rearrangements can vary by more than an order of magnitude both within and between kingdoms, and the rate of rearrangement within a lineage appears to be a variable that can change over time. What evolutionary pressures are responsible for these dramatic rate changes?

Over the next few years, our understanding of these phenomena will be greatly aided by sequencing deeply into the evolutionary tree. The National Human Genome Research Institute (http://www.genome.gov/10002154) and the US Department of Energy (http://www.jgi.doe.gov/sequencing) together list ~20 fungal species, ~40 invertebrates and ~25 vertebrates that are in various phases of preparation for or execution of whole-genome sequencing. In plants, sequencing is under way for the maize (http://www.jgi.doe.gov/sequencing/) and *Medicago trunculata* (legume) genomes.

This collection of sequenced genomes, together with those that are already published, will provide us with a tremendous range of variation in genome size and rearrangement rate both within and between species. For example, among the 12 species of the genus *Drosophila* scheduled for sequencing are those with numerous fixed and polymorphic inversions (e.g. *D. willistoni*), those with few (*D. simulans*) and some that have chromosomes with disparate numbers of inversions

on different chromosomes (e.g. *D. grimshawi* [13]). Analysis of these genomes might reveal why inversions proliferate in some species and often cluster along dipteran chromosomes [13,120]. The *Drosophila* sequences should also shed light on the evolution of genome size, which varies from ~150 Mb (*D. erecta*) to ~330 Mb (*D. virilis*) among the twelve species. In plants, it will be of great interest to compare the genome of rice, a 450-Mb diploid, with that of maize, an ancient tetraploid with a repeat-rich 2.2-Gb genome. Another project that is likely to help elucidate the biology of genome organization is the *Oryza* Map Alignment Project (http://www.omap.org), an effort to develop genome maps of 11 wild relatives of domestic rice. The collection of species to be sequenced is evenly split between tetraploids (e.g. *O. minuta*), and their diploid relatives (e.g. *O. punctata*), thereby enabling multiple independent polyploidization events to be studied.

Acknowledgements

The following funding sources supported this work: A.C is supported by the Wellcome Trust; E.E.E. holds a NIH grant (GM58815); A.H.P holds grants from the National Science Foundation (DBI-0115903 and DBI-0211700) and USDA-NRI (02–01412); S.G.O holds grants from the BBSRC and the Wellcome Trust; L.S holds grants from NSF (DBI-0321685), USDA (CSREES-00–52100–9622) and NIH (HG02639). A.C. thanks Richard Durbin and Des Higgins for generously allowing her to complete this work in their laboratories.

References

1 Darwin, C. (1859) *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*, John Murray
2 Powell, J.R. (1987) In the air – Theodosius Dobzhansky's 'Genetics and the Origin of Species'. *Genetics* 117, 363–366
3 Sturtevant, A.H. (1921) A case of rearrangement of genes in *Drosophila*. *Proc. Natl. Acad. Sci. U. S. A.* 7, 235–237
4 Dobzhansky, T. (1937) *Genetics and the Origin of Species*, Columbia University Press
5 Otto, S.P. and Whitton, J. (2000) Polyploid incidence and evolution. *Annu. Rev. Genet.* 34, 401–437
6 Goffeau, A. (2004) Evolutionary genomics: seeing double. *Nature* 430, 25–26
7 Kellis, M. *et al*. (2004) Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. *Nature* 428, 617–624
8 Wolfe, K.H. and Shields, D.C. (1997) Molecular evidence for an ancient duplication of the entire yeast genome. *Nature* 387, 708–713
9 Dietrich, F.S. *et al*. (2004) The *Ashbya gossypii* genome as a tool for mapping the ancient *Saccharomyces cerevisiae* genome. *Science* 304, 304–307
10 Dujon, B. *et al*. (2004) Genome evolution in yeasts. *Nature* 430, 35–44
11 Delneri, D. *et al*. (2003) Engineering evolution to study speciation in yeasts. *Nature* 422, 68–72
12 Aulard, S. *et al*. (2004) Mitotic and polytene chromosomes: comparisons between *Drosophila melanogaster* and *Drosophila simulans*. *Genetica* 120, 137–150
13 White, M.J.D. (1973) *Animal Cytology and Evolution*, Cambridge University Press
14 Ranz, J.M. *et al*. (2001) How malleable is the eukaryotic genome? Extreme rate of chromosomal rearrangement in the genus *Drosophila*. *Genome Res.* 11, 230–239
15 Zdobnov, E.M. *et al*. (2002) Comparative genome and proteome analysis of *Anopheles gambiae* and *Drosophila melanogaster*. *Science* 298, 149–159
16 Stein, L.D. *et al*. (2003) The genome sequence of *Caenorhabditis briggsae*: a platform for comparative genomics. *PLoS Biol.* 1, E45
17 Triantaphyllou, A.C. (1984) Polyploidy in meiotic parthenogenetic populations of *Meloidogyne hapla* and a mechanism of conversion to diploidy. *Révue Nematol.* 7, 65–72
18 Adamowicz, S.J. *et al*. (2002) New insights into the distribution of polyploid *Daphnia*: the Holarctic revisited and Argentina explored. *Mol. Ecol.* 11, 1209–1217
19 Becak, M.L. and Kobashi, L.S. (2004) Evolution by polyploidy and gene regulation in *Anura*. *Genet. Mol. Res.* 3, 195–212
20 Noonan, J.P. *et al*. (2004) Coelacanth genome sequence reveals the evolutionary history of vertebrate genes. *Genome Res.* 14, 2397–2405
21 Murphy, W.J. *et al*. (2004) Mammalian phylogenomics comes of age. *Trends Genet.* 20, 631–639
22 Bourque, G. *et al*. (2004) Reconstructing the genomic architecture of ancestral mammals: lessons from human, mouse, and rat genomes. *Genome Res.* 14, 507–516
23 Nadeau, J.H. and Taylor, B.A. (1984) Lengths of chromosomal segments conserved since divergence of man and mouse. *Proc. Natl. Acad. Sci. U. S. A.* 81, 814–818
24 Pevzner, P. and Tesler, G. (2003) Human and mouse genomic sequences reveal extensive breakpoint reuse in mammalian evolution. *Proc. Natl. Acad. Sci. U. S. A.* 100, 7672–7677
25 Ventura, M. *et al*. (2001) Centromere emergence in evolution. *Genome Res.* 11, 595–599
26 Eder, V. *et al*. (2003) Chromosome 6 phylogeny in primates and centromere repositioning. *Mol. Biol. Evol.* 20, 1506–1512
27 Bailey, J.A. *et al*. (2002) Recent segmental duplications in the human genome. *Science* 297, 1003–1007
28 She, X. *et al*. (2004) The structure and evolution of centromeric transition regions within the human genome. *Nature* 430, 857–864
29 Bailey, J.A. *et al*. (2004) Hotspots of mammalian chromosomal evolution. *Genome Biol.* 5, R23
30 Bowers, J.E. *et al*. (2003) Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature* 422, 433–438
31 Stebbins, G. (1966) Chromosome variation and evolution. *Science* 152, 1463–1469
32 McGrath, J.M. *et al*. (1993) Duplicate sequences with a similarity to expressed genes in the genome of *Arabidopsis thaliana*. *Theor. Appl. Genet.* 86, 880–888
33 Kowalski, S.P. *et al*. (1994) Comparative mapping of *Arabidopsis thaliana* and *Brassica oleracea* chromosomes reveals islands of conserved organization. *Genetics* 138, 499–510
34 Kishimoto, N. *et al*. (1994) Identification of the duplicated segments in rice chromosomes 1 and 5 by linkage analysis of cDNA markers of known functions. *Theor. Appl. Genet.* 88, 722–726
35 Nagamura, Y. *et al*. (1995) Conservation of duplicated segments between rice chromosomes 11 and 12. *Breed. Sci.* 45, 373–376
36 Chittenden, L.M. *et al*. (1994) A detailed RFLP map of *Sorghum bicolor* x *S. propinquum* suitable for high-density mapping suggests ancestral duplication of *Sorghum* chromosomes or chromosomal segments. *Theor. Appl. Genet.* 87, 925–933
37 Arabidopsis Genome Initiative. (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408, 796–815
38 Blanc, G. *et al*. (2000) Extensive duplication and reshuffling in the *Arabidopsis* genome. *Plant Cell* 12, 1093–1101
39 Vision, T.J. *et al*. (2000) The origins of genomic duplications in *Arabidopsis*. *Science* 290, 2114–2117
40 Paterson, A.H. *et al*. (2000) Comparative genomics of plant chromosomes. *Plant Cell* 12, 1523–1540
41 Vandepoele, K. *et al*. (2002) The automatic detection of homologous regions (ADHoRe) and its application to microcolinearity between *Arabidopsis* and rice. *Genome Res.* 12, 1792–1801
42 Vandepoele, K. *et al*. (2003) Evidence that rice and other cereals are ancient aneuploids. *Plant Cell* 15, 2192–2202
43 Paterson, A.H. *et al*. (2003) Structure and evolution of cereal genomes. *Curr. Opin. Genet. Dev.* 13, 644–650
44 Simillion, C. *et al*. (2002) The hidden duplication past of *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U. S. A.* 99, 13627–13632
45 Paterson, A.H. *et al*. (2004) Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. *Proc. Natl. Acad. Sci. U. S. A.* 101, 9903–9908

46 Song, R. *et al*. (2002) Mosaic organization of orthologous sequences in grass genomes. *Genome Res.* 12, 1549–1555
47 Bennetzen, J.L. (2000) Comparative sequence analysis of plant nuclear genomes: microcolinearity and its many exceptions. *Plant Cell* 12, 1021–1029
48 Bennetzen, J.L. (2002) Mechanisms and rates of genome expansion and contraction in flowering plants. *Genetica* 115, 29–36
49 Feuillet, C. and Keller, B. (2002) Comparative genomics in the grass family: molecular characterization of grass genome structure and evolution. *Ann. Bot. (Lond.)* 89, 3–10
50 Ramakrishna, W. *et al*. (2002) Different types and rates of genome evolution detected by comparative sequence analysis of orthologous segments from four cereal genomes. *Genetics* 162, 1389–1400
51 Brunner, S. *et al*. (2003) A large rearrangement involving genes and low-copy DNA interrupts the microcollinearity between rice and barley at the *Rph7* locus. *Genetics* 164, 673–683
52 Ilic, K. *et al*. (2003) A complex history of rearrangement in an orthologous region of the maize, sorghum, and rice genomes. *Proc. Natl. Acad. Sci. U. S. A.* 100, 12265–12270
53 Wicker, T. *et al*. (2003) Rapid genome divergence at orthologous low molecular weight glutenin loci of the A and A^m genomes of wheat. *Plant Cell* 15, 1186–1197
54 Brunner, S. *et al*. (2005) Evolution of DNA sequence nonhomologies among maize inbreds. *Plant Cell* 17, 343–360
55 SanMiguel, P. *et al*. (1996) Nested retrotransposons in the intergenic regions of the maize genome. *Science* 274, 765–768
56 Waterston, R.H. *et al*. (2002) Initial sequencing and comparative analysis of the mouse genome. *Nature* 420, 520–562
57 Eyre-Walker, A. *et al*. (2002) Quantifying the slightly deleterious mutation model of molecular evolution. *Mol. Biol. Evol.* 19, 2142–2149
58 Dernburg, A.F. (2001) Here, there, and everywhere: kinetochore function on holocentric chromosomes. *J. Cell Biol.* 153, F33–F38
59 Lander, E.S. *et al*. (2001) Initial sequencing and analysis of the human genome. *Nature* 409, 860–921
60 Gibbs, R.A. *et al*. (2004) Genome sequence of the Brown Norway rat yields insights into mammalian evolution. *Nature* 428, 493–521
61 Hillier, L.W. *et al*. (2004) Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature* 432, 695–716
62 Postlethwait, J.H. *et al*. (2000) Zebrafish comparative genomics and the origins of vertebrate chromosomes. *Genome Res.* 10, 1890–1902
63 Burt, D.W. *et al*. (1999) The dynamics of chromosome evolution in birds and mammals. *Nature* 402, 411–413
64 Warmington, J.R. *et al*. (1986) A 'hot-spot' for Ty transposition on the left arm of yeast chromosome III. *Nucleic Acids Res.* 14, 3475–3485
65 Melnick, L. and Sherman, F. (1993) The gene clusters ARC and COR on chromosomes 5 and 10, respectively, of *Saccharomyces cerevisiae* share a common ancestry. *J. Mol. Biol.* 233, 372–388
66 Kupiec, M. and Petes, T.D. (1988) Allelic and ectopic recombination between Ty elements in yeast. *Genetics* 119, 549–559
67 Wicksteed, B.L. *et al*. (1994) A physical comparison of chromosome III in six strains of *Saccharomyces cerevisiae*. *Yeast* 10, 39–57
68 Leh-Louis, V. *et al*. (2004) Expansion and contraction of the *DUP240* multigene family in *Saccharomyces cerevisiae* populations. *Genetics* 167, 1611–1619
69 Rachidi, N. *et al*. (1999) Multiple Ty-mediated chromosomal translocations lead to karyotype changes in a wine strain of *Saccharomyces cerevisiae*. *Mol. Gen. Genet.* 261, 841–850
70 Fischer, G. *et al*. (2000) Chromosomal evolution in *Saccharomyces*. *Nature* 405, 451–454
71 Dunham, M.J. *et al*. (2002) Characteristic genome rearrangements in experimental evolution of *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U. S. A.* 99, 16144–16149
72 Yoshida, J. *et al*. (2003) Positive and negative roles of homologous recombination in the maintenance of genome stability in *Saccharomyces cerevisiae*. *Genetics* 164, 31–46
73 Goffeau, A. *et al.* (1996) Life with 6000 genes. *Science* 274, 546, 563–547
74 Fierro, F. and Martin, J.F. (1999) Molecular mechanisms of chromosomal rearrangement in fungi. *Crit. Rev. Microbiol.* 25, 1–17
75 Davière, J.M. *et al*. (2001) Potential role of transposable elements in the rapid reorganization of the *Fusarium oxysporum* genome. *Fungal Genet. Biol.* 34, 177–192
76 Cáceres, M. *et al*. (1999) Generation of a widespread *Drosophila* inversion by a transposable element. *Science* 285, 415–418
77 Cáceres, M. *et al*. (2001) Molecular characterization of two natural hotspots in the *Drosophila buzzatii* genome induced by transposon insertions. *Genome Res.* 11, 1353–1364
78 Lyttle, T.W. and Haymer, D.S. (1992) The role of the transposable element *hobo* in the origin of endemic inversions in wild populations of *Drosophila melanogaster*. *Genetica* 86, 113–126
79 Mathiopoulos, K.D. *et al*. (1998) Cloning of inversion breakpoints in the *Anopheles gambiae* complex traces a transposable element at the inversion junction. *Proc. Natl. Acad. Sci. U. S. A.* 95, 12444–12449
80 Coghlan, A. and Wolfe, K.H. (2002) Fourfold faster rate of genome rearrangement in nematodes than in *Drosophila*. *Genome Res.* 12, 857–867
81 Dehal, P. *et al*. (2001) Human chromosome 19 and related regions in mouse: conservative and lineage-specific evolution. *Science* 293, 104–111
82 Armengol, L. *et al*. (2003) Enrichment of segmental duplications in regions of breaks of synteny between the human and mouse genomes suggest their involvement in evolutionary rearrangements. *Hum. Mol. Genet.* 12, 2201–2208
83 Nickerson, E. *et al*. (1999) Sequence analysis of the breakpoints of a pericentric inversion distinguishing the human and chimpanzee chromosomes 12. *Am. J. Hum. Genet.* 65, A291
84 Stankiewicz, P. *et al*. (2001) The evolutionary chromosome translocation 4;19 in *Gorilla gorilla* is associated with microduplication of the chromosome fragment syntenic to sequences surrounding the human proximal CMT1A-REP. *Genome Res.* 11, 1205–1210
85 Fan, Y. *et al*. (2002) Genomic structure and evolution of the ancestral chromosome fusion site in 2q13-2q14.1 and paralogous regions on other human chromosomes. *Genome Res.* 12, 1651–1662
86 Locke, D.P. *et al*. (2003) Refinement of a chimpanzee pericentric inversion breakpoint to a segmental duplication cluster. *Genome Biol.* 4, R50
87 Dennehey, B.K. *et al*. (2004) Inversion, duplication, and changes in gene context are associated with human chromosome 18 evolution. *Genomics* 83, 493–501
88 Bonierbale, M.W. *et al*. (1988) RFLP maps based on a common set of clones reveals modes of chromosomal evolution in potato and tomato. *Genetics* 120, 1095–1103
89 Gebhardt, C. *et al*. (1991) RFLP maps of potato and their alignment with the homeologous tomato genome. *Theor. Appl. Genet.* 83, 49–57
90 Tanksley, S.D. *et al*. (1992) High density molecular linkage maps of the tomato and potato genomes. *Genetics* 132, 1141–1160
91 Reinisch, A.J. *et al*. (1994) A detailed RFLP map of cotton, *Gossypium hirsutum*×*Gossypium barbadense*: chromosome organization and evolution in a disomic polyploid genome. *Genetics* 138, 829–847
92 Rong, J. *et al*. (2004) A 3347-locus genetic recombination map of sequence-tagged sites reveals features of genome organization, transmission and evolution of cotton (*Gossypium*). *Genetics* 166, 389–417
93 Delneri, D. *et al*. (2000) Exploring redundancy in the yeast genome: an improved strategy for use of the *cre-loxP* system. *Gene* 252, 127–135
94 Colson, I. *et al*. (2004) Effects of reciprocal chromosomal translocations on the fitness of *Saccharomyces cerevisiae*. *EMBO Rep.* 5, 392–398
95 Pérez-Ortín, J.E. *et al*. (2002) Molecular characterization of a chromosomal rearrangement involved in the adaptive evolution of yeast strains. *Genome Res.* 12, 1533–1539
96 Coluzzi, M. *et al*. (2002) A polytene chromosome analysis of the *Anopheles gambiae* species complex. *Science* 298, 1415–1418
97 Noor, M.A. *et al*. (2001) Chromosomal inversions and the reproductive isolation of species. *Proc. Natl. Acad. Sci. U. S. A.* 98, 12084–12088
98 Feder, J.L. *et al*. (2003) Evidence for inversion polymorphism related to sympatric host race formation in the apple maggot fly, *Rhagoletis pomonella*. *Genetics* 163, 939–953
99 Jiggins, C.D. and Bridle, J.M. (2004) Speciation in the apple maggot fly: a blend of vintages? *Trends Ecol. Evol.* 19, 111–114

100 Coluzzi, M. *et al*. (1979) Chromosomal differentiation and adaptation to human environments in the *Anopheles gambiae* complex. *Trans. R. Soc. Trop. Med. Hyg.* 73, 483–497
101 Nath, B.B. (2000) Dobzhansky and evolutionary cytogenetics: pioneering evolutionary studies with chromosomes. *Resonance* 5, 61–65
102 Fernández Iriarte, P.J. *et al*. (2003) Chromosomal inversions effect body size and shape in different breeding resources in *Drosophila buzzatii*. *Heredity* 91, 51–59
103 Petrarca, V. and Beier, J.C. (1992) Intraspecific chromosomal polymorphism in the *Anopheles gambiae* complex as a factor affecting malaria transmission in the Kisumu area of Kenya. *Am. J. Trop. Med. Hyg.* 46, 229–237
104 Schaeffer, S.W. *et al*. (2003) Evolutionary genomics of inversions in *Drosophila pseudoobscura*: evidence for epistasis. *Proc. Natl. Acad. Sci. U. S. A.* 100, 8319–8324
105 Puig, M. *et al*. (2004) Silencing of a gene adjacent to the breakpoint of a widespread *Drosophila* inversion by a transposon-induced antisense RNA. *Proc. Natl. Acad. Sci. U. S. A.* 101, 9013–9018
106 White, M.J.D. (1978) *Modes of speciation*, W.H. Freeman & Company
107 Pardo-Manuel de Villena, F. and Sapienza, C. (2001) Female meiosis drives karyotypic evolution in mammals. *Genetics* 159, 1179–1189
108 Navarro, A. and Barton, N.H. (2003) Accumulating postzygotic isolation genes in parapatry: a new twist on chromosomal speciation. *Evolution Int. J. Org. Evolution* 57, 447–459
109 Navarro, A. and Barton, N.H. (2003) Chromosomal speciation and molecular divergence – accelerated evolution in rearranged chromosomes. *Science* 300, 321–324
110 Bowers, E.J. (2003) Chromosomal speciation. *Science* 301, 764–765
111 Lu, J. *et al*. (2003) Comment on 'Chromosomal speciation and molecular divergence-accelerated evolution in rearranged chromosomes'. *Science* 302, 988
112 Sebat, J. *et al*. (2004) Large-scale copy number polymorphism in the human genome. *Science* 305, 525–528
113 Iafrate, A.J. *et al*. (2004) Detection of large-scale variation in the human genome. *Nat. Genet.* 36, 949–951
114 Gonzalez, E. *et al*. (2005) The influence of *CCL3L1* gene-containing segmental duplications on HIV-1/AIDS susceptibility. *Science* 307, 1434–1440
115 Stefansson, H. *et al*. (2005) A common inversion under selection in Europeans. *Nat. Genet.* 37, 129–137
116 Rieseberg, L.H. *et al*. (2003) Major ecological transitions in wild sunflowers facilitated by hybridization. *Science* 301, 1211–1216
117 Feuillet, C. and Keller, B. (1999) High gene density is conserved at syntenic loci of small and large grass genomes. *Proc. Natl. Acad. Sci. U. S. A.* 96, 8265–8270
118 Paterson, A.H. (2002) What has QTL mapping taught us about plant domestication? *New Phytol.* 154, 591–608
119 Le Thierry d'Ennequin, M. *et al*. (1999) Plant domestication: a model for studying the selection of linkage. *J. Evol. Biol.* 12, 1138–1147
120 Sharakhov, I.V. *et al*. (2002) Inversions and gene order shuffling in *Anopheles gambiae* and *A. funestus*. *Science* 298, 182–185
121 Sankoff, D. (1999) Comparative mapping and genome rearrangement. In *From Jay L. Lush to Genomics: Visions for Animal Breeding and Genetics* (Dekkers, J.C.M. *et al*., eds), pp. 124–134, CAB International
122 Graur, D. and Martin, W. (2004) Reading the entrails of chickens: molecular timescales of evolution and the illusion of precision. *Trends Genet.* 20, 80–86
123 Blanchette, M. *et al*. (1996) Parametric genome rearrangement. *Gene* 172, 11–17
124 Adams, M.D. *et al*. (2000) The genome sequence of *Drosophila melanogaster*. *Science* 287, 2185–2195
125 Holt, R.A. *et al*. (2002) The genome sequence of the malaria mosquito *Anopheles gambiae*. *Science* 298, 129–149
126 Xia, Q. *et al*. (2004) A draft sequence for the genome of the domesticated silkworm (*Bombyx mori*). *Science* 306, 1937–1940
127 The *C. elegans* Sequencing Consortium. (1998) Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* 282, 2012–2018
128 International Human Genome Sequencing Consortium. (2004) Finishing the euchromatic sequence of the human genome. *Nature* 431, 931–945
129 Aparicio, S. *et al*. (2002) Whole-genome shotgun assembly and analysis of the genome of *Fugu rubripes*. *Science* 297, 1301–1310
130 Goff, S.A. *et al*. (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science* 296, 92–100
131 Yu, J. *et al*. (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). *Science* 296, 79–92
132 Kellis, M. *et al*. (2003) Sequencing and comparison of yeast species to identify genes and regulatory elements. *Nature* 423, 241–254
133 Cliften, P. *et al*. (2003) Finding functional features in *Saccharomyces* genomes by phylogenetic footprinting. *Science* 301, 71–76