The plant genome: an evolutionary view on structure and function. Organisation of the plant genome in chromosomes

J. S. (Pat) Heslop-Harrison* and Trude Schwarzacher*

Department of Biology, University of Leicester, Leicester LE1 7RH, UK

Received 31 January 2011; revised 11 February 2011; accepted 15 February 2011.
* For correspondence (fax xxxxxx; e-mail phh4@le.ac.uk or ts32@le.ac.uk).

SUMMARY

The plant genome is organized into chromosomes that provide the structure for the genetic linkage groups and allow faithful replication, transcription and transmission of the hereditary information. Genome sizes in plants are remarkably diverse, with a 2350-fold range from 63 to 149 000 Mb, divided into $n = 2$ to $n = \approx 1200$ chromosomes. Despite this huge range, structural features of chromosomes like centromeres, telomeres and chromatin packaging are well-conserved. The smallest genomes consist of mostly coding and regulatory DNA sequences present in low copy, along with highly repeated rDNA (rRNA genes and intergenic spacers), centromeric and telomeric repetitive DNA and some transposable elements. The larger genomes have similar numbers of genes, with abundant tandemly repeated sequence motifs, and transposable elements alone represent more than half the DNA present. Chromosomes evolve by fission, fusion, duplication and insertion events, allowing evolution of chromosome size and chromosome number. A combination of sequence analysis, genetic mapping and molecular cytogenetic methods with comparative analysis, all only becoming widely available in the 21st century, is elucidating the exact nature of the chromosome evolution events at all timescales, from the base of the plant kingdom, to intraspecific or hybridization events associated with recent plant breeding. As well as being of fundamental interest, understanding and exploiting evolutionary mechanisms in plant genomes is likely to be a key to crop development for food production.

Keywords: genome, nucleus, chromosomes, cytogenetics, genome size, evolution, polyploidy, centromeres, plant breeding, heterochromatin.

THE ORGANIZATION OF THE PLANT GENOME

Plant nuclear genomes

The plant nuclear genome, consisting of the DNA and associated proteins, is organized into discrete chromosomes. Each unreplicated chromosome and metaphase chromatid consists of a single DNA molecule that is linear and unbroken from one end to the other (Figure 1). At metaphase of mitosis, the DNA is condensed into mitotic chromosomes – short, rod like bodies – while at interphase, the chromosomes are decondensed within the interphase nucleus (Figure 2). The study of the chromosome and its organization involves cytogenetics, and the field of molecular cytogenetics has developed to understand DNA sequence and the molecular structure of the chromosome and chromatin. Both the size of the plant genome and the number of chromosomes vary widely between species. In this article we will discuss the nature and consequences of these differences in an evolutionary context.

The Arabidopsis genome sequencing initiative was established partially on the basis that the genes and gene sequences found in Arabidopsis would be substantially similar to those in all other plants (Meyerowitz, 1989; Somerville, 1989). Rice, because of its nutritional importance as a crop, was the next target for genomic sequencing following an initiative to identify all genes by sequencing. The similarity of gene sequences across all plants has been found to be true, although an initial surprise was the low total number of genes (27 206 protein-coding genes in Arabidopsis, The Arabidopsis Information Resource...
Figure 1. The organization and features of a plant chromosome. Top: A fluorescent light micrograph of a metaphase chromosome stained blue with the DNA-binding fluorochrome 4',6-diamidino-2-phenylindole (DAPI). In situ hybridization shows the location of two tandemly repeated DNA sequences detected by red and green fluorescence. Centre: A diagram of the structure of a metaphase chromosome with two chromatids. Centromeres or primary constrictions are seen as gaps in cytological chromosome preparations (see Figure 3b,d). They nucleate the proteinaceous kinetochore plate to which the spindle microtubules attach and are characterised by specific centromeric histones. DNA sequences at the centromeres are not conserved between species (see text). The centromere and the regions surrounding it, called paracentromeric regions, contain large arrays of tandem repeats and are often enriched in transposable elements.

- **Euchromatin.** Lightly stained in cytological preparations (Figure 3a); generally gene rich, showing high transcriptional activity and higher levels of recombination at meiosis. Transposable elements may be dispersed widely through euchromatin.

- **Heterochromatin.** Strongly stained in cytological preparations (Figure 3a); rich in highly repeated tandemly organised DNA sequence families and sometimes transposable elements. Heterochromatin generally lacks meiotic recombination and is relatively deficient in genes, and those present often have decreased transcriptional activity.

**NOR.** Nucleolus organising regions contain the long arrays of 45S rDNA repeat units, including the 18S–5.8S–26S rRNA genes and intergenic spacers. Most genomes have several major and minor rDNA loci (Figure 4c,d). Expression of the rRNA genes generates the nucleolus at interphase (Figure 4b); when the chromosomes are spread, NORs are often visible as secondary constrictions or telomere-associated sequences (TAS). Telomeres, at the ends of chromosomes, are relatively short arrays usually of conserved simple repeat (TTAGGG)n motifs, and are species specific and often chromosomal specific.

Telomeres, the end of chromosomes, are relatively short arrays usually of conserved simple repeat (TTAGGG)n motifs. They maintain chromosome integrity by stabilizing chromosome termini.

Chromosome organization is related to genome function within the cell nucleus (Spector, 2003), with physical organization relating to regulation and gene expression, cell division, recombination and replication. There are genes involved in aspects of chromosome organization. The Gene Ontology (GO) project aims to generate descriptions of gene products in their database consisting of a controlled vocabulary of terms covering biological concepts (http://amigo.geneontology.org). It defines ‘chromosome organization’ as ‘a process that is carried out at the cellular level that results in the assembly, arrangement of constituent parts, or disassembly of chromosomes, structures composed of a very long molecule of DNA and associated proteins that carries hereditary information’. Many of these genes are related to chromatin (see Fransz and de Jong, 2011), or meiosis and recombination, rather than the structural and evolutionary aspects of chromosome organization that are discussed here.

Non-nuclear genomes and DNA sequences

Along with the nuclear genome, genes are also carried in the organelles (chloroplasts or plastids, and mitochondria) and the genomes of viruses, mycoplasmas, bacteria and fungi may be present within or in close association with plant nuclei or cells. These genomes interact and impact on the organization and evolution of the associated plant nuclear genome. Furthermore, the possible presence and effects of non-nuclear genomes (which may be transmitted to the next generation) must be considered in genomic and evolutionary studies. Increasing amounts of data obtained after the first plant genome sequences were completed have shown that transfer of genes into the plant nuclear genome, while...
Transfer of genes from both mitochondria (see Goremykin et al., 2008) and chloroplasts or other plastids (see Cullis et al., 2009) to the nucleus over evolutionary time has led to the loss of many genes from organelles (see Green, 2011). There is also evidence for transfer of genes from mitochondria to chloroplast (grape: Goremykin et al., 2009). These authors, and Bock and Timmis (2008), review the continuing nature of transfer of genes into the nucleus, with the increased regulatory ability, and the variation in genes that have been transferred in different evolutionary groups of plants. These gene transfers have led to many incongruent evolutionary trees from analysis of nuclear copies of organelar genes, where short PCR products have not distinguished the origin of the gene. Large insert (e.g. BAC) sequences can identify DNA sequences flanking the organelle-origin genes, or in situ hybridization can show their location on chromosomes rather than in organelles (e.g. Vaughan et al., 1999).

Viral genomes, particularly from the pararetroviruses with a DNA genome, can transfer from the episomal virus into the nucleus, and can be expressed as infective virus particles that cause disease. The banana streak virus was the first example to be characterized (Gayral and Iskra-Caruana, 2009; Harper et al., 1998), and petunia and tobacco vein clearing virus was identified soon after (Lockhart et al., 2000; Richert-Pöggeler et al., 2003). Solanaceous species are particularly rich in endogenous pararetroviruses (EPRVs; Hansen et al., 2005), where the majority show homology to Cavemoviruses and in some cases reach several thousand integrants. Host genome invasion by pararetroviruses has occurred several times during the evolution of Solanaceae (Staginnus and Richert-Pöggeler, 2006) and repeatedly in banana (Gayral and Iskra-Caruana, 2009). Non-functional sequences as well as complete and inducible integrants have been isolated indicating that integrated sequences decay and can be highly degenerated; they tend to be concentrated in pericentromeric heterochromatin associated with retrotransposable elements (Metaviridae sequences; Gregor et al., 2004; Hansen et al., 2005; Staginnus et al., 2007), and may play a role in host defence against virus infection through RNAi silencing (Staginnus and Richert-Pöggeler, 2006).

In the 1970s, Agrobacterium species were shown to be able to transfer genes for hormone and opine synthesis.
Figure 4. Large and small chromosomes share features of rDNA loci, centromeres and telomeres. 
(a) 4′,6-diamidino-2-phenylindole (DAPI)-stained metaphase and interphase chromosomes of Arabidopsis thaliana; enlarged inset after fluorescence in situ hybridization (FISH) with the abundant centromeric 180-bp repeat (red). (b) Interphase of Medicago truncatula after FISH with 45S rDNA probe (red). Unexpressed condensed sites of the rDNA are visible at the periphery of the nucleolus while decondensed, expressed strands run though the lighter volume of the nucleolus with less DAPI-stained DNA (micrograph from Matheus Mondin). (c) A wheat/rye recombinant line carrying a 1DL.1RS translocation as identified by genomic rye DNA hybridization (FISH) with the abundant centromeric 180-bp repeat (red). Unexpressed condensed sites of the rDNA are visible at the periphery of the nucleolus while decondensed, expressed strands run though the lighter volume of the nucleolus with less DAPI-stained DNA (micrograph from Matheus Mondin). (d) Oil palm, Elaeis guineensis (2n = 32) metaphase chromosomes after FISH with the telomeric sequence (TTTAGGG), (green) located in variable copy numbers at the end of each chromatid, and the 45S rDNA on one chromosome pair (red). Part of an interphase is visible lower right hand side (Castilho et al., 2000). Bar = 10 μm.

into the plant nuclear genome, and Schell and van Montagu showed how this property could be used in plant transformation (see, for example, Zambryski et al., 1983). Subsequently, technology to transfer genes from outside the nucleus into the genome of the host plant has been developed using the Agrobacterium or other approaches. Molecular cytogenetic analysis including fluorescent in-situ hybridization is very appropriate to locate the transgene in the genome, and even to determine copy numbers (Fransz et al., 1996; Leggett et al., 2000; Pedersen et al., 1997; Salvo-Garrido et al., 2001; Schwarzacher, 2008; Svitash and Somers, 2002; Wolters et al., 1998). Considerable efforts are required for analysis of low or single copy sequences, but these are justified as verification of nuclear integration may be difficult by Southern hybridization or PCR, particularly in slow-growing, sterile or non-intercrossable hybrids or polyploids where transmission and segregation analysis is impractical. Chromosomal analysis of transgenic lines can also establish whether the plants have maintained their chromosomal integrity or whether aneuploidy, polyploidy or rearrangements have occurred.

Composition of nuclear DNA

The nuclear DNA of plants consists of the single- or low-copy coding sequences, introns, promoters and regulatory DNA sequences, but also of various classes of repetitive DNA motifs that are present in hundreds or even thousands of copies in the genome (Heslop-Harrison and Schmidt, 1998). Repetitive DNA motifs include characteristic sequences at chromosome centromeres and telomeres (see below; Figures 1 and 4a,c), and the rDNA (rRNA genes and intergenic spacers) at the 45S and 5S loci (Figures 1 and 4b–d). Tandemly repeated or satellite DNA consists of a motif (as short as two bases, a microsatellite or simple sequence repeat, but sometimes 10 000 bp long) that is repeated in many copies at one or more genomic locations (Figure 3b,c). Satellite DNA in plants typically consists of motifs of about 180 bp and can be seen either as deep-staining heterochromatin that does not decondense during interphase (blue condensed chromatin in Figures 3a and 4b) or by in situ hybridization of the sequence after labelling (Figure 3b,c); these satellite sequences are abundant but their function in the genome is not known. Transposable elements are the third class of repetitive DNA sequences; both class I (retrotransposons) and class II (DNA transposons) elements may amplify and the elements and recognizable degraded remnants may represent half or more of the entire DNA present in the genome. Both classes of transposable elements include sequences that encode enzymes related to their own replication and integration into the nuclear DNA.

Genome size or nuclear DNA content

Each plant species has a characteristic number of base pairs in its nuclei, known as its genome size or nuclear DNA content. Work from Swift (1950) onwards has shown that the nuclear DNA content is largely constant within a species. Measuring and cataloguing the size of genomes, number of chromosomes and range of chromosome sizes and
morphism (karyotypes) has been carried out over many decades. Karyotype data have proven useful for evolutionary and phylogenetic studies at taxonomic levels between the species and family. In contrast to DNA sequence data, karyotype data often do not allow to infer higher levels of relationships. Indeed, the significance, any selective constraints, or other ‘reasons’ for differences in genome organization above the family level between species groups remain unknown.

Genome sizes are now normally estimated by using flow cytometry, replacing earlier methods of measuring absorbance of stained nuclei (microdensitometry). Nuclear genome size has been widely measured and cited in pg (picograms) of DNA, but in the context of molecular biology is now most frequently given in number of base pairs for the 1C DNA content. A nucleus immediately after meiosis but before DNA replication will have the 1C DNA content, while a replicated nucleus entering mitosis in the vegetative part of an angiosperm would have four times this amount, the 4C DNA content. Bennett and Leitch (2011) have assembled the diverse measurements of plant genome sizes into online databases (http://data.kew.org/cvalues/cvalOrigReference.html); the algae (see Bowler, 2011), pteridophyte and bryophyte data are not considered here. The databases of plant and animal genome sizes have been discussed in a broader context by Gregory et al. (2006; http://www.genomsizecom).

Published measurements of genome sizes and chromosome numbers often need critical assessment as they can be made for purposes where rigorous checking and replication is not required, may be field-based, carried out on a large scale, use techniques which are unproven or of limited reliability, or have technical errors (Greilhuber, 2005). Hence individual reports should be compared with measurements from multiple sources or observation. Reports of extreme values are particularly prone to represent errors. Casual examination of stained chromosome preparations by light microscopy – preferably of metaphase spreads, but even of stained interphase nuclei – will avoid mistakes in measurement of genome size by four-fold or more, and ensure that diploids are separated from polyploids with 50% (3n) or more (4x and above) chromosomes.

Bennett and Leitch (2011) report angiosperm genome sizes as varying from the smallest reported higher plant genome size of 63 Mb in two species of the carnivorous Genlisea, G. aurea (2n = approximately 52) and G. margaretae (2n = approximately 40), to the largest of Paris japonica (2n = 8x = 40) at 149 000 Mb, a 2350-fold range among measurements of 6288 species. For a diploid rather than polyploid species, Fritillaria platyptera (2n = 2x = 24) has the highest value at 84 150 Mb. Species with the smallest genomes of <200 Mb belong to one monocot and 13 diverse eudicot families. Many species with largest genomes are in the order Liliales (Liliaceae, Melanthiaceae and Alstroemeriaceae), with only nino eudicot families having species with genomes over 15 000 Mb. The average angiosperm genome size is 5800 Mb, with the major groups (Angiosperm Phylogeny Group III, 2009) of basal angiosperms (average 2300 Mb) and eudicots (2800 Mb) being smaller than the monocots (10 200 Mb, reduced to 8500 Mb if the order Liliales is excluded). Interestingly, gymnosperm genomes are larger with an average genome size of 18 200 Mb, and a range from 2200 to 35 200 Mb. Figure 5 illustrates this wide range of nuclear DNA contents in angiosperms.

Among eukaryotic genomes which have been sequenced, the average length of the coding sequences (excluding introns) has been reported as 1346 bp (with little variation between groups; Xu et al., 2006), while the number of genes in diploid higher plants has been found to be about 30 000 (see Ming et al., 2008), accounting for a total of 40 Mb of DNA. With the requirement for structural regions of chromosomes (centromeres and telomeres), rRNA, regulatory sequences and introns, this suggests 60 Mb is close to the minimum genome size. Lysak et al. (2009) studied genome size evolution in the Brassicaceae (showing a 16-fold range in 185 taxa studied) in the context of the phylogenetic relationships within the family. They concluded that half the species had a decreased genome size compared with the
common ancestor, despite the occurrence of dynamic genomic processes (transposition of transposable elements and polyploidization) that can increase genome size; the mechanisms to eliminate amplified DNA remain to be elucidated. Knowledge of genome size is important for choice of strategies for genomic projects including library construction, cloning, and genome sequencing. In general terms the collection of this data has not revealed general principles related to consequences of variation in genome size, nor suggested constraints, nor the mechanisms or selection pressures that modulate genome size over evolutionary time.

Greilhuber (2005) remarked that the occurrence and extent of genome size variation below the species level is controversial, pointing out faults in a number of studies reporting differences. Nevertheless, unless speciation is driven by genome size changes, differences between species show that intraspecific differences in DNA content are present and have consequences for chromosome behaviour including meiotic pairing. Chromosomal polymorphisms caused by differences in repetitive DNA sequences can occur rapidly. In maize, there are differences in the sizes of terminal heterochromatic knobs, consisting of repetitive DNA sequences (Aguiar-Perecin and de Vosa, 1985; Laurie and Bennett, 1985). The extensive variation in heterochromatin contents in rye – seen as chromosome polymorphisms even within the two homologues (see Figure 1b) – also gives differences in nuclear DNA content (Alkhimova et al., 2004). Under some conditions, repetitive sequences at the terminal regions of chromosomes are lost during mitotic divisions. Özkan et al. (2010) have shown limited variation in genome size in wheat, with substantial interspecific variation, due to the activity of retroelements.

Copy number variations (CNV) have been demonstrated to arise in the rRNA arrays of flax given different treatments by Cullis (2005). CNVs involving chromosome segments more than 1 kb in size with insertions, deletions and duplications, have been found across all chromosome arms in maize (Beló et al., 2010). Such polymorphisms in the genome, in plants like animals, are likely to have important consequences for populations and their adaptation (Biemont, 2008), disease response and heterosis (Beló et al., 2010).

Chromosome number

Every species has a characteristic number of chromosomes in the nucleus. Numbers vary extensively between species, and examples of both increases and decreases during evolution and speciation are frequent. Within the eudicots, the lowest and highest chromosome numbers, \(2n = 4\) and \(2n = \text{around 640}\) have both been reported in the single genus *Sedum* (Crassulaceae; in a flora by ’t Hart and Bleij, 2003; source and reliability unknown), although few species have more than 200 chromosomes. Several other eudicots and monocots have \(2n = 4\), while \(2n \approx 596\) has been reported in the monocot palm *Voanioala gerardii* and \(2n \approx \text{around 1200}\) in the fern *Ophioglossum reticulatum*. In genetic mapping and DNA sequencing projects, chromosome number is critical to know as it defines the number of independent linkage groups.

There are a few exceptions to the constancy of chromosome number within a species where species include several cytotypes, like members with different ploidy levels. For example, individuals of *Hordeum murinum* may be diploid \((2n = 2x = 14)\) or tetraploid \((2n = 4x = 28)\) plants (Taketa et al., 1999); there are even a few tetraploid populations of *Arabidopsis thaliana* \((2n = 4x = 20)\; \text{Heslop-Harrison and Maluszynska, 1994; Steinitz-Sears, 1963.} \) Another source of variation in chromosome number (and genome size) is the presence of supernumerary or B chromosomes (review: Jones et al., 2008) in addition to the normal chromosome complement. These usually small chromosomes are derived from the standard chromosomes in the complement, and apparently lack genes although there is a ‘drive’ process which ensures their survival and indeed amplification in number within some plants despite having detectable and often negative effects on the phenotype.

In contrast to the wide chromosome number range seen among the angiosperms, gymnosperms (characterized by large genomes) have no species with extreme chromosome numbers (typically \(2n = 2x = 14–28\)), and there are very few polyploid species in the group. Chromosome number can be stable across families: of the 232 species in 11 genera in the Pinaceae, all those studied have \(2n = 2x = 24\) chromosomes except for Douglas fir (*Pseudotsuga menziesii*, \(2n = 26\); Krutovsky et al., 2004). The 400–500 species of grasses (Poaceae) in the subtribe Triticeae, including barley, rye, wheat and a number of forage grasses (Barkworth, 2010), all have a basic chromosome number of \(x = 7\) (Figure 3a,c), although many are polyploids (Figure 4d; see below). In contrast, the *Brassica* genus has a wide range in chromosome number, and the changes, discussed below, may be driving speciation.

Chromosome size

Average chromosome size for a species is derived from chromosome number and genome size. Based on Bennett and Leitch (2011), taking unreplicated haploid genome sizes (1C) for angiosperms and dividing by haploid number \((n)\) of chromosomes reveals that 18 of the 5163 species have chromosomal DNA molecules (as would, for example, be analysed by pulse field gel electrophoresis, PFGE) \(< 10\) Mb in average size, while 118 species have an average size of more than \(3000\) Mb. The double-stranded DNA molecule in each chromatin of a metaphase chromosome of *Genlisea aurea*, averaging \(2.4\) Mb, is only half the size of the 4.6 Mb genome of the bacteria *Escherichia coli*. In species with small chromosomes, stained bacteria (where the genome may be replicated several times) can be confused with...
chromosomes in microscope preparations. Figure 4 shows A. thaliana chromosomes averaging 30 Mb in size together with wheat chromosomes averaging 800 Mb and oil palm (Elaeis guineensis) chromosomes of 114 Mb.

Despite the stability of chromosome number in the Pinaeae (2n = 24), genome size varies over a three-fold range up to 35 000 Mb, and in the Triticaceae, the haploid, x, genome size varies from about 3300 to more than 8000 Mb. Like genome size and chromosome number, these differences in average chromosome size, and the nature of the differences involving amplification or DNA and RNA transposable elements, tandemly repeated DNA sequences, and perhaps segmental duplications of the genome, can be described accurately from several complementary methods. Detailed sequence analysis (e.g. International Brachypodium Initiative, 2010) indicates that footprints of centromeric repeats and peaks in retroelement frequency are seen at the junctions of ancestral chromosome insertions. Both single-generation chromosomal changes and long-term accumulation of repetitive DNA have evolutionary roles in reproductive isolation and restriction of gene flow between newly evolving species, with consequences for understanding genome and gene evolution, as well as for the population biology, acquisition, loss or modification of gene function, and allele diversity.

As chromosomes within a species can be of different sizes, they can be sorted using flow-cytometry based on their fluorescence. In bread wheat, the first DNA library was made by Wang et al. (1992) from wheat chromosome 4A. A flow sorted BAC library of chromosome 1B was made by Janda et al. (2006), and many other chromosomes have been sorted and characterized (Paux et al., 2006; Šafář et al., 2004). The International Wheat Genome Sequence Consortium (IWGSC – http://www.wheatgenome.org) is using these flow sorted chromosomes to partition the wheat genome before chromosome-by-chromosome sequencing of the 17 000 Mb genome.

CHROMOSOMAL AND KARYOTYPE EVOLUTION

Chromosome evolution and structural variation

Chromosomes evolve by fission and fusion (leading to a change in chromosome number, or to inversions of segments within one chromosome; Jones, 1998), events that may be accompanied by duplication and inversions of chromosome arms. As an example, the chromosomes of the native European orchid Cephalanthera (see Figure 3d) with species having 2n = 32, 36 or 44, are thought to have evolved by palaetetraploidy from x = 9 followed by centric (Robertsonian) fusions leaving interstitial telomeres (Moscone et al., 2007).

With genomic data involving both genetic mapping and genome sequencing, it is now possible to identify the large scale chromosomal rearrangements that have occurred during evolution. Chromosome numbers in the Brassicaceae vary from 2n = 8 to 2n = 256 (Lysak et al., 2006). A. thaliana, with 2n = 10 (Figure 4a), has one of the smallest chromosome numbers, an advanced character representing reduction from its ancestors in the clade including A. lyrata and Capsella rubella (both 2n = 16). An impressive use of comparative chromosome painting to meiotic chromosomes using groups of BAC probes to identify each chromosome segment allowed Lysak et al. (2006) to show the origin of each chromosome in A. thaliana relative to the ancestral n = 8 karyotype, involving four chromosomal inversions, two translocations and three chromosome fusion events. In Brassica, Mandakova and Lysak (2008) used multiple selected BACs as probes to reveal the monophyletic origin of the x = 7 tribes, some of which included a translocation where chromosomal segments are exchanged between two chromosomes. The results also suggest that structure of the ancestral karyotype of the Brassica, with a reduction in chromosome number from n = 8 to n = 7 has happened more than once, with different fusion and intrachromosomal inversion events. Xiong and Pires (2011) have developed an in situ chromosome painting method to identify all chromosomes in Brassica napus and its diploid progenitors, showing a chromosomal translocation in one B. napus cultivar. They suggest that this approach will be useful to understand chromosome reorganization, genome evolution and recombination; sequence analysis would not be appropriate for the detection of single translocation breakpoints.

While some of the chromosome number changes occur through doubling of chromosome numbers or polyploidy (see below), many involve fusion or fission of chromosomes, as shown in the Brassicaceae, grasses and many other families. Through sequence comparisons, multiple orthologous gene sequences are found to show a conserved order (synteny) along chromosomes over large taxonomic distances. Data of this nature are accumulating rapidly, and syntenic comparisons are now an essential part of most genome sequence papers. For example, Jaillon et al. (2007) compared Vitis (grape vine) genomic regions to their orthologues in Populus trichocarpa, A. thaliana and Oryza sativa, a taxonomic range where direct comparisons were hardly conceivable before sequence-based comparisons became possible. In Vitis, their analysis showed that the genome has been triplicated during its early evolution, before the split of the poplar/Arabidopsis/Vitis lineages, but after the monocot/eudicot split as it was not shared with rice. The analysis identified an additional duplication in the poplar lineage, and two whole genome duplication events in the Arabidopsis lineage, as well as global duplications in the rice lineage. In the grass Brachypodium distachyon (2n = 10), sequencing of the 272 Mb genome (International Brachypodium Initiative, 2010) revealed a complex evolutionary history with six major interchromosomal
duplications within the genome, the five *Brachypodium* chromosomes originating from a five-chromosome ancestral genome through a 12-chromosome intermediate involving seven major chromosome fusions. Sets of collinear genes along all ten *Brachypodium* chromosome arms can be identified easily in the other grasses where detailed genetic maps are available (rice, barley, wheat, sorghum, and *Aegilops tauschii*). Twelve separate syntenic blocks of orthologous genes from *Brachypodium* are present in rice, sorghum and barley, with nested insertions of some *Brachypodium* ancestral groups into centromeres of the other species. In the Triticeae, a detailed analysis of syntenic regions by Luo et al. (2009) has shown how the basic number of \( x = 7 \) has been derived from \( x = 12 \) in the ancestral species (represented by rice and sorghum) not through end-to-end chromosome fusions, or translocations and loss of microchromosomes, but by the insertion of four whole chromosomes into breaks in the centromeric region of at least four other chromosomes, along with another fusion and translocation event.

Analysis of the nature of the rearrangements using whole genome sequence comparisons is enabling the history of genome evolution to be reconstructed with unprecedented accuracy. For plant breeders, knowledge of the nature of the changes shows the types of changes which might be introduced in the future, and suggests strategies and candidate accessions for crossing programmes. Parallel work across the mammals (Nagarajan et al., 2008) is also showing the evolutionary chromosome rearrangements across diverse species. Similar chromosomal fusion, fission and elimination events to those discussed in *Brassica* have been reported in cattle and the Artiodactyla (Chaves et al., 2003). In mammals, *in situ* hybridization and chromosome painting is widely used (Froenicke et al., 2006). Despite some successes (Mandakova and Lysak, 2008), this technique has been less used in plants, presumably because of the more rapid homogenization of DNA sequences from retrotransposons, so probes from large amounts of DNA become genome-specific rather than chromosome-specific. Recent advances in large-insert (BAC or fosmid) hybridization suggest it will be increasingly used to address chromosome evolution (Lysak et al., 2006) and physical linkage mapping of sequences (Han et al., 2011).

**Aneuploidy – chromosome loss or gain**

Aberrant cell division is relatively frequent, and chromosomes are lost or gained during mitosis or meiosis leading to aneuploidy. Figure 2, an intergeneric hybrid, shows nuclei at all phases of the cell cycle, but includes some cells with micronuclei (arrows) from mis-divisions. In many cases, these cells will not divide further, but the mis-division can occur in gametes or cells which regenerate to a whole organism. In mammals, most such aneuploids do not develop. Many plant aneuploids grow to generate adult plants, not least because plant genomes are often polyploid (see below) and have higher plasticity and mechanisms for gene dosage compensation. Chromosome addition lines, with an extra copy of a chromosome, occur naturally (first found in *Datura* by Blakeslee and Avery, 1919). They are also made by crossing tetraploid and diploid plants, or crossing different species, followed by backcrossing to derive lines with one or a few extra chromosomes. These hybrids have proved valuable to transfer alien chromosomes from wild relatives to crop species; recombination between the alien and crop chromosome can then reduce the chromosome number while still transferring the required characters. Particularly in wheat, such lines (Figure 4d) have a long history of use in breeding programmes (see, e.g. Bardsley et al., 1999; Heslop-Harrison et al., 1991), and a number of programmes are exploiting the transfer of important disease resistance genes into wheat (Ayala-Navarrete et al., 2007; Graybosch et al., 2009; Molnár et al., 2011; Sepsí et al., 2008).

Monosomic plants are regularly found in species with a recognizable polyploid ancestry and are missing one (of a pair) of chromosomes. These have proved extremely valuable for genetic analysis, as the phenotype of the plant reflects modified expression of the genes carried by that monosomic chromosome; substantial amounts of genetic analysis in wheat and in maize have involved monosomic analysis (Helentjaris et al., 1986). Trisomic lines, with an additional single chromosome, are also valuable for genetic analysis of diploid species to assign linkage groups to chromosomes (rice: McCouch et al., 1988).

**Polyploidy**

Whole genome duplication or polyploidy has probably played a major role in the evolution of all angiosperms by enabling fertile interspecific hybrids to be generated with multiple gene alleles at each locus, through freeing duplicated genes to mutate, and through reproductive isolation of new polyploids leading to speciation with limited gene flow (see, for example, Proost et al., 2011; Soltis and Burleigh, 2009). Polyploidy can arise by multiplication of the genome in one plant – autoploidy – or through hybridization of two species with doubling of the chromosomes of one or more of the species involved – allopolyploidy. Autopolyploids may be recognized as a different species from their diploid progenitor, or may be placed in the same taxon, despite usually having some morphological differences including size and pollen morphology, and being reproductively isolated.

Cytological evidence for polyploidy includes the occurrence of an incrementing series of chromosome numbers within a species group (e.g., *Cephalanthera*; Moscone et al., 2007), the behaviour of hybrids with chromosome pairing at meiosis, and the existence of monosomic plants. In the 1990s, this evidence suggested that perhaps 30% of plants...
were polyploid, although some questioned whether species such as maize were polyploids or palaeopolyploids. However, with DNA sequence and genetic map data showing the presence of copies of multiple genes in the same order on two or more chromosomes, evidence for whole genome duplications or polyploidy in the ancestry of species becomes unequivocal (Tang et al., 2010). Schnable et al. (2009) show that every chromosome arm in maize carries blocks of genes duplicated in order on another chromosome, and the results clearly show chromosomes involved in translocations. It is now obvious that ‘diploid’ Brassica species including B. oleracea and B. rapa are ancient hexaploids (Lagercrantz and Lydiate, 1996), with three different genomes. The analysis of sequence data in combination with physical and genetic mapping shows the complex nature of the collinear genome segments, translocations and inversions (Trick et al., 2009) and the amplification of repetitive elements after separation of the ancestral species (Alix et al., 2008).

Many of the polyploid events, recent and ancient, have involved autoploidy or hybridization of species which are evolutionarily close. For these plants to be fertile, meiotic chromosome pairing must lead to regular formation of bivalents, rather than multivalents involving more than one homologous pair of chromosomes where recombination and segregation would lead to unbalanced gametes. In wheat, Riley and Chapman (1958) described the effect of a single locus, Pairing homoeologous (Ph), which ensures strict bivalent formation, showing that homology search mechanisms are under genetic control. We can speculate that the widespread and early occurrence of polyploidy in the angiosperm lineage is due to the group’s unique ability to achieve strict bivalent pairing at meiosis, which could be a consequence of very sensitive homology matching (Schwarzacher, 1997). Evidence suggests mediation by cyclin-dependent kinase-like genes (reviewed in Yousaizafi et al., 2010).

Recent work by Fawcett et al. (2009) and associated commentary by Soltis and Burleigh (2009) has dated whole genome duplication events across 13 diverse angiosperm families to the Cretaceous–Tertiary (K–T) boundary when 60% of plant species went extinct; Fawcett et al. (2009) speculate that the new polyploids had a substantial evolutionary advantage over their diploid ancestors (Proost et al., 2011). It will be interesting to see if more recent events are found, or whether polyploidy is ultimately an evolutionary dead-end except following catastrophic climate change. Interestingly, the K–T adaptation through polyploidy seems to be restricted to the angiosperms. The pteridophytes include polyploids and many high chromosome numbers that potentially represent higher ploidy, but the K–T extinction event marked the extinction of the fern forests; in contrast, the gymnosperms survived and remain a very successful group although they include few polyploids (except in the genus Ephedra). There are not enough sequence data from these large genomes to identify older polyploids, although the similar and low chromosome number in most gymnosperms provides weak evidence against whole genome duplication.

Chromosome changes and speciation

Occasional chromosomal mutations can become fixed in a population, thus establishing reproductive barriers and leading to the emergence of new species. The diverged species may later form hybrids, often in a limited geographic area, a hybrid or tension zone, where otherwise selectively disadvantaged hybrids with reduced fitness survive in an environment not optimal for either of the parental species (Hewitt, 1988). Analysing the gene flow and differential introgression of genomes in such hybrid zones allows identifying genomic regions involved in speciation (Payseur, 2010). Furthermore, the seemingly random changes found in chromosomal sets of individuals are often of a similar nature to those found between species. They can be seen as the first step in speciation through chromosome evolution.

THE STRUCTURE OF THE CHROMOSOME

Chromosome packaging

The packaging of the double-stranded DNA helix into the nucleosomes is similar in all organisms (Richmond et al., 2002); coiling into the next level of fibre is discussed by Fransz and de Jong (2011). Neither the detailed nature nor the consequences of packaging of the DNA fibres into the chromosome at higher levels are clear. Many biology textbooks include diagrams with a hierarchy of coiled-coils, but evidence for this is weak and inconsistent. There are technical reasons why investigation has been difficult, including the fact that the DNA is in a hydrated matrix with salts and proteins which is rapidly disturbed by fixation protocols, while the structures are too polymorphic to be understood by crystallography. However, study of higher levels chromatin packaging, its genetic control and the access by replication, transcription and condensation proteins will lead to better understanding of normal and abnormal nuclear development and the genetic and epigenetic regulation processes.

Morphological features of chromosomes

In most species, chromosomes have three structural features that have been identified since the earliest microscopy work: the telomeres at the ends of each chromosome, the centromere or primary constriction and, on some chromosomes, a secondary constriction at the nucleolar organizing region (NOR) (Figure 1). Using conventional DNA stain Feulgen these features are particularly well distinguishable (Figure 3d). Chromosome shape is defined by the position of the centromere along its length: it can be at one end of the chromosome (a telocentric chromosome), close to the end
(acrocentric), near the middle (metacentric), or somewhere between the physical middle and the end (submetacentric).
The description of the chromosome sizes, usually given as measurements of physical length made in a microscope, and the position of the centromeres, gives the karyotype of a species. Karyotypes can include a set of very similar sized chromosomes such as seen in rye and wheat (Figures 3a,c and 4d), but bimodal karyotypes with several large and a number of smaller chromosomes (Figure 3d) are frequently seen.

Telomeres
The Nobel prize-winning work of Blackburn and Szostak discovered that a unique DNA sequence in the telomeres protects the chromosomes from degradation in many species, and confirmed that indeed each chromosome was a single, double-stranded DNA molecule. In work with A. thaliana, Richards and Ausubel (1988) showed that chromosomes ended with the repeated 7-bp long DNA motif (TTAGGG)n, which is added by a telomerase enzyme, rather than through semi-conservative replication. This event solves the capping and replication problem of the ends of a DNA double helix (reviews: Fajkus et al., 2005; Watson and Riha, 2010). Because of this mode of addition to chromosomes, the copy number of the repeat unit has been found to vary both between different cells and different chromosomes (Figure 4c; Schwarzacher and Heslop-Harrison, 1991). The repetitive motif is not universal, but a 6 bp motif, as found in many mammals, (TTAGGG)n is present in some groups of plants (Sykorova et al., 2003a,b).

Centromeres
The centromere of plant metaphase chromosomes is normally visible as a sharp constriction along its length (Figures 1 and 3a,d), if not present near the end on acrocentric chromosomes. It acts as the focus where the proteinaceous kinetochore plate forms, to which the spindle microtubules attach. The centromeres of most plant species include large arrays of tandemly repeated DNA (Figure 4a; Heslop-Harrison et al., 1996; Maluszynska and Heslop-Harrison, 1991). Often, retrotransposon sequences (Gindullis et al., 2001). Genomic analysis has shown the presence of actively transcribed genes (Jiang et al., 2003; Mutti et al., 2010). However, despite the conservation of the function, the kinetochore proteins and the CenH3 histone that forms part of the nucleosomes core at centromeres of metaphase chromosomes, the DNA sequences at the centromere in different species are highly diverged and show considerable size variation (Ma et al., 2007). It is now clear that epigenetic mechanisms establish and propagate active centromeres on chromosomes, independent of their sequence (Jiang et al., 2003; Morris and Moazed, 2007).

Because of the epigenetic nature of centromeres, it is possible for a chromosome to have a ‘neo-centromere’ that is not always functional (Carvalho et al., 2008). It is also found that centromeres from one species may not nucleate microtubules strongly in another species background, and hence the chromosomes of one species do not segregate efficiently and are lost (Figure 2). In the hybrid Hordeum vulgare × Hordeum bulbosum, the chromosomes from many genotypes of H. bulbosum are lost during division (Bennett et al., 1976; mechanism investigated by Gernand et al., 2006), giving a haploid H. vulgare plant where the chromosome number can be doubled to generate homozygous plants. A very exciting approach to generating haploids came from Ravi and Chan (2010): noting that the centromeres of the eliminated genome were less able to interact with spindle microtubules, they made transgenic Arabidopsis plants with a CenH3 protein modified to be less efficient. When crossed to wild-type plants, chromosomes from the modified genome were eliminated, leading to the formation of haploids.

While the monocentric centromere as above is very widespread in the plant kingdom, two other types of centromere structure have been identified in eukaryotes. The localized point centromere from budding yeast Saccharomyces cerevisiae, with a DNA sequence of about 125 bp that provides specific kinetochore protein binding sites (Morris and Moazed, 2007), seems not to have any sequence similarity with the centromeres of plant and animal eukaryotes. The second centromere type is not localized on the chromosome, but functions to allow microtubules to bind along their complete length. The first animal to be fully sequenced, Caenorhabditis elegans, had these diffuse or holocentric centromeres, where the microtubules attach along the whole chromosome. Six families of plants (three monocots and three eudicots), have holocentric chromosomes. The association of microtubules along the whole chromosome length was observed by Guerra et al. (2006) in Rhynchospora tenuis (2n = 4; Cyperaceae). In this family, chromosome number varies up to 2n = circa 200, including many chromosomes <10 Mb in size, suggesting that chromosome fragmentation may have occurred during evolution, but the chromosomes are still able to segregate at division by binding microtubules. In contrast to these exceptionally small chromosomes, another genus with holocentric chromosomes, Cuscuta, has a large average chromosome size ranging up to 1000 Mb.

The rRNA sites and the nucleolus
As well as the centromeres, another constriction or gap is usually seen on some metaphase chromosomes in a complement – the secondary constriction at the NOR (Figures 1 and 3a, arrow). The NOR corresponds to major sites of the 45S rDNA, consisting of a tandem repeat of a unit with the 18S–5.8S–26S rRNA genes and their transcribed and untranscribed spacer regions (Figure 4b–d). The repeat unit is typically about 10 kb long, and in Arabidopsis it is present...
about 360 times on two pairs of chromosomes, representing about 5% of the DNA (Copenhaver and Pikaard, 1996; Heslop-Harrison and Maluszynska, 1994). In other species with larger genomes, such as wheat, the rRNA genes are present at a small number of discrete sites on the chromosomes (Figure 4d), with a larger number of copies of the repeat – 1200 at one locus in hexaploid wheat.

At interphase, the nucleolus, the most conspicuous structure within the nucleus, is the site of transcription of the rRNA repeat units and there is little stained DNA within the volume of the nucleolus. Untranscribed copies of the rDNA are often condensed and locate just outside the nucleolus, while in situ hybridization shows the transcribed genes as a decondensed thread running through the nucleolus (Figure 4b).

The 18S, 5.8S and 26S rRNA products come together with the 5S rRNA and the ribosomal proteins to make the ribosomes. The 5S rRNA genes, like the 18S–5.8S–26S rRNA genes, are present in the genome as a tandem repeat. Both the 45S and the 5S rRNA loci are often found to have ‘rearranged’ as blocks during evolution. In A. thaliana, the sites of the 5S rDNA are on different chromosomes in the Landsberg and Columbia ecotypes (Murata et al., 1997). In cereals, both the sites of the rDNA and the order of the loci, varies extensively between related species (Castilho and Heslop-Harrison, 1995). Where genetic maps are available, the change in position of the loci is not accompanied by transfer of regions of genes flanking the moved rRNA genes. (Dubcovsky and Dvorak, 1995).

THE CELL CYCLE AND THE INTERPHASE NUCLEUS

The physical structure of the plant cell nucleus changes through the cell cycle (Figure 2). The ‘framework’ within which these physical events happen can be regarded as the architecture of the nucleus. It is this architecture, in combination with the linear order of genes along the chromosomes, that is responsible for the higher-level organization of the nucleus, and the processes related to interactions between independent molecules or parts of macromolecules. The degree to which this framework involves a physical scaffold or is self-organizing remains uncertain. The processes involved in ‘decondensation’ of the chromosome to the interphase nucleus are also, in general, poorly understood, although likely to involve loops of chromatin extending from more condensed axes that are visible by light or electron microscopy. During interphase there may be a gradient across the nucleus in the proportion that is filled with chromatin, and chromatin may be more dense adjacent to the nuclear envelope, particularly in species with small genomes. The interphase nucleus itself is a dynamic environment, and both structural components and the DNA move during the interphase. Most obviously, soon after division, rRNA gene expression from multiple chromosomes (the homologous pair if only one pair of sites is present, or sites on several different chromosomes) form individual nucleoli. At later stages of the cell cycle, these have normally moved and fused to a smaller number of larger nucleoli. Interphase nucleus size varies within a single plant: the egg cell is often characterized by a large volume, with the chromatin being much dispersed through the whole volume, while the male sperm cell nucleus is highly condensed (Cao and Russell, 1997; Russell et al., 1996).

In 2003, Cremer and Cremer wrote ‘there is increasing agreement that the study of the functional architecture of the eukaryotic nucleus will be one of the most important post-genomic research areas’. Since writing this, chromatin research, involving understanding of the interactions of DNA and proteins has expanded, and the epigenetic consequences of chromatin modification have become clear (see Fransz and deJong, 2011). However, the relationship between nuclear organization, gene expression, higher-order chromatin arrangements and their interactions with other nuclear components, as considered by Cremer and Cremer (2001) remains a challenge to understand. Shopland and Bewersdorf (2008) discuss how recent advances in light microscopy are likely to reveal more information about chromosome structure and function, and point out that relatively little is known about the structural, dynamic, and mechanical properties of these macromolecular assemblies. Figure 6 illustrates the application of superresolution microscopy to resolve the synaptonemal complex at meiosis, where conventional light microscopy is unable to resolve the two lateral elements that are closer than 300 nm. Gustafsson et al. (2008) show that advanced systems have wide application to study chromosomal organization at high resolution, so in great detail.

SEX CHROMOSOMES AND SEX DETERMINATION IN PLANTS

More than 95% of angiosperm and gymnosperm species are hermaphroditic, bearing flowers with both pollen and ovules (as in Arabidopsis or wheat), or monoecious where both male and female flowers are carried on the same plant (as in maize) (Dellaporta and Calderon-Urrea, 1993). Some 4% of plants are dioecious, where male and female flowers are carried on different plants and, in most of these, sex is determined genetically. Dioecy is thought to have evolved relatively recently and independently in a number of plant families. In a few cases, dimorphic sex chromosomes were found such as in the ‘classic’ examples of Rumex species and Silene latifolia, as well as Humulus, Cannabis and Coccinia (see Figure 3b; Kejnovsky and Vyskot, 2010; Navajas-Perez et al., 2005a,b; Navajas-Pérez et al., 2009; Vyskot and Hobza, 2004). When cytologically homomorphic sex chromosomes are present, gene differences and sex-determining genes, including a MSY (male specific Y) region are found in male and female plants. Such non-heteromorphic sex-chromosome-like regions have been described in

several crop plants whose genomes have been sequenced such as papaya, grape and poplar (grape: Jaillon et al., 2007; papaya, Ming et al., 2008; poplar, Yin et al., 2008), as well as asparagus, kiwi and spinach.

Papaya is trioecious with XX female, XY male, and XyY hermaphroditic (Liu et al., 2004; Zhang et al., 2004). The Y is evolutionarily young and is estimated to have diverged from the X 2–3 million years ago. Within its male specific region, some 13% of the Y, including the centromere and highly methylated heterochromatic knobs have been found (Zhang et al., 2008) and numerous chromosomal rearrangements have been detected (Yu et al., 2008). In poplar, Yin et al. (2008) have identified a region of one chromosome showing characteristics of a sex chromosome with a gender-associated locus. Reduced recombination, distorted segregation and haplotype divergence was only observed in the female and consequently sex determination in *Populus* is an incipient ZW chromosome system where males are ZZ and females are the ZW heterogametic sex.

Plant sex chromosome evolution occurred recently, and is still ongoing, so provides an excellent model to study DNA sequence and chromosome evolution. It is believed that the process started with the emergence of sex determining genes (X has male sterility and female fertility; Y has maleness factor and female suppressor) followed by suppression of recombination in their surrounding region (for review see Bergero and Charlesworth, 2009; Kejnovsky and Vyskot, 2010; Navajas-Perez et al., 2005a,b). Thus cytological homomorphic sex chromosomes with their heteromorphic DNA regions could represent this first step and are indeed often found to be younger than dimorphic sex chromosomes. The expansion of suppression of recombination to the majority of the chromosome is postulated to lead to accumulation of deleterious mutations, erosion of genes caused by insertion of retroelements or DNA transposons and finally degeneration. As a result heteromorphic sex chromosomes emerge that are often larger than the autosomes in plants (Figure 3b) due to accumulation of repetitive DNA elements (see below) and are in contrast to the small mammalian Ys that are much older and have been allowed to loose genes by rearrangements (Bergero and Charlesworth, 2009).

Molecular investigations have shown that the Y chromosome of *Silene latifolia* estimated to be about 10 million years old shows all of above signs of sex chromosome evolution including genetic degeneration, reduction of DNA polymorphism, accumulation of mutations at important functional sites coding for proteins, and gene expression changes (see Armstrong and Filatov, 2008; Filatov et al., 2009). Analysis of the repetitive DNA distribution and comparing female and male DNA sequences on *S. latifolia* sex chromosomes, has revealed that parts of the Y chromosome have diverged from the X at different times and can be divided into ‘strata’ similar to the human Y. Different amounts of various DNA sequence families, from almost all classes of repeats known in plants, are present on the Y in large numbers. Cermak et al. (2008) undertook a survey of all repeats on the Y of *S. latifolia* and found in decreasing abundance, subtelomeric tandem repeats, gypsy and copia like retroelements, followed by LINEs and SINEs and DNA transposons including hATs and MITEs. Interestingly, they and Filatov et al. (2009) found a transposable element (TE) abundant on autosomes that is excluded on the Y indicating a divergent evolution of DNA sequences on sex and autosomal chromosomes.

Accumulation of repetitive DNA sequences has also been seen in the genus *Rumex*, which contains several species with dimorphic sex chromosomes and a derived complex XX/XY1Y2 system in *R. acetosa*, *R. papilaris* and...
R. hastatus (Navajas-Pérez et al., 2005a,b; see also Figure 3b). The Y degeneration in XX/XY1Y2 system was accompanied by massive accumulation of repetitive DNA followed by chromosomal rearrangements giving rise to the multiple Y chromosomes (Mariotti et al., 2009; Navajas-Pérez et al., 2009). The loss of recombination between X and Y chromosomes would reduce the evolutionary rate of Y-specific satDNAs, but also hinders intra-specific homogenization processes. As a consequence, different rates of evolution have been found for autosomal and sex chromosome variants of repeats, as well as differential patterns of Y-heterochromatin as well as the presence of different subfamilies and related satDNAs in different regions of the Y chromosomes (Mariotti et al., 2009; Navajas-Pérez et al., 2005a,b; Navajas-Pérez et al., 2006, 2009). Further the Y chromosome experienced many inversions of various extents.

Additional evidence of repeat accumulation at different times during the evolution of the Y chromosomes, comes from the studies of simple sequence repeats that have accumulated in the Y chromosome of Silene especially in the longer arm which has stopped recombining relatively recent and harbours no other repeats yet (Kejnovsky et al., 2009). In Rumex acetosa several simple sequence repeats including (ACC) (see Figure 3b) are found highly amplified throughout both Y chromosomes except towards one telomere, presumably the pseudoautosomal regions. The autosomes and X chromosome show much lower levels with several distinct bands along most chromosomes similar to the pattern found in wheat and rye chromosomes (see Figure 3c; Cuadrado and Schwarczacher, 1998).

THE SIGNIFICANCE OF CHROMOSOME ORGANIZATION

The chromosome is a key level of organization of the plant genome, providing the structure for the genetic linkage groups, allowing replication, transcription and transmission of the genome, and allowing whole genome duplication and physical reorganization. Following completion of the Arabidopsis and other genome sequences, the widespread presence of segmental and whole genome duplications across angiosperms is much more frequent than was suspected from earlier studies. Comparative genomics using whole genome sequencing complemented by molecular cytogenetics has provided new insights into the nature of chromosomal rearrangements including fusions, fissions, inversions, deletions and duplications, across a much wider groups of plants than has been possible with cytogenetic approaches alone. These episodic events combine with continuous processes including sequence mutation, transposable element accumulation, tandem repeat amplification and sequence homogenization. Improved methods of chromosomal analysis with in-situ hybridization and use of antibodies are assisting characterization of genome-wide and chromosome-level changes in the genome. The fundamental insights gained from these studies are now showing how genomes evolve and how diversity can be generated.

So far, the controls on many features of chromosome organization and their variability remain to be elucidated. Why should different species have genomes varying in size by more than 2000-fold, and both chromosome number and chromosome sizes vary by 300-fold? The behaviour of these genomes seems to be similar in terms of replication, gene expression, control and evolution, or at least differences do not reflect the huge variation in genome organization. Indeed, it is remarkable that the same genetic, segregation, expression, replication and evolutionary mechanisms seem to be applicable over this large range. Crop plants represent an intensively selected subset of <0.1% of the 400 000 angiosperm species, and fewer than 30 species provide more than 97% of the world’s food (FAOstat, 2010). Even among the top crops, the variation in nature of genomes is evident with diploids, recent polyploids, and hybrid species, and genome sizes between 465 Mb in rice to 17 000 Mb in wheat. Exploiting the diversity and evolutionary mechanisms in plant genomes is likely to be a key to crop development for food production.

REFERENCES


Dear Author,

During the copy-editing of your paper, the following queries arose. Please respond to these by marking up your proofs with the necessary changes/additions. Please write your answers on the query sheet if there is insufficient space on the page proofs. Please write clearly and follow the conventions shown on the attached corrections sheet. If returning the proof by fax do not write too close to the paper’s edge. Please remember that illegible mark-ups may delay publication.

Many thanks for your assistance.

<table>
<thead>
<tr>
<th>Query reference</th>
<th>Query</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>AUTHOR: Please check article title.</td>
<td>Title changed; somehow included Special issue title</td>
</tr>
<tr>
<td>Q2</td>
<td>AUTHOR: Journal style requires that the first forenames of all authors should be given. Please provide missing details.</td>
<td></td>
</tr>
<tr>
<td>Q3</td>
<td>AUTHOR: Please provide fax number for corresponding author</td>
<td></td>
</tr>
<tr>
<td>Q4</td>
<td>AUTHOR: Please check keywords.</td>
<td></td>
</tr>
<tr>
<td>Q5</td>
<td>AUTHOR: Please check all website address and confirm that it is correct. (Please note that it is the responsibility of the author(s) to ensure that all URLs given in this article are correct and useable.)</td>
<td>FAX added</td>
</tr>
<tr>
<td>Q6</td>
<td>AUTHOR: Goremykin et al., 2008 has not been included in the Reference List, please supply full publication details.</td>
<td>Websites work should be G. 2009</td>
</tr>
<tr>
<td>Q7</td>
<td>AUTHOR: Gayral et al., 2009 has been changed to Gayral and Iskra-Caruana, 2009 so that this citation matches the Reference List. Please confirm that this is correct.</td>
<td></td>
</tr>
<tr>
<td>Q8</td>
<td>AUTHOR: Schwarzacher et al., 2008 has been changed to Schwarzacher, 2008 so that this citation matches the Reference List. Please confirm that this is correct.</td>
<td></td>
</tr>
<tr>
<td>Q9</td>
<td>AUTHOR: Aguiar-Perecin and Vosa, 1985 has been changed to Aguiar-Perecin and de Vosa, 1985 so that this citation matches the Reference List. Please confirm that this is correct.</td>
<td></td>
</tr>
<tr>
<td>Q10</td>
<td>AUTHOR: Hart and Bleij 2003 has been changed to ‘t Hart and Bleij, 2003 so that this citation matches the Reference List. Please confirm that this is correct.</td>
<td>10. correct</td>
</tr>
<tr>
<td>Q11</td>
<td>AUTHOR: Paux et al., 2007 has been changed to Paux et al., 2006 so that this citation matches the Reference List. Please confirm that this is correct.</td>
<td>Correct 2006</td>
</tr>
<tr>
<td>Q12</td>
<td>AUTHOR: Šafář et al., 2009 has been changed to Šafár et al., 2004 so that this citation matches the Reference List. Please confirm that this is correct.</td>
<td></td>
</tr>
<tr>
<td>Q13</td>
<td>AUTHOR: That’s not enough to reduce 12 to 7</td>
<td></td>
</tr>
<tr>
<td>Q14</td>
<td>AUTHOR: Froenicke et al., 2005 has been changed to Froenicke et al., 2006 so that this citation matches the Reference List. Please confirm that this is correct.</td>
<td></td>
</tr>
<tr>
<td>Q15</td>
<td>AUTHOR: Heslop-Harrison et al., 1991 has not been included in the Reference List, please supply full publication details.</td>
<td></td>
</tr>
<tr>
<td>Q16</td>
<td>AUTHOR: McCouch et al., 1998 has been changed to McCouch et al., 1988 so that this citation matches the Reference List. Please confirm that this is correct.</td>
<td></td>
</tr>
<tr>
<td>Q17</td>
<td>AUTHOR: Heslop-Harrison et al., 1996 has not been included in the Reference List, please supply full publication details.</td>
<td></td>
</tr>
<tr>
<td>Q18</td>
<td>AUTHOR: Murata et al., 1998 has been changed to Murata et al., 1997 so that this citation matches the Reference List. Please confirm that this is correct.</td>
<td></td>
</tr>
<tr>
<td>Q19</td>
<td>AUTHOR: Please check whether Navajes-Perez et al., 2005 belongs to 2005a or 2005b.</td>
<td></td>
</tr>
<tr>
<td>Q20</td>
<td>AUTHOR: Zhang et al., 2004 has not been included in the Reference List, please supply full publication details.</td>
<td></td>
</tr>
<tr>
<td>Q21</td>
<td>AUTHOR: Bergero and Charlesworth 2008 has been changed to Bergero and Charlesworth, 2009 so that this citation matches the Reference List. Please confirm that this is correct.</td>
<td></td>
</tr>
<tr>
<td>Q22</td>
<td>AUTHOR: Anhalt et al. (2008) has not been cited in the text. Please indicate where it should be cited; or delete from the Reference List.</td>
<td></td>
</tr>
<tr>
<td>Q23</td>
<td>AUTHOR: Anhalt et al. (2009) has not been cited in the text. Please indicate where it should be cited; or delete from the Reference List.</td>
<td></td>
</tr>
<tr>
<td>Q24</td>
<td>AUTHOR: Please provide the volume number, page range for reference Bennett and Leitch (2011) if now on print.</td>
<td></td>
</tr>
<tr>
<td>Q25</td>
<td>AUTHOR: Please confirm that Journal title ‘J. Hered.’ has been abbreviated correctly.</td>
<td></td>
</tr>
<tr>
<td>Q26</td>
<td>AUTHOR: Please provide the volume number, page range for reference Bowler (2011).</td>
<td></td>
</tr>
<tr>
<td>Q27</td>
<td>AUTHOR: Brandes et al. (1997) has not been cited in the text. Please indicate where it should be cited; or delete from the Reference List.</td>
<td></td>
</tr>
<tr>
<td>Q28</td>
<td>AUTHOR: Carroll and Straight (2005) has not been cited in the text. Please indicate where it should be cited; or delete from the Reference List.</td>
<td></td>
</tr>
</tbody>
</table>

12. correct

14. correct

15. Ref added

16. correct

17. Changed to 1999; brandes added

18. correct

19. b deleted and corrected

20. changed to 2008

21. changed to 2009

22. inserted

23. deleted

24. Vol/page added

25. yes

26. This special issue - needs adding by PlantJ

27. cited now

28. cited now
| Q29 | AUTHOR: Contento et al. (2005) has not been cited in the text. Please indicate where it should be cited; or delete from the Reference List. |
| Q30 | AUTHOR: Cuadrado and Jouve (2010) has not been cited in the text. Please indicate where it should be cited; or delete from the Reference List. |
| Q31 | AUTHOR: Dolezel et al. (2007) has not been cited in the text. Please indicate where it should be cited; or delete from the Reference List. |
| Q32 | AUTHOR: Endo and Gill (1996) has not been cited in the text. Please indicate where it should be cited; or delete from the Reference List. |
| Q33 | AUTHOR: Please provide the document title, volume number, page range for reference Fransz and deJong (2011). |
| Q34 | AUTHOR: Please check the journal name for ‘J. Plant Registrations’ for correctness. |
| Q35 | AUTHOR: Please provide the volume number, page range for reference Green (2011). |
| Q36 | AUTHOR: Harper (1999) has not been cited in the text. Please indicate where it should be cited; or delete from the Reference List. |
| Q37 | AUTHOR: Please provide the city location of publisher for reference ’t Hart and Bleij (2003). |
| Q40 | AUTHOR: Heslop-Harrison et al. (1999) has not been cited in the text. Please indicate where it should be cited; or delete from the Reference List. |
| Q41 | AUTHOR: Ishii et al. (2010) has not been cited in the text. Please indicate where it should be cited; or delete from the Reference List. |
| Q42 | AUTHOR: Please provide the city location of publisher for reference Karwur (2001). |
| Q43 | AUTHOR: Karwur (2001) has not been cited in the text. Please indicate where it should be cited; or delete from the Reference List. |
| Q44 | AUTHOR: Kubis et al. (1998) has not been cited in the text. Please indicate where it should be cited; or delete from the Reference List. | 44. now cited |
| Q45 | AUTHOR: Leitch et al. (2010) has not been cited in the text. Please indicate where it should be cited; or delete from the Reference List. | 45 now cited |
| Q46 | AUTHOR: Lysak et al. (2005) has not been cited in the text. Please indicate where it should be cited; or delete from the Reference List. | should be 2005 not 2006 at one point in text - changed. |
| Q47 | AUTHOR: Murray et al. (2002) has not been cited in the text. Please indicate where it should be cited; or delete from the Reference List. | 47 cited |
| Q48 | AUTHOR: Nagaki et al. (2005) has not been cited in the text. Please indicate where it should be cited; or delete from the Reference List. | 48 sentence added |
| Q49 | AUTHOR: Please check author names for reference Navajas-Pérez et al. (2009) | 49 Ruiz added |
| Q50 | AUTHOR: Please provide the volume number, page range for reference Özkan et al. (2010). | 50 cited as in journal |
| Q51 | AUTHOR: Pearce et al. (1997) has not been cited in the text. Please indicate where it should be cited; or delete from the Reference List. | 51 ref added |
| Q52 | AUTHOR: Please provide the volume number, page range for reference Proost et al. (2011). | 52 this special issue plant journal added citation. |
| Q53 | AUTHOR: Schwarzacher et al. (1992) has not been cited in the text. Please indicate where it should be cited; or delete from the Reference List. | |
| Q54 | AUTHOR: Sharp et al. (1989) has not been cited in the text. Please indicate where it should be cited; or delete from the Reference List. | |
| Q55 | AUTHOR: Stoilov et al. (2007) has not been cited in the text. Please indicate where it should be cited; or delete from the Reference List. | 55 deleted citation |
| Q56 | AUTHOR: Suda and Leitch (2010) has not been cited in the text. Please indicate where it should be cited; or delete from the Reference List. | added in fig. legend |
| Q57 | AUTHOR: Wang et al. (2009) has not been cited in the text. Please indicate where it should be cited; or delete from the Reference List. | added citation. |
| Q58 | AUTHOR: Wolfgruber et al. (2009) has not been cited in the text. Please indicate where it should be cited; or delete from the Reference List. | added citation. |
| Q59 | AUTHOR: Yan et al. (2006) has not been cited in the text. Please indicate where it should be cited; or delete from the Reference List. | 59 added citation |
Q60  AUTHOR: Cuadrado et al., 1998 has not been included in the Reference List, please supply full publication details.
USING E-ANNOTATION TOOLS FOR ELECTRONIC PROOF CORRECTION

Required Software
Adobe Acrobat Professional or Acrobat Reader (version 7.0 or above) is required to e-annotate PDFs. Acrobat 8 Reader is a free download: http://www.adobe.com/products/acrobat/readstep2.html

Once you have Acrobat Reader 8 on your PC and open the proof, you will see the Commenting Toolbar (if it does not appear automatically go to Tools>Commenting>Commenting Toolbar). The Commenting Toolbar looks like this:

If you experience problems annotating files in Adobe Acrobat Reader 9 then you may need to change a preference setting in order to edit.
In the “Documents” category under “Edit – Preferences”, please select the category ‘Documents’ and change the setting “PDF/A mode:” to “Never”.

Note Tool — For making notes at specific points in the text
Marks a point on the paper where a note or question needs to be addressed.

Replacement text tool — For deleting one word/section of text and replacing it
Strikes red line through text and opens up a replacement text box.

Cross out text tool — For deleting text when there is nothing to replace selection
 Strikes through text in a red line.
Approved tool — For approving a proof and that no corrections at all are required.

Highlight tool — For highlighting selection that should be changed to bold or italic. Highlights text in yellow and opens up a text box.

Attach File Tool — For inserting large amounts of text or replacement figures as a file. Inserts symbol and speech bubble where a file has been inserted.

Pencil tool — For circling parts of figures or making freeform marks

Create freeform shapes with a pencil tool. Particularly with graphics within the proof it may be useful to use the Drawing Markups toolbar. These tools allow you to draw circles, lines and comment on these marks.

How to use it:
1. Select Tools > Drawing Markups > Pencil Tool
2. Draw with the cursor
3. Multiple pieces of pencil annotation can be grouped together
4. Once finished, move the cursor over the shape until an arrowhead appears and right click
5. Select Open Pop-Up Note and type in details of required change
6. Click the X in the top right hand corner of the note box to close.
Help
For further information on how to annotate proofs click on the Help button to activate a list of instructions: