GENOMICS ARTICLE

Comparative Genome Organization in Plants: From Sequence and Markers to Chromatin and Chromosomes

J. S. Heslop-Harrison[1]

John Innes Centre, Norwich NR4 7UH, United Kingdom

INTRODUCTION

Comparative studies have provided the basis for some of the most important discoveries in biology. The study of differences, whether at the level of gene alleles or living kingdoms, has shown the critical features and function of most biological structures. The framework for comparative studies of organisms was perhaps laid out by the earliest taxonomists and medics: Dioscorides (80) and others used chemosystematic properties and morphology to group plants with similar medicinal properties. In the first half of the twentieth century, cytologists such as Darlington (1931; Darlington and LaCour, 1942), Kihara (1924), and Sears (1941) studied plant chromosomes, making key discoveries such as the importance of polyploidy in plant evolution and showing that whole chromosomes carried similar groups of genes in different species. Soon thereafter, an understanding of evolutionary changes at the cytological level was established that included an appreciation for chromosome translocation, fusion, fission, and the correlation between chromosomal particularities and phylogeny. Today, molecular markers show that gene orders are conserved over substantial evolutionary distances (Gebhardt et al., 1991; Ahn et al., 1993; Devos and Gale, 1993, 1997; see also Devos and Gale, 2000, in this issue), with the number of chromosomal and genetic differences between species generally increasing with evolutionary distance (Bennetzen et al., 1998; Tikhonov et al., 1999).

Comparative studies are useful in elucidating the function of biological structures and in providing markers for evolutionary investigation, whether in the context of plant breeding, ecology, or biodiversity. Evolutionary comparisons, moreover, encompassing the very origin of life, complement other genomic studies. As pointed out by Gesteland and Atkins (1993), many relics of the "RNA world," existing 3.5 billion years ago, have been discovered in modern organisms, demonstrating the extraordinary conservation of nucleic acid sequence and function. Examples of such relics include ribosomes (where the RNA aggregate in the absence of proteins is able to synthesize peptide bonds; Nitta et al., 1998), ribozymes, and features of codon usage (see Heslop-Harrison, 2000). Comparative studies enable the application of data from one species to investigations of taxonomically disparate species, as exemplified in the use of *Escherichia coli*, human, Arabidopsis, and rice in understanding the wheat genome. The genome sequencing projects are in fact based on the premise that knowledge of the whole sequence of Arabidopsis, for instance, will aid in the isolation from crop plants of agronomically important genes that bear homology to Arabidopsis genes. Work from one kingdom often suggests models to test in another, although techniques easily applicable in one species may be impossible or inappropriate in another.

Many features of plant, animal, fungal, and even prokaryotic genomes are remarkably similar, but there are some elements that are not conserved (e.g., centromeres; see below). A notable feature of angiosperms is the widespread occurrence of polyploidy—even over experimentally observable time frames—involving either the doubling of chromosome numbers within a species or the interspecific union of chromosome sets. In contrast, recent and evolutionarily significant polyploidy is unusual in gymnosperms, most vertebrates, and well-studied species such as Caenorhabditis or Drosophila. One might suggest that the genes enabling regular meiosis with strict bivalent formation between homologous but not homeologous pairs of chromosomes are an early feature of angiosperm evolution. Such genes might conceivably control the stringency and timing of meiotic chromosome pairing as well as influence the disparate organization of repetitive DNA sequences in plants and animals.

THE LINEAR DNA SEQUENCE

With the first sequences of complete plant chromosomes now published (Lin et al., 1999; Mayer et al., 1999), it is

[1] E-mail Pat.heslop-harrison@bbsrc.ac.uk; fax 44-1603-450045.

appropriate to consider the relationship of linear sequence information to the organization and function of chromosomes within the context of the genome. DNA sequence data have averred the general model of the structure of the DNA component of the chromosome. Sequencing provides direct evidence that the double-helical DNA molecule is continuous from telomere to telomere, and there is no evidence for bases other than A, T, G, and C. These conclusions, although long anticipated, are not trivial, because the resolution from sequencing is magnitudes greater than can be observed by microscopy, and biochemical assays are hampered when handling DNA molecules of several million kilodaltons. It is also worth noting that the prokaryotic vectors and enzymes used in sequencing can process the angiosperm DNA in its entirety, confirming the universal nature of DNA.

Arabidopsis was chosen as the first plant target for complete sequencing because its genome size (130 to 140 Mbp) is rather small, $\sim$200 times smaller than other plant genomes (see http://www.rbgkew.org.uk/cval/database1.html); the Arabidopsis genome is diploid, with five pairs of chromosomes ($2n = 2x = 10$). There is significant correlation between genome size and plant niche (e.g., Bennett et al., 1998), although this is not immediately obvious: rye (9000 Mbp genome size; $2n = 2x = 14$) and Arabidopsis are short-lived, annual plants, whereas oak (200 Mbp) and pines (23,000 Mbp) are temperate trees. Admittedly, the angiosperms represent an unusually broad taxonomic spectrum, but among birds, a taxon that is also quite large, genome sizes vary by only a factor of two (from 2000 to 3800 Mbp; Tiersch and Wachtel, 1991).

The division of genomic DNA into independent chromosomes is a fundamental feature of genome architecture. Like genome size, chromosome number varies widely among plant species, such that $2n$ ranges in value from 4 to more than 1000, although the number within any given species, with the exception of supernumerary or B chromosomes, is usually constant. Some taxa, such as the family Cruciferae, have highly variable chromosome numbers, whereas the number is conserved in others. Polyploids tend to have higher chromosome numbers, and species in which n is a multiple of 6 or 7 are frequent. Nevertheless, except for polyploidy, there are few plant characteristics that correlate clearly with chromosome number. Of course, chromosome number has a genetic consequence in that genes are reassorted at meiosis on different chromosomes.

The placement of genes and their introns within the broader genomic context is an important area of research, involving detailed annotation of the genome (see www.tigr.org for the current status in Arabidopsis). Most of the genes found in species with much larger genomes are present in Arabidopsis, which is estimated to contain $\sim$25,000 genes. But the smallness of the Arabidopsis genome means that other characteristic sequence regions, such as those that are highly repeated, are less abundant than in other species. DNA motifs, ranging in length from a single base to thou-

sands of bases, repeated many hundreds or thousands of times, are a characteristic of all eukaryotic genomes and represent between 50 and 90% or more of all DNA. In Arabidopsis, many duplications of gene sequences have been found both within (e.g., $\sim$250 tandem duplications each $\sim$10 kb on chromosome 2) and between chromosomes (e.g., regions $\sim$4 Mb long between chromosomes 2 and 4, or 700 kb long between chromosomes 1 and 2; Lin et al., 1999). Furthermore, a large part of the mitochondrial genome, $\sim$270 kb, is inserted into chromosome 2. The transfer of genes from organelles to nucleus over evolutionary time is now well established in plants and animals (Martin and Herrmann, 1998; Vaughan et al., 1999).

To fully exploit sequence information from Arabidopsis, Caenorhabditis (*C. elegans* Sequencing Consortium, 1998), yeast (Goffeau et al., 1996), and human chromosomes (Dunham et al., 1999), we must also appreciate what sequence data alone cannot tell us. Modified bases, and particularly 5-methylcytosine, are not distinguished from their unmodified equivalents, and there is no information about chromatin packaging and three-dimensional organization, topics that are essential to a complete understanding of the genome. There are, moreover, a few gaps at which DNA sequence is not available. In Arabidopsis, such regions occur at the centromeres and at nucleolar organizer regions in the two chromosomes sequenced to date, and there are more extensive gaps at 11 sites in human chromosome 22 (Dunham et al., 1999). Below, I discuss the localization of key sequence motifs along the plant chromosome, using Arabidopsis, the triticeae cereals, and pine as examples to support a general model of sequence distribution along plant chromosomes (Schmidt and Heslop-Harrison, 1998).

REPETITIVE DNA SEQUENCES AND THE LARGE-SCALE ORGANIZATION OF THE CHROMOSOME

The regions that have remained inaccessible from the otherwise fully sequenced chromosomes consist, at least in the Arabidopsis chromosomes and most of human chromosome 22, of long and relatively homogeneous stretches of repetitive DNA motifs. Such stretches are not tractable by current technologies that read sequences of only a few hundred base pairs that must then be ordered so that they represent the complete chromosome. It is important to know about the length of the sequence gaps, the homogeneity of repeat motifs, and the level of variation within the motifs before one can to begin to hypothesize how they evolve and function in the context of the genome. In Arabidopsis, multiple random fragments of individual bacterial artificial chromosomes, averaging 80 kb long, show the repetitive sequence to be homogeneous, with no interspersion of low-copy sequence among the tandem 180-bp repeats (Lin et al., 1999); however, some genes exist amid the repeated sequences, found in both forward and reverse orientations, in

the centromeric regions (Figure 1). The extent to which repeat motif variants relate to chromosomal function as opposed to satisfying some "bulk" requirement remains to be determined.

Of sequence motifs that are highly repeated, some are highly conserved from one species to another: the rRNA genes are present as hundreds of copies with only a small percentage of variation in all eukaryotes. Other sequence motifs are extremely variable, even between accessions of a species, providing tools to assess potential functions of particular aspects of genome architecture and for studying interorganismal relationships. Hence, the study of repetitive DNA sequence motifs and their chromosomal distribution in a comparative context—comparing, for instance, sequences from Arabidopsis and wheat (Figure 2) versus conifers and Crocus—has considerable potential for understanding genome evolution and sequence components. Specifically, individual sequences of a particular repeated motif may vary
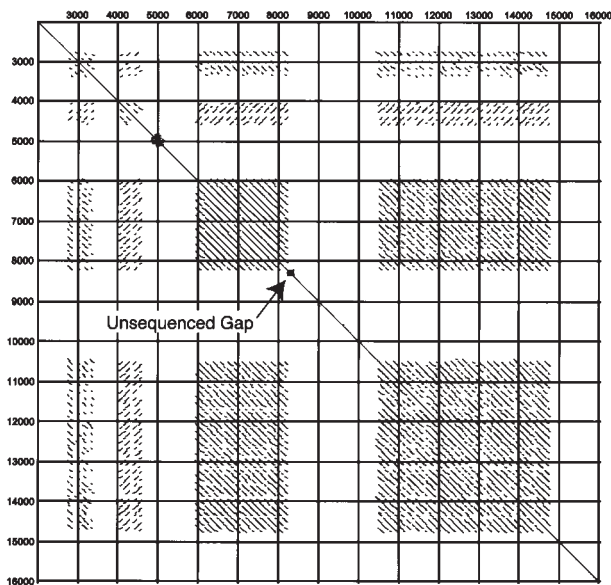
in both their copy number and exact sequence, giving rise to the concept of sequence families (see Figure 1). Various classes of repeated sequences (see below) are easily recognized: (1) tandemly repeated sequences, in which one copy follows another in an array of many tens or even thousands of copies; (2) retroelements, in which amplification occurs through an RNA intermediate (acting as a template for protein translation as well as DNA transcription) before reinsertion into the genome; and (3) those that are special classes such as telomeric sequences or rDNA units.

Cytogenetic methods offer a powerful system for looking at the organization of DNA repeat motifs along a chromosome using in situ hybridization of labeled probe sequences to the denatured DNA of chromosomes spread on microscope slides. The techniques are robust and reliable, with chromosomal target regions containing a few kilobases (even if dispersed over much longer chromosome segments) of sequence homologous to the probe (Schmidt and Heslop-Harrison, 1998; Schwarzacher and Heslop-Harrison, 2000). Sequences suspected to be abundant because of their frequent occurrence in a library or strength of membrane hybridization may prove difficult to interpret in terms of chromosome localization because hybridization probes to size-separated restriction fragments often give multiple dense bands or smears. Where sequence information is available, results are consistent with those from in situ hybridization. For example, the completed sequence of Arabidopsis chromosome 2 (Lin et al., 1999) confirms in situ hybridization data that had unexpectedly shown *copia*-like retroelements to be dispersed along the chromosome arms and clustered at centromeric regions (Figure 2; Brandes et al., 1997a). In polyploid species such as bread wheat, some sequences are much more abundant in one chromosome set than in another (Figure 2).

In situ hybridization methods also offer advantages in comparing different accessions or species. In addition, viral and mitochondrial sequences within the nucleus of various accessions can be located by in situ hybridization without knowledge about the nuclear flanking DNA. Analytical difficulties by in situ hybridization are encountered in relation to neither genome size nor repetition of a sequence motif.

Figure 1. Tandem Repetition of a 180-bp Motif in Arabidopsis Chromosome 2.

The dot plot shows the tandem repetitition of a 180-bp motif within 14 kb of sequence around the centromeric region (sequence data from Lin et al., 1999; see www.tigr.org). A dot is printed whenever 90% of bases within a window of 30 as numbered along the abscissa match those numbered on the ordinate. The continuous diagonal shows that the sequences along the two axes are identical, whereas the many other diagonal lines show that particular motifs, mostly the 180-bp motif, are repeated in tandem at multiple sites. Diagonals that represent obtuse angles relative to the abscissa indicate forward repeats; diagonal hits representing acute angles represent reversed complementary sequences. The repeat motifs are uniform in length (parallel lines have equal spacing), but repeat blocks vary in length.

rDNA

The 45S rDNA loci consist of tandem arrays of repeating units of the 18S, 5.8S, and 26S rRNA genes and the transcribed and nontranscribed spacers, each unit being typically 10 kb long in plants. Hundreds or thousands of copies of the repeat units may be present, together representing up to ~10% of the genome (8% in Arabidopsis; Pruitt and Meyerowitz, 1986). The units, along with the 5S rRNA genes (occurring as tandem repeats independent of the 45S rDNA), are localized at one or more sites per chromosome set, and their characteristic positions along chromosomes provide useful markers for chromosome identification (see,
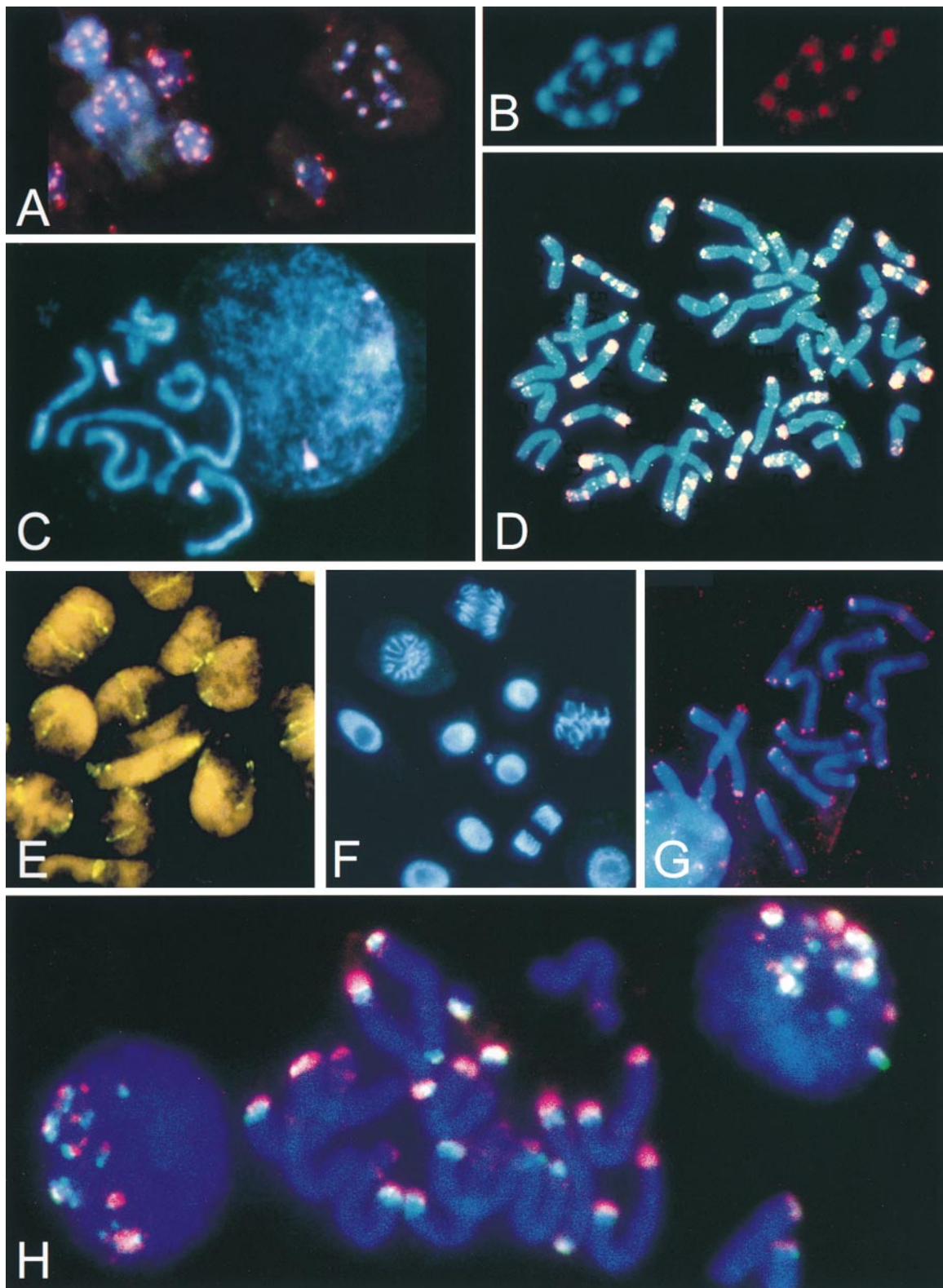
Figure 2. In Situ Hybridization of Metaphase Chromosomes and Interphase Nuclei from Various Species.

e.g., Doudrick et al., 1995). The units themselves are highly conserved, and probes isolated originally from wheat can be used to localize the 45S and 5S genes in most eukaryotic species (Figure 2). Changes in chromosomal distribution of the units generally correlate with the rates of speciation, and they have been used, for example, to examine evolutionary trends in the Triticeae (Figure 3; Castilho and Heslop-Harrison, 1995; de Bustos et al., 1996; Taketa et al., 1999).

Telomeres

Telomeres are specialized structures that stabilize chromosome ends and enable replication (see Zakian, 1995). The telomeric region is highly conserved and consists of a short repeat, the sequence of which is similar to TTTAGGG, in tandem arrays many hundreds of units long at the physical ends of chromosomes in most eukaryotes (Drosophila is a notable exception; see Fuchs et al., 1995). Unlike most chromosomal DNA, terminal sequences cannot be fully replicated by a semiconservative mechanism but rather require the enzyme telomerase to supply an RNA template at the DNA terminus. The number of telomeric repeats is a species-specific characteristic, equivalent to 2 to 5 kb in Arabidopsis (Richards and Ausubel, 1988), 12 to 15 kb in cereals (Figure 2; see also Schwarzacher and Heslop-Harrison, 1990), and up to 60 to 160 kb in tobacco (Fajkus et al., 1995). The number of copies of the repeat also differs among the chromosome arms of the karyotype (Figure 2; Schwarzacher and Heslop-Harrison, 1990) and possibly varies from cell to cell and tissue to tissue (Kilian et al., 1995). Through its ability to attach the telomeric sequences to new chromosomal ends, telomerase also provides a mechanism to stabilize and repair broken chromosomes (Wang et al., 1992). In a sugar beet (*Beta vulgaris*) hybrid line incorporating an alien chromosome fragment from *B. procumbens*, telomeric sequences were detectable by in situ hybridization on all chromosome ends except one terminus of the alien fragment. Perhaps the particular conformation of the DNA at this end precludes the action of telomerase and thereby leads to lack of stability of the line (Schmidt et al., 1997).

Subtelomeric repetitive sequences have often been revealed by staining patterns of chromosomes. Analysis of these sequences on rye chromosomes shows that they are able to evolve in copy number rapidly (Alkhimova et al., 1999) and may be part of a complex chromosome end structure (Vershinin et al., 1995; Figure 2). Zhong et al. (1998) have used in situ hybridization to show that each chromosome end in tomato has a unique organization of the telomeric and a particular subtelomeric repeat, with large differences in lengths of each array. On chromosome 4 of Arabidopsis, the tandem repeats of the rDNA abut the telomeric repeats with <500 bp intervening (Copenhaver and Pikaard, 1996). Using the telomeric sequence to probe discrete chromosomal fragments resolved by pulsed-field gel electrophoresis, Ganal et al. (1992) were able to determine the genetic ends of chromosomes and hence show the complete map of tomato in terms of centimorgans.

Centromeres

During mitosis and meiosis, chromosomal segregation depends on the attachment of microtubules, however indirectly,

---

Figure 2. (Continued).

Chromosomes are counterstained in light blue with the DNA stain DAPI. Sites of hybridization of labeled DNA probes to homologous sequences along the chromosomes are detected with red or green fluorochromes.

(A) Metaphase and interphase chromosomes from Arabidopsis probed with the 180-bp tandem repeat motif (see Figure 1). The sequence is located around the centromeres of all five chromosome pairs.

(B) A metaphase Arabidopsis cell probed with a fragment of a *copia* group retroelement. The retroelement is abundant in the centromeric region of all chromosomes but, unlike the 180-bp repeat, shows hybridization along all the chromosome arms, which is indicative of dispersed sites.

(C) The large chromosomes of Crocus (with a genome some 50 times larger than Arabidopsis) probed with the 45S rDNA sequence from wheat. This sequence is highly conserved, and homologous sequences are present in most species.

(D) Metaphase chromosomes of hexaploid wheat ($2n = 6x = 42$, with A, B, and D genomes) probed with the tandemly repeated DNA sequence dpTa1 (red), which hybridizes to multiple sites. The sequence characterizes each chrommomsome but is predominantly located on the D-genome chromosomes. The sequence labeled green-white, pSc119.2, is clustered in the terminal regions of many chromosomes.

(E) Multiple interphase nuclei of a wheat variety that carries a rye chromosome arm (1BL.1RS translocation). DNA is counterstained in orange, and the rye arm is seen in yellow. One pole of each nucleus has a higher proportion of the volume filled with DNA.

(F) Nuclei from a Triticeae cereal fixed with aldehyde and stained with DAPI, showing nuclei at different stages of the cell cycle during which chromosomes show different organization and activity.

(G) Telomeric sequences in a rye metaphase probed with the synthetic oligomeric sequence $(TTTAGGG)_6$. The sequence is present at both ends of all seven chromosome pairs, but differences in copy number are reflected by the intensities of hybridization at each chromosome terminus.

(H) The locations of two nonhomologous subtelomeric sequences in rye (red and cyan) on metaphase chromosomes and within two interphase nuclei. All telomeric sequences are near one pole of the nucleus, away from the centromeric pole.
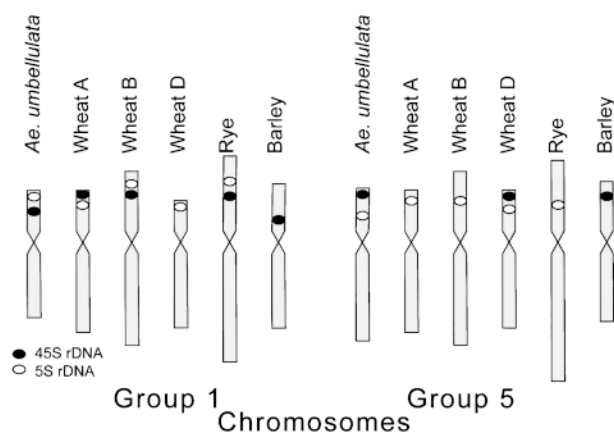
Figure 3. Diagrammatic Representation of 45S and 5S rDNA on Chromosomes from Groups 1 and 5 in Various Triticeae.

The sites differ in location, size, and order in the six genomes. The variations do not always reflect those found by genetic mapping of other molecular markers that may show greater conserved synteny among the species. No inversions have been detected among wheat, rye, and barley, although the rDNA genes show both orders on group 1 chromosomes. The rDNA sites provide useful markers for following the evolution of cereal chromosomes and also assist in the identification of individual chromosomes.

to the centromeres. This function is highly conserved, and in most species of plants and animals, the centromeres are regions of the chromosomes defined cytologically by a primary constriction. A few species, such as the sedge Luzula and the nematode worm Caenorhabditis (*C. elegans* Sequencing Consortium, 1998), have holocentric chromosomes such that microtubules attach throughout the length of the chromosome. The best-characterized centromeres are in the budding yeast *Saccharomyces cerevisiae* (see Clarke, 1990), where a functional centromere is contained within a 125-bp sequence characterized by three centromere DNA elements (CDEI, 8 bp; CDEII, ~80 bp; and CDEIII, 26 bp, where even a single nucleotide change may alter function). Nevertheless, yeast is not a good model for centromere function in plants and animals, in which the DNA at the centromere often, but by no means always, consists of a tandemly repeated sequence. A considerable fraction of the genomic DNA can in fact be represented by the centromere-associated repeats: 0.3% of the human genome is represented by the α satellite, and 3% of the Arabidopsis genome consists of the 180-bp centromeric repeat (Murata et al., 1994).

Despite insightful analyses of the structure and proteins associated with the centromere (Pluta et al., 1990), comprehensive information about centromeric DNA sequences is lacking. In mammals, key sequences are under study (Craig et al., 1999), and many but not all authors regard the tan-

demly repeated sequences as playing a key role in centromere function and chromosome segregation (Kipling and Warburton, 1997; Tyler-Smith et al., 1998). Such sequences have been isolated from many plants and localized to the centromeres by in situ hybridization. The major 180-bp satellite sequence in Arabidopsis is located at the centromeres of all five chromosome pairs (Maluszynska and Heslop-Harrison, 1991), although several other repetitive DNA sequences have also been located in this region (Brandes et al., 1997b; Fransz et al., 1998, 2000). Harrington et al. (1997) have synthesized human microchromosomes from synthetic arrays of the α-satellite DNA. Many of the tandemly repeated sequences, whether in Arabidopsis (Heslop-Harrison et al., 1999), rice (Aragon-Alcaide et al., 1996; Nonomura and Kurata, 1999), millet (Kamm et al., 1994), or animals, include a 17-bp motif that might act as a binding site for centromeric protein B (CENP-B; see Heslop-Harrison et al., 1999). In cereals, retrotransposon-like repeated elements have been documented at the centromeric regions, and several authors have speculated about their role in karyotype evolution and centromere function (Miller et al., 1998; Presting et al., 1998; Ananiev et al., 1999; see also largely homologous sequences reported by Aragon-Alcaide et al. [1996] and Jiang et al. [1996]).

A combination of approaches is under way to elaborate centromere structure in Arabidopsis. Detailed analysis of the 180-bp repeat units indicates that there are variants localized in particular chromosomes (Figures 1 and 2; see also Heslop-Harrison et al., 1999). Sequencing of nearly 2 Mb within the genetically defined centromere has revealed a few recognizable genes and a high density and diverse range of vestigial and presumably inactive mobile elements (Lin et al., 1999). Copenhaver et al. (1999) have used the sequence data from chromosomes 2 and 4 in combination with accurate genetic mapping to define DNA sequences responsible for centromere function. The centromeres consist of a central, repetitive core, flanked by moderately repetitive DNA that has a low rate of recombination, which in turn is flanked by regions with mobile elements and normal recombination rates. Because some repeats are even more abundant in extracentromeric DNA, the repeats alone are probably not sufficient for centromere function (Copenhaver et al., 1999).

Transposable Elements and Retroelements

Retroelements (class I transposable elements) are discrete components of the plant nuclear genome that replicate and reinsert at multiple sites in a complex process that involves activation of excision, DNA-dependent RNA transcription, translation of the RNA into functional proteins, RNA-dependent DNA synthesis (reverse transcription), and reintegration of newly generated retroelement copies into the genome (reviewed in Kumar and Bennetzen, 1999). Major classes of retroelements include LINEs, SINEs, *copia*- and *gypsy*-like elements, and retroviruses (Hull and Covey, 1996; Kumar,

1998; Harper et al., 1999; Jakowitsch et al., 1999; Kumar and Bennetzen, 1999; Schmidt, 1999). Retroelements, typically including two or three open reading frames extending over 5 kb, tend to be highly amplified and frequently represent half of the nuclear DNA (Pearce et al., 1996; SanMiguel et al., 1996; Smit, 1996). Retroelements have been found in all plants investigated and are very heterogeneous (Flavell et al., 1992), suggesting that they are an ancient component of genomes. They are generally dispersed over plant chromosomes, consistent with their mode of amplification, but may associate with particular genomic regions (Figure 2). Most frequently, the rDNA and centromeric regions, consisting of tandemly repeated DNA elements, show a lower proportion of *gypsy*- and *copia*-like retroelements than do other regions (Kamm et al., 1996; Heslop-Harrison et al., 1997; Kubis et al., 1998a; Schmidt, 1999). It is hypothesized that retroelements are more abundant around the centromeres of Arabidopsis chromosomes so as to limit the disruption of genes (Figure 2; Brandes et al., 1997a). Relatively little is known about the chromosomal organization of LINEs (Kubis et al., 1998b).

As they insert themselves into the genome, retroelements act as mutagenic agents, thereby providing a putative source of biodiversity (Hirochika et al., 1996; Heslop-Harrison et al., 1997; Ellis et al., 1998; Flavell et al., 1998) and serving as markers of diversity. Regulatory mechanisms may act to protect genomes from insertional mutagenesis (Lucas et al., 1995), and it has been suggested that transgene-induced gene silencing reflects mechanisms aiming to prevent genome invasion by retroelements. Plant retrotransposon activity can be regulated at any step of the replication cycle, including transcription, translation, reverse transcription, nuclear import, and integration. Along with DNA (class II) transposable elements and other elements such as miniature inverted tandem elements (MITES; Wessler et al., 1995; Casacuberta et al., 1998), insertion of retrotransposon elements can inactivate or alter gene function (Wessler et al., 1995). Indeed, transposition is estimated to account for 80% of the mutations detected in Drosophila (Capy, 1998). Transposons can excise, partially or completely restoring gene function, and can also lead to chromosome rearrangements such as inversions or translocations. Transposable elements can also act to move elements such as exons and promoters into existing sequences so as to create new gene functions and contribute to evolution (Plasterck, 1998; Moran et al., 1999). Indeed, retroelements are activated under stress conditions (Wessler, 1996; Grandbastien, 1998; Kumar and Bennetzen, 1999; Walbot, 1999). Alternative splicing of genes caused by transposable elements has been shown in maize (Bureau and Wessler, 1994a, 1994b). Methylation of retroelements can also affect adjacent sequences and lead to transcriptional repression (Yoder et al., 1997; Goubely et al., 1999).

The sequences of degenerate and potentially active retroelements give valuable data about genome evolution and phylogenetic relationships (Figure 4). In three species in the

Vicia genus, *copia* retroelement copy number varies from 1000 to 1,000,000, with more sequence heterogeneity being present in species with higher copy number (Pearce et al., 1996). Although in part due to random mutation of the high number of copies present in most plant genomes, sequence variability is often nonuniformly distributed along the retroelement: regulatory regions (including the long terminal repeats of *copia* elements) can evolve faster than coding regions, perhaps enabling elements to coexist with their host genomes without detriment (Vernhettes et al., 1998). Although retroelement amplification leads to large genomes (Bennetzen and Kellogg, 1997), it is probable that retroelement turnover and loss can occur in a directed manner (Tatout et al., 1998), leading to different retroelement compositions between species. For example, chromosome sets in the cultivated hexaploid oat, *Avena sativa*, can be discriminated by the presence of retroelement families (Katsiotis et al., 1996).

Simple Sequence Repeats (Microsatellites)

Runs of single nucleotides or motifs of up to ~5 bp, described as microsatellites or simple sequence repeats (SSRs), are ubiquitous elements of eukaryotic genomes (Tautz and Renz, 1984). Genetic mapping using microsatellites as markers involves amplification of repeat arrays by
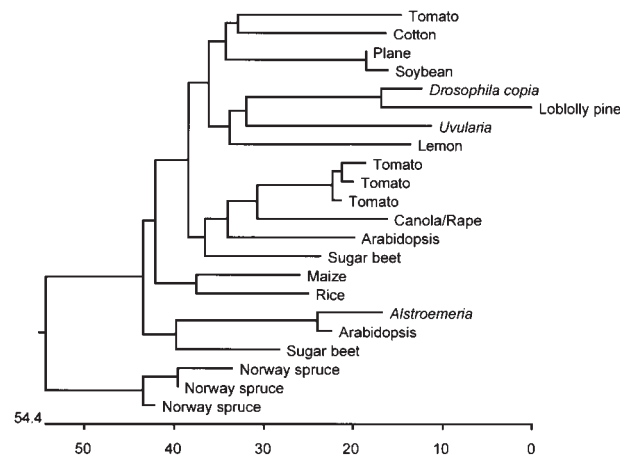

Figure 4. A Phylogenetic Tree (Clustal Method) According to Repeat Sequences.

Fourteen plant species and Drosophila are arranged in accordance with genomic representation of *copia* group retroelements (Genbank EMBL database). The *copia* elements are dispersed along the chromosomes (see Brandes et al., 1997a; see also Figure 2B), consistent with their mode of amplification through an RNA intermediate. Units (bottom) indicate the number of substitution events over ~260 bp.

the polymerase chain reaction with primers flanking the arrays. SSRs also provide highly informative and polymorphic markers for plant, fungal, and animal fingerprinting (Weising et al., 1991). Synthetic oligonucleotide SSRs have been used for in situ hybridization to chromosomes, revealing that microsatellite sequences vary widely with regard to genomic organization, raising implications for amplification and dispersion mechanisms and hence evolution. In some cases, synthetic SSRs have been used to detect sites within previously characterized repeat motifs. For example, a tandemly repeated motif near the centromeres of all 16 pairs of sugar beet chromosomes includes an $(AC)_8$ motif (Schmidt and Heslop-Harrison, 1996), and the polypurine motif $(GAA)_7$ has been correlated with the positions of C-bands in barley (Pedersen and Linde-Laursen, 1994). Notably, although conventional staining systems give very different chromosome bands in wheat and rye, the hybridization pattern of the motif GACA with some 40 amplified sites is very similar in the two species, suggesting that the pattern was established before their evolutionary separation (Cuadrado and Schwarzacher, 1998). In the human genome, changes in copy number of different microsatellite classes may occur through interallelic replication slippage of AT-rich sequences or complex, conversion-like events of GC-rich regions, with recombination in DNA flanking the repeat array (Bois and Jeffreys, 1999).

Tandem Arrays of Repetitive DNA

Many repetitive sequence motifs occur as tandem repeats at a number of discrete sites—typically between one and 30—in the genome. Using in situ hybridization, these tandem repeats can provide useful markers for chromosome identification, and their presence and distribution can reveal evolutionary changes (Figure 2; Kubis et al., 1997). Both the site distribution and sequence of tandemly repeated sequences may show polymorphism between species and accessions of a species. However, the evolution of tandem repeats does not show characteristics of a "molecular clock" with a constant mutation rate. All evidence points to its occurrence in bursts or evolutionary waves, perhaps occurring during periods of rapid speciation or stress.

In many species, the distribution of different repetitive DNA sequences closely follows their taxonomic relationships: eight different sequences isolated from *Beta* spp can be used to elucidate the relationships between the four related sections of the genus (Schmidt and Heslop-Harrison, 1994). In contrast, taxonomy within the genus *Crocus* shows little correlation with the distribution of repetitive sequence, reflecting not only a disparity between taxonomy and actual phylogeny but also the explosive speciation occurring at one evolutionary period (Frello and Heslop-Harrison, 2000).

A family of repetitive sequences originally isolated from rye, named pSc119.2 (Bedbrook et al., 1980), is abundant in all species of the tribe Triticeae, and even in related tribes such as Avenae, but absent from cultivated barley and close relatives. Because it is likely that the sequence was present in the common ancestor of the Triticeae tribe, its absence from barley implies that high-copy sequences may be superfluous to the genome and again suggests there is no molecular clock to gauge evolution. In rye itself, more distal subtelomeric sequences, pSc200 and pSc250, are relatively species specific (Vershinin et al., 1995) and have presumably evolved more recently.

Tandem repeats are normally regarded as transcriptionally silent (Radic et al., 1987), although a significant proportion of RNA in rice has been shown to represent a particular subtelomeric tandem repeat (Wu et al., 1994). It is possible that such RNA is due to read-through transcription in which a stop codon is ignored, which might occur more frequently under stressful conditions.

Frequently, unequal crossover and recombination of chromosome strands within the tandem arrays are considered to be involved in the evolution and amplification of repeat units (Dover, 1982; Charlesworth et al., 1994). McAllister and Werren (1999) have presented experimental evidence for the unequal crossover model and also suggest how turnover of repeats allows migration of retroelements toward the ends of arrays. In yeast, Paques et al. (1998) conclude that the expansion and contraction mechanisms for tandem arrays have their origin in DNA repair rather than genome replication mechanisms. It is also evident that genome-scanning mechanisms can homogenize different units of a tandem repeat, making all sequences identical. Much work showing homogenization has been performed on rDNA repeat units, and it is possible that similar mechanisms may act in other repeats. Schlotterer and Tautz (1992) have shown that intrachromosomal homogenization occurs rapidly in Drosophila rDNA, whereas interchromosomal homogenization occurs at a slower rate. In cotton, a tetraploid species, Wendel et al. (1995) have shown that the rDNA has become homogenized to resemble the variant found in only one of the ancestral *Gossypium* spp.

DNA SEQUENCE IN THE CHROMOSOME

Within the nucleus, DNA is modified by the addition of methyl groups, and most DNA is wrapped around histone proteins, forming nucleosomes and the 30-nm fiber as the fundamental structural subunit of chromosomes (Manuelidis and Chen, 1990; Wolffe, 1995). Higher levels of packaging, often very dynamic, result in chromatin fibers such that varying chromatin density is seen within the three dimensions of the interphase nuclei (Figure 5) as metaphase chromosomes appear. The packing of the genomic DNA can directly affect aspects of RNA transcription, DNA replication, recombination, DNA repair, and chromosome segregation (Cremer et al., 1993; Heslop-Harrison et al., 1993).
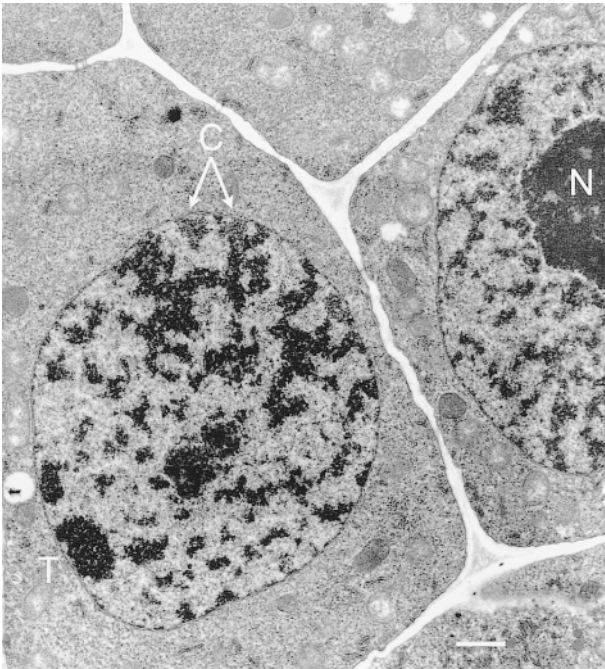
Figure 5. Nuclear Architecture of Rye Seedling Root-Tip Cells.

Chromatin is visible as electron-dense material in this electron micrograph. The nucleolus (N) is seen within one nucleus, and centromeric (C) and telomeric (T) chromatin is visible as large, electron-dense, condensed blocks of heterochromatin adjacent to the nuclear envelope. The pole of the nucleus near the centromeres (at C) contains a greater proportion of chromatin (dark) than the pole near the telomere (at T), consistent with the light micrographs in Figure 2E. Bar = 1 μm.

Methylation

In plants, as well as in most prokaryotes and animals (except for Drosophila), modification of DNA by cytosine methylation is extensive (Finnegan et al., 1996): ~80% of cytosines in CG dinucleotides are modified (Gruenbaum et al., 1981). Plants, like animals, may contain unmethylated CG-rich regions (CpG islands) related to transcriptionally active genes (Antequera and Bird, 1993), and extensive evidence suggests that methylation is a mechanism for regulating gene expression. Numerous reports have correlated hypermethylation near genes, or in gene promoters, with reduced levels of gene expression (Barlow, 1993; Razin and Cedar, 1993; Sardana et al., 1993; Neves et al., 1995; Finnegan et al., 1996). Repression occurs at the level of transcription initiation (Tate and Bird, 1993), although methylation does not seem to repress the activity of all genes, including those borne by transposons (Martienssen, 1998). Many DNA methylation patterns are established during ontogeny and may remain stable through later development (Jahner and Jaenisch, 1984; Razin and Cedar, 1993; Neves

et al., 1997). Studies of floral homeotic mutants (Finnegan et al., 1996; Ronemus et al., 1996) suggest a direct correlation between DNA methylation and normal regulation of developmentally important genes (Jacobsen and Meyerowitz, 1997).

In animals, most methylation seems to occur at symmetrical sites in the DNA molecule, where the nucleotide combinations CG or CNG (N is any nucleotide) occur on both DNA strands. After DNA replication, methylation patterns are copied by maintenance methylases that respond to the methylation status of diagonally opposite Cs in the newly replicated DNA strand. In plants, it appears that methylation does not always occur at symmetrical positions (Fulnecek et al., 1998; Goubely et al., 1999); methylation sites must be established de novo after each replication cycle, perhaps by a DNA–DNA (Matzke et al., 1994) or RNA–DNA (Pelissier et al., 1999) pairing process. Wassenegger et al. (1994) indicate that overexpressed mRNAs might direct sequence-specific de novo methylation of the DNA template and thus regulate gene activity. Such mechanisms may be involved in gene-silencing phenomena.

DNA methylation usually represents a terminal stage of differentiation but may be modulated, as is apparent by the activation in tissue culture of previously inactive retroelements (Grandbastien, 1998). Some methylation patterns change during plant development, particularly through meiosis (Silva et al., 1995) and embryogenesis (Castilho et al., 1999). Progressive reduction in methylation levels can occur upon DNA replication so as to result in hemimethylated and subsequently unmethylated DNA in daughter nuclei (Matzke et al., 1989; Kilby et al., 1992; Jeddeloh et al., 1998). For the experimental reduction of DNA methylation, the cytosine analog 5-azacytidine, with a nitrogen atom rather than carbon atom at the 5-position of the pyrimidine ring, has revealed that reduced methylation of tandem DNA repeats in tobacco is maintained during protoplasting and plant regeneration (Bezdek et al., 1991; Koukalova et al., 1994).

Henikoff and Comai (1998) have found that Arabidopsis, like mouse and pea, has multiple methyltransferase specificities, probably resulting from multiple genes, and certain specificities may be tissue specific. In pea, methylase activities that recognize CG and CWG (where W is A or T) probably arise from the post-translational modification of a single gene product (Pradhan and Adams, 1995; Pradhan et al., 1995). Different enzymes are most likely to be involved in methylation of asymmetrical sites as opposed to maintenance of methylation of symmetrical sites (Goubely et al., 1999). In mammals, a CG demethylase has been identified (Bhattacharya et al., 1999), revealing a new mechanism of gene regulation presumably also present in plants.

Smith (1998) has suggested that the function of DNA methyltransferases and DNA methylation is in maintenance of eukaryotic chromosome stability. DNA methyltransferases participate in DNA repair complexes and also stabilize nucleoprotein assemblies required in the inactivation and imprinting of chromosomes. Methyltransferases may incorporate a
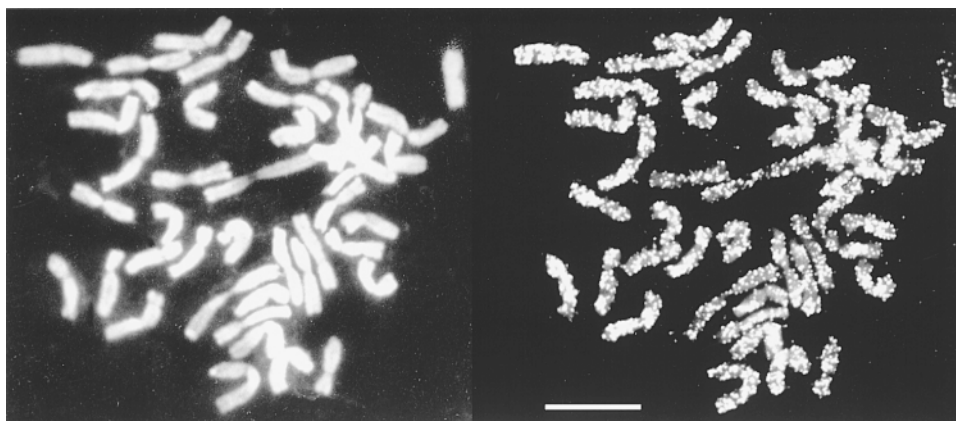
Figure 6. Antibody Labeling of Methylcytosine in Metaphase Chromosomes of Triticale.

The chromosomes of the wheat–rye hybrid ($2n = 6x = 42$) are counterstained with DAPI (left). The antibody-labeled chromosomes (right) show widespread, punctate labeling with many gaps and regions of reduced labeling. See Castilho et al. (1999) for more details. Bar = 10 μm.

chromodomain, a protein module that mediates interactions between key chromatin proteins (Henikoff and Comai, 1998). Antibodies to methylcytosine have shown that different regions of chromosomes have different levels of methylation both in humans (De Capoa et al., 1995) and in plants (Figure 6; Frediani et al., 1996; Oakeley et al., 1997; Siroky et al., 1998; Castilho et al., 1999).

Structure and Packaging of Linear DNA into Chromosomes

The DNA double helix is wrapped around histone core particles, with ~146 bp of DNA forming the two turns around each nucleosome. Nucleosomes are connected by linker DNA, typically 20 to 35 bp long. Using micrococcal nuclease to cleave DNA in the linker region between nucleosomal core particles (Figure 7), it has become clear that chromatin higher-order structures and nucleosomal organization are not homogeneous along chromosomes (Fischer et al., 1994; Wolffe and Pruss, 1996) and that the dynamic chromatin structure found in animal systems applies also to plants. For example, Vershinin and Heslop-Harrison (1998) have shown small but significant variation in the nucleosomal organization and linker DNA length between telomeric DNA and various repetitive DNA sequence motifs in the bulk chromatin of rye, wheat, and their relatives. Furthermore, differences in linker DNA length and the sensitivity of cereal chromatin to micrococcal nuclease were observed in rye and wheat despite their relatively close taxonomic relationships.

Repetitive sequences, in particular tandem arrays, probably play a key role in stabilizing DNA packaging and higher-order chromatin condensation. Repetitive DNA motifs usually show a strictly defined arrangement (phasing) around nucleosomes. Gazdova et al. (1995) determined the position

of the nucleosomal core next to a 10- to 11-bp AT track in a monomer of a tobacco tandem repeat, and Vershinin and Heslop-Harrison (1998) showed the defined phasing of tandem repeat motifs of 120, 360, and 550 bp. Nucleotide base stacking and twisting angles have been derived by Calladine et al. (1988) and provide the basis for predicting natural curvature of DNA molecules. Radic et al. (1987) hypothesized that bends in satellite DNA represent an essential structural signal for complete heterochromatin condensation. Repeated tracts of four to six adenines in phase with the helix produce bends, and bent DNA preferentially assembles into nucleosomes. After the packing of the repeats, the small proportion of single-copy DNA, regardless of its natural curvature preferences, can be fitted.

The frequent occurrence of sequence motifs ~180 bp long, or multiples of this length, indicates that the natural fit of the DNA molecule to the nucleosome core may be an important feature with respect to selection of lengths of repetitive DNA motifs. Breakage is observed in a small percentage of metaphase chromosomes and is often enhanced in divisions in interspecific hybrids: one might speculate that the poor fit of linkers between nucleosomes increases the breakage frequency, and repair mechanisms may be less efficient in the hybrid background.

Chromatin Remodeling and Histone Acetylation

Along with DNA methylation, chromatin remodeling and histone acetylation have been implicated in the modification of gene transcription (Martienssen and Henikoff, 1999). Histone acetylation per se may both change the relative positions of nucleosomes and influence the structure of chromatin (Turner, 1991). Chromatin remodeling involves specific enzymes affecting nucleosome structure and posi-

tioning along the DNA molecule (Cairns, 1998). Tazi and Bird (1990) have suggested that DNA methylation silences transcription through assembly of a repressive nucleosomal array. Wade et al. (1997) have suggested that nucleosome positioning may be critical in regulating the rate of transcription by modulating access and procession rate of the polymerase complex. Furthermore, methylation could suppress gene expression through an indirect mechanism affecting chromatin structure (Kass et al., 1997; Bergman and Mostoslavsky, 1998). Methylation may also mediate interaction between transposon sequences and chromatin factors, which conceal the sequences from the rest of the genome (Kass et al., 1997). The *DDM1* gene in Arabidopsis causes rapid hypomethylation of repetitive DNA (Vongs et al., 1993; Kakutani et al., 1995, 1996), followed by hypomethylation of genes over many plant generations (Jeddeloh et al., 1998). Sequence analysis indicates that the DDM1 protein does not have a direct role as a methyltransferase but rather modifies the accessibility of chromatin to methylation (Jeddeloh et al., 1999). The *DDM1* gene is similar to sequences from animals and fungi, which act to modify or disrupt protein–DNA

interactions of multiprotein complexes that include the DDM1-like component (Jeddeloh et al., 1999). Thus, the DDM1 protein probably functions in the DNA methylation system by affecting chromatin structure, perhaps by directing certain sequences to the methylation machinery or by modulating nucleosome remodeling.

Chromatin remodeling might have an ancient origin in the modulation of genome organization and may be a general requirement for replication of condensed, inactive regions of the genome. In vertebrates, it is well known that methylation of CG dinucleotides correlates with alterations in chromatin structure and gene silencing (Antequera et al., 1990; Antequera and Bird, 1993). Prymakowska-Bosak et al. (1996) have argued that genes involved in basal cellular functions are probably influenced relatively little by alterations in chromatin structure, whereas genes involved in specific developmental programs are likely to be regulated by factors related to chromatin constitution. The classic phenomenon of position-effect variegation in Drosophila occurs as chromosomes become heterochromatic (see Henikoff et al., 1993), and related phenomena have been associated with pea (Kass and Adams, 1993; Sabl and Henikoff, 1996). Johnson et al. (1995) have established interrelationships between gene transcription and methylation and DNA packaging.

Local chromatin structure and its modification in early meiosis are important in the positioning and frequency of meiotic double-strand breaks in DNA that enable recombination in yeast (Ohta et al., 1994; Wu and Lichten, 1994). Earlier studies (Chandley and McBeath, 1987; Raman and Nanda, 1986) had also discussed that the regions of the human genome where the chromatin undergoes conformational changes from mitosis to meiosis could encompass recombinational hot spots. The lack of condensation of early replicating chromosomal segments during premeiotic interphase could be a prerequisite for crossover at pachytene.
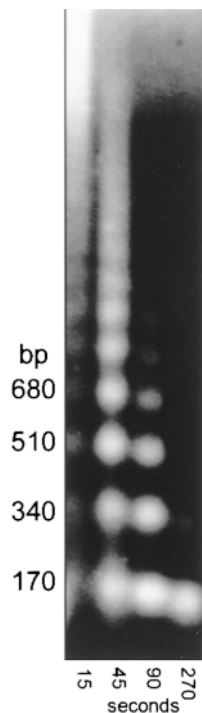
THE THREE-DIMENSIONAL NUCLEUS

Genome Architecture

Genome architecture refers to the structural organization of the plant genome in the three-dimensional nucleus and can be extended to describe its dynamics and the relationship between structure and function. Cockell and Gasser (1999) concur with the emerging view that gene regulation cannot be fully explained by linear, two-dimensional models involving merely the binding of factors to regulatory elements. It has been widely suggested that nuclear architecture is related directly to the control of gene expression and that the multiple levels of organization of the chromatin provide functional regulation of DNA behavior. DNA packing and unpacking, replication, repair, mutation, and transcription are all regarded as cell-type specific aspects of a dynamic architecture. The scientific literature is now full of direct and

Figure 7. Nucleosomal Structure of Rye Chromatin.

Micrococcal nuclease digestion of extracted chromatin followed by size separation by agarose-gel electrophoresis and probing with a telomeric sequence results in a ladder of discrete bands that changes with the course of the digestion reaction. The bands differ in increments of ∼170 bp. See Vershinin and Heslop-Harrison (1998) for more details.

indirect acknowledgment of the importance of nuclear architecture, including unpredictable positional effects and rearrangements. However, much of the literature about nuclear architecture and chromatin structure is based on mammalian, insect, or yeast models, often using cultured or model cell types such as fibroblasts or Drosophila polytene nuclei.

Electron micrographs show that DNA is largely condensed in plant interphase nuclei and that this condensed interphase chromatin is similar in appearance to chromosomes (Figure 5); the chromatin of cereals is largely condensed even in interphase nuclei (Muller et al., 1980). Measurement of chromosome volume, although inaccurate because of edge effects, indicates that volumes are similar in G2 interphase nuclei compared with mitotic chromosomes (Heslop-Harrison et al., 1988). Thus, little nuclear DNA is truly "decondensed."

Packaging of Nuclear DNA

The traditional twentieth-century view of the nucleus as an unstructured jumble of spaghetti-like chromatin fibers is largely discounted, and most researchers agree that there are intranuclear frameworks that provide the dynamic genome with functional organization. Various levels of intranuclear compartmentalization can be regarded: individual chromosomes (Figure 2), euchromatic and heterochromatic regions, the nucleolus (Figure 5), and regions of active RNA synthesis and processing. Furthermore, telomeres and centromeres may be attached to or closely adjacent to the nuclear envelope (Schwarzacher and Heslop-Harrison, 1990; Rawlins et al., 1991) and occupy defined parts (poles) of the nucleus in many species (Rabl, 1885; Cremer et al., 1982; Anamthawat-Jónsson and Heslop-Harrison, 1990).

Cook (1997) has argued that each chromosome in a haploid set has a unique array of transcription units strung along its length and that chromatin fibers will therefore be folded into unique arrays of loops, with homologs sharing similar arrays. At meiosis, homologous chromosomes come together; this occurs when they are transcriptionally active, so that pairing may be an inevitable consequence of the transcription of partially condensed chromosomes (Cook, 1997). Similarly, Karpen et al. (1996) proposed that DNA–protein structures inherent to heterochromatin in Drosophila could produce a self-complementary chromosome "landscape" that ensures partner recognition and alignment by "best-fit" mechanisms. Specific coiling patterns that could promote pairing, showing apparent denser and weaker zones presumably reflecting more or less condensed chromatin, were observed at stages before meiotic prophase in the homologous chromosome domains of wheat (Figure 1A in Schwarzacher, 1997).

The existence of a nuclear matrix, or chromosomal skeleton, or both, following models for the cytoskeleton, remains controversial. Numerous papers describe features that appear under conditions that are far from the in vivo situation so that the relevance of such nuclear scaffolds, matrices, cages, and compartments remains questionable. As with any responsive and precisely regulated system, even small changes in hydration, ion concentration, and tonicity during experimentation are certain to have major effects on structure (Jackson and Cook, 1995). There are, moreover, complex controls on traffic between cytoplasm and nucleus (Jackson and Cook, 1995). Most of the major cytoskeletal proteins, including tubulins and actin, have been found within the nucleus, but their function and significance are unclear. It is accepted that the lamins (intermediate filaments) have a key role at the periphery of the nucleus and also extend deep into the nuclear volume. The nuclear pore complex and its dynamics have been worked out in detail (Allen et al., 1998; Goldberg et al., 1999), and it is clear that proteins permeate several microns from the pore complex into the nucleus in yeast, insects, and vertebrates. Advanced microscopic methods and antibody technology show that there are unexpected subnuclear localization patterns of many proteins (including transcription factors) in Arabidopsis and other species.

The higher-order structure of the chromatin fiber and the organization of chromatin domains in the nucleus appear to have a profound influence on gene expression. Good evidence exists that in most interphase nuclei, individual chromosomes occupy discrete domains, but the internal structure of these territories and the relation of their organization to presumptive higher-order functional compartments are difficult to investigate (Heslop-Harrison and Bennett, 1990; Sadoni et al., 1999). However, there is reasonable evidence in mammals that active genes tend to locate on the periphery of the territories, where RNA transcripts are formed. The chromosome domains may be elongated or subspherical, and DNA fibers may stretch away—possibly many microns—from the surface of the domain.

Although perhaps not a model for all aspects of nuclear behavior, incontrovertible evidence for nuclear compartmentalization is provided by the nucleoli. Nucleoli are subspherical compartments in the nucleus; there are no defined boundaries to nucleoli, although their composition is very different from the rest of the nucleus, and they move and fuse during interphase of the cell cycle. Soon after cell division in most species, multiple nucleoli (each originating from one rRNA locus) are the norm, but they often fuse to a smaller number during development of the cell. They are located at different positions within the nucleus, depending on cell type: peripheral and very close to the nuclear envelope in pollen mother cells at early meiotic prophase, but more central in other cell types.

GENOMICS, CHROMOSOMES, EVOLUTION, AND THE NUCLEUS

Genomes evolve at the level of the chromosome, chromosome segment, gene, and DNA sequence. Biotechnologists

and plant breeders aim to control and direct evolution, although limiting the impact of experimentally imposed genome evolution is an objective for the conservation of biodiversity and the environment. As Capy (1998) has stated, studies of the molecular basis of genome evolution are still young, but identification of the many processes in genome evolution, from molecular events to population dynamics, shows the impressive plasticity of the genome and the rapid amplification and fixation of advantageous novelties. An understanding of the functional and genetic bases of the major sources of variation at the genomic level (including retroelements; Kumar and Bennetzen, 1999) will

have important applications. An appreciation of the types of changes that have occurred during species evolution will enable us to understand what can be done in plant breeding with respect to the changing environment.

Chromosome organization has a fundamental influence on processes as diverse as chromosome pairing, segregation, gene organization, and expression and has a direct impact on the aims of plant breeders in understanding genome evolution and genetics. The current model of the chromosome in the nucleus (Figure 8) is very different from that of five years ago. We now have complete sequences of chromosomes, and we can build a picture of the organization of
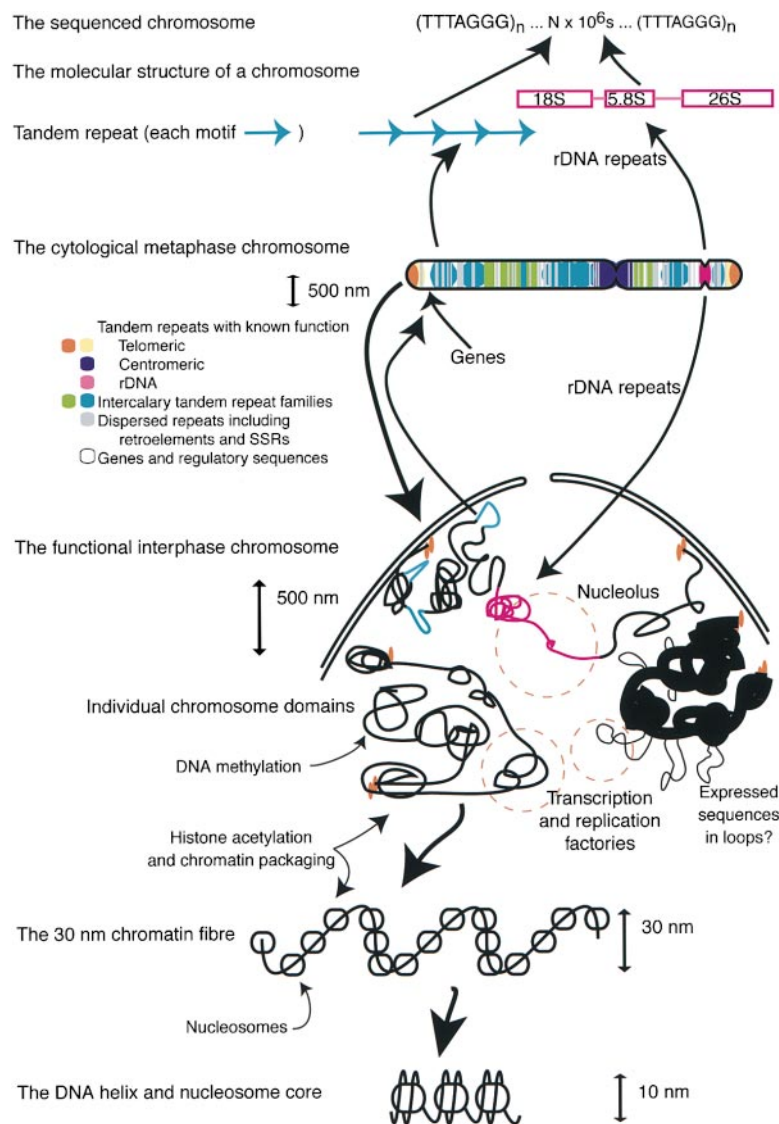

Figure 8. A Model of the Sequence and Structural Organization of Plant Chromosomes.

See the text for details.

the different sequence motif types, each with characteristic locations along the linear structure (Schmidt and Heslop-Harrison, 1998). We now know about tandem repeats in relation to chromosome structure at telomeres and centromeres, about the abundance of intercalary tandem repeats, and in plant species with larger genomes than Arabidopsis, about largely dispersed retroelements that represent 50% of the DNA in the genome. The observed organization may be a consequence of the varying constraints on the evolution of each sequence motif, with many evolutionary changes reflective of the amplification or depletion of sequence motifs at particular genomic locations. The current chromosome model will be helpful for designing strategies for gene cloning, evolutionary studies, gene transfer, and perhaps gene expression.

The functional interphase nucleus is divided into compartments that are dynamic with respect to changes in chromosome structure and locations of genomic components (Figure 8). Subdomains of interphase chromosome territories are related to the different domains of mitotic chromosomes (Zink et al., 1998). We can suggest that there is a proteinaceous framework to the interphase nucleus, with many enzymatic and structural processes involved in chromatin modification and packaging. Together, these structures undoubtedly contribute to the observed and complex interactions between different genes and their regulation and inheritance of expression patterns. Nevertheless, this interphase model is based largely on vertebrate and insect systems; it must be treated with caution because the linearly compartmentalized mammalian chromosome (Holmquist, 1992; Saccone et al., 1993) shows differences in evolution and structure from the plant metaphase chromosome, as revealed by banding and molecular cytogenetics methods (Figure 8). There are parallels between plants and animals in chromatin modification and packaging in the nucleus (Figure 8, bottom), but details of the modulation and the significance of these structures remain to be studied.

Bennetzen et al. (1998) has shown that knowledge of the structure of the (grass) genome is useful in gene discovery and isolation. Comparative analysis of different species over narrow as well as wider taxonomic ranges has been instrumental in understanding genome organization in general. Promising technological developments to investigate the nucleus include in silico analyses, high-throughput DNA and protein studies, digital imaging methods, and in vivo methods to study dynamic chromosome structures and interactions. Such technologies will ensure that advances in understanding the behavior, function, and evolution of the chromosome in the nucleus will come at an increasing rate.

ACKNOWLEDGMENTS

REFERENCES

Ahn, S., Anderson, J.A., Sorrells, M.E., and Tanksley, S.D. (1993). Homoeologous relationships of rice, wheat and maize chromosomes. Mol. Gen. Genet. **241,** 483–490.

Alkhimova, A.G., Heslop-Harrison, J.S., Shchapova, A.I., and Vershinin, A.V. (1999). Rye chromosome variability in wheat–rye addition and substitution lines. Chromosome Res. **7,** 205–212.

Allen, T.D., Rutherford, S.A., Bennion, G.R., Wiese, C., Riepert, S., Kiseleva, E., and Goldberg, M.W. (1998). Three-dimensional surface structure analysis of the nucleus. Methods Cell Biol. **53,** 125–138.

Anamthawat-Jónsson, K., and Heslop-Harrison, J.S. (1990). Centromeres, telomeres, and chromatin in the interphase nucleus of cereals. Caryologia **43,** 205–213.

Ananiev, E.V., Phillips, R.L., and Rines, H.W. (1999). Chromosome-specific molecular organization of maize (*Zea mays* L.) centromeric regions. Proc. Natl. Acad. Sci. USA **95,** 13073–13078.

Antequera, F., and Bird, A. (1993). DNA methylation and CpG islands. In The Chromosome, J.S. Heslop-Harrison and R.B. Flavell, eds (Oxford, UK: Bios), pp. 127–133.

Antequera, F., Boyes, J., and Bird, A. (1990). High levels of *de novo* methylation and altered chromatin structure at CpG islands in cell lines. Cell **62,** 503–514.

Aragon-Alcaide, L., Miller, T., Schwarzacher, T., Reader, S., and Moore, G. (1996). A cereal centromeric sequence. Chromosoma **105,** 261–268.

Barlow, D.P. (1993). Methylation and imprinting—From host defence to gene regulation. Science **260,** 309–310.

Bedbrook, J.R., Jones, J., O'Dell, M., Thompson, R.D., and Flavell, R.B. (1980). A molecular description of telomeric heterochomatin in *Secale* species. Cell **19,** 545–560.

Bennett, M.D., Leitch, I.J., and Hanson, L. (1998). DNA amounts in two samples of angiosperm weeds. Ann. Bot. **82,** 121–134.

Bennetzen, J.L., and Kellogg, E.A. (1997). Do plants have a one-way ticket to genomic obesity? Plant Cell **9,** 1509–1514.

Bennetzen, J.L., SanMiguel, P., Chen, M.S., Tikhonov, A., Francki, M., and Avramova, Z. (1998). Grass genomes. Proc. Natl. Acad. Sci. USA **95,** 1975–1978.

Bergman, Y., and Mostoslavsky, R. (1998). DNA demethylation: Turning genes on. Biol. Chem. **379,** 401–407.

Bezdek, M., Koukalova, B., Brzobohaty, B., and Vyskot, B. (1991). 5-Azacytidine-induced hypomethylation of tobacco HRS60 tandem DNA repeats in tissue culture. Planta **184,** 487–490.

Bhattacharya, S.K., Ramchandani, S., Cervoni, N., and Szyf, M. (1999). A mammalian protein with specific demethylase activity for mCpG DNA. Nature **397,** 579–583.

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

Brandes, A., Heslop-Harrison, J.S., Kamm, A., Kubis, S., Doudrick, R.L., and Schmidt, T. (1997a). Comparative analysis of the chromosomal and genomic organization of *Ty1-copia*-like retrotransposons in pteridophytes, gymnosperms and angiosperms. Plant Mol. Biol. **33**, 11–21.

Brandes, A., Thompson, H., Dean, C., and Heslop-Harrison, J.S. (1997b). Multiple repetitive DNA sequences in the paracentromeric regions of *Arabidopsis thaliana* L. Chromosome Res. **5**, 238–246.

Bureau, T.E., and Wessler, S.R. (1994a). Transduction of a cellular gene by a plant retroelement. Cell **77**, 479–480.

Bureau, T.E., and Wessler, S.R. (1994b). Mobile inverted–repeat elements of the Tourist family are associated with the genes of many cereal grasses. Proc. Natl. Acad. Sci. USA **91**, 1411–1415.

Cairns, B.R. (1998). Chromatin remodelling machines: Similar motors, ulterior motives. Trends Biochem. Sci. **23**, 20–25.

Calladine, C.R., Drew, H.R., and McCall, M.J. (1988). The intrinsic curvature of DNA in solution. J. Mol. Biol. **201**, 127–137.

Capy, P. (1998). A plastic genome. Nature **396**, 522–523.

Casacuberta, E., Casacuberta, J.M., Puigdomènech, P., and Monfort, A. (1998). Presence of miniature inverted-repeat transposable elements (MITEs) in the genome of *Arabidopsis thaliana*: Characterisation of the Emigrant family of elements. Plant J. **16**, 79–85.

Castilho, A., and Heslop-Harrison, J.S. (1995). Physical mapping of 5S and 18S–25S rDNA and repetitive DNA sequences in *Aegilops umbellulata*. Genome **38**, 91–96.

Castilho, A., Neves, N., Rufini-Castiglione, M., Viegas, W., and Heslop-Harrison, J.S. (1999). 5-Methylcytosine distribution and genome organization in Triticale before and after treatment with 5-azacytidine. J. Cell Sci. **112**, 4397–4404.

***C. elegans* Sequencing Consortium.** (1998). Genome sequence of the nematode *C. elegans*: A platform for investigating biology. Science **282**, 2012–2018.

Chandley, A.C., and McBeath, S. (1987). DNase-I hypersensitive sites along the XY bivalent at meiosis in man include the XPYP pairing region. Cytogenet. Cell Genet. **44**, 22–31.

Charlesworth, B., Sniegowski, P., and Stephan, W. (1994). The evolutionary dynamics of repetitive DNA in eukaryotes. Nature **371**, 215–220.

Clarke, L. (1990). Centromeres of budding and fission yeasts. Trends Genet. **6**, 150–154.

Cockell, M., and Gasser, S.M. (1999). Nuclear compartment and gene regulation. Curr. Opin. Genet. Dev. **9**, 199–205.

Cook, P.R. (1997). The transcriptional basis of chromosome pairing. J. Cell Sci. **8**, 1033–1040.

Copenhaver, G.P., and Pikaard, C.S. (1996). RFLP and physical mapping with an rDNA-specific endonuclease reveals that nucleolus organizer regions of *Arabidopsis thaliana* adjoin the telomeres on chromosomes 2 and 4. Plant J. **9**, 259–272.

Copenhaver, G.P., et al. (1999). Genetic definition and sequence analysis of Arabidopsis centromeres. Science **286**, 2468–2474.

Craig, J.M., Earnshaw, W.C., and Vagnarelli, P. (1999). Mammalian centromeres: DNA sequence, protein composition, and role in cell cycle progression. Exp. Cell Res. **246**, 249–262.

Cremer, T., Cremer, C., Baumann, H., Luedtke, E.K., Sperling, K., Teuber, V., and Zorn, C. (1982). Rabl's model of the interphase chromosome arrangement tested in Chinese hamster cells by premature chromosome condensation and laser-UV-microbeam experiments. Hum. Genet. **60**, 46–56.

Cremer, T., et al. (1993). Role of chromosome territories in the functional compartmentalization of the cell nucleus. Cold Spring Harbor Symp. Quant. Biol. **58**, 777–791.

Cuadrado, A., and Schwarzacher, T. (1998). The chromosomal organization of simple sequence repeats in wheat and rye genomes. Chromosoma **107**, 587–594.

Darlington, C.D. (1931). The analysis of chromosome pairing in *Triticum* hybrids. Cytologia **3**, 21–25.

Darlington, C.D., and LaCour, L. (1942). The Handling of Chromosomes. (London: Allen).

de Bustos, A., Cuadrado, A., Soler, C., and Jouve, N. (1996). Physical mapping of repetitive DNA sequences and 5S and 18S–26S rDNA in five wild species of the genus Hordeum. Chromosome Res. **4**, 491–499.

De Capoa, A., Menendez, F., Poggesi, I., Giancitti, P., Grapelli, C., Marotta, M., Niveleau, A., Reynaud, C., Archidiacono, N., and Rocchi, M. (1995). Labelling by anti 5-MeC antibodies as a measure of the methylation status of human constitutive heterochromatin. Chromosome Res. **3** (suppl. 1), 45.

Devos, K.M., and Gale, M.D. (1993). Extended genetic maps of the homoeologous group 3 chromosomes of wheat, rye and barley. Theor. Appl. Genet. **85**, 649–652.

Devos, K.M., and Gale, M.D. (1997). Comparative genetics in the grasses. Plant Mol. Biol. **35**, 3–15.

Dioscorides, P. (80). Ridis anazarbei, de medicinali materia. Athens.

Doudrick, R.L., Heslop-Harrison, J.S., Nelson, C.D., Schmidt, T., Nance, W.L., and Schwarzacher, T. (1995). Karyotype of Slash Pine (*Pinus elliottii* var. *elliottii*) using patterns of fluorescence *in situ* hybridization and fluorochrome banding. J. Hered. **86**, 289–296.

Dover, G.A. (1982). Molecular drive: A cohesive mode of species evolution. Nature **299**, 111–117.

Dunham, I., et al. (1999). The DNA sequence of human chromosome 22. Nature **402**, 489–495.

Ellis, T.H.N., Poyser, S.J., Knox, M.R., Vershinin, A.V., and Ambrose, M.J. (1998). Ty1-*copia* class retrotransposon insertion site polymorphism for linkage and diversity analysis in pea. Mol. Gen. Genet. **260**, 9–19.

Fajkus, J., Kovarik, A., Kralovics, R., and Bezdek, M. (1995). Organization of telomeric and subtelomeric chromatin in the higher plant *Nicotiana tabacum*. Mol. Gen. Genet. **247**, 633–638.

Finnegan, E.J., Peacock, W.J., and Dennis, E.S. (1996). Reduced DNA methylation in *Arabidopsis thaliana* results in abnormal plant development. Proc. Natl. Acad. Sci. USA **93**, 8449–8454.

Fischer, T.C., Groner, S., Zentgraf, U., and Hemleben, V. (1994). Evidence for nucleosomal phasing and a novel protein specifically binding to cucumber satellite DNA. Z. Naturforsch. **49c**, 79–86.

Flavell, A.J., Smith, D.B., and Kumar, A. (1992). Extreme heterogeneity of *Ty1-copia* group retrotransposons in plants. Mol. Gen. Genet. **231**, 233–242.

Flavell, A.J., Knox, M., Pearce, S.R., and Ellis, T.H.N. (1998). Retrotransposon-based insertion polymorphisms (RBIP) for high throughput marker analysis. Plant J. **16**, 643–650.

Fransz, P., Armstrong, S., Alonso-Blanco, C., Fischer, T.C., Torres-Ruiz, R.A., and Jones, G. (1998). Cytogenetics for the model system *Arabidopsis thaliana*. Plant J. **13,** 867–876.

Fransz, P.F., Armstrong, S., de Jong, J.H., Parnell, L.D., van Drunen, G., Dean, C., Zabel, P., Bisseling, T., Jones, G.H. (2000). Integrated cytogenetic map of chromosome arm 4S of *A. thaliana*: Structural organization of heterochromatic knob and centromere region. Cell **100,** 367–376.

Frediani, M., Giraldi, E., and Ruffini Castiglione, M. (1996). Distribution of 5-methylcytosine rich regions in the metaphase chromosomes of *Vicia faba*. Chromosome Res. **4,** 141–146.

Frello, S., and Heslop-Harrison, J.S. (2000). Chromosomal variation in *Crocus vernus* Hill (Iridaceae) investigated by *in situ* hybridization of rDNA and a tandemly repeated sequence. Ann. Bot., in press.

Fuchs, J., Brandes, A., and Schubert, I. (1995). Telomere sequence localization and karyotype evolution in higher plants. Plant Syst. Evol. **196,** 227–241.

Fulnecek, J., Matyasek, R., Kovarik, A., and Bezdek, M. (1998). Mapping of 5-methylcytosine residues in *Nicotiana tabacum* 5S rRNA genes by genomic sequencing. Mol. Gen. Genet. **259,** 133–141.

Ganal, M.W., Broun, P., and Tanksley, S.D. (1992). Genetic-mapping of tandemly repeated telomeric DNA sequences in tomato (*Lycopersicon esculentum*). Genomics **14,** 444–448.

Gazdova, B., Siroky, J., Brzobohaty, B., Kenton, A., Parokonny, A., Heslop-Harrison, J.S., Palme, K., and Bezdek, M. (1995). Characterization of a new family of tobacco highly repetitive DNA, GRS, specific for the *Nicotiana tomentosiformis* genomic component. Chromosome Res. **3,** 245–254.

Gebhardt, C., Ritter, E., Barone, A., Debener, T., Walkemeier, B., Schachtschabel, U., Kaufmann, H., Thompson, R.D., Bonierbale, M.W., Ganal, M.W., Tanksley, S.D., and Salamini, F. (1991). RFLP maps of potato and their alignment with the homoeologous tomato genome. Theor. Appl. Genet. **83,** 49–57.

Gesteland, R.F., and Atkins, J.F. (1993). The RNA World. (New York: Cold Spring Harbor Laboratory Press).

Goffeau, A., et al. (1996). Life with 6000 genes. Science **274,** 546.

Goldberg, M.W., Cronshaw, J.M., Kiseleva, E., and Allen, T.D. (1999). Nuclear-pore-complex dynamics and transport in higher eukaryotes. Protoplasma **209,** 144–156.

Goubely, C., Arnaud, P., Tatout, C., Harrison, G., Heslop-Harrison, J.S., and Deragon, J.-M. (1999). S1 SINE retroelements are methylated at symmetrical and non-symmetrical positions in *Brassica napus*: Identification of a prefered target site for asymmetrical methylation. Plant Mol. Biol. **39,** 243–255.

Grandbastien, M.-A. (1998). Activation of plant retrotransposons under stress conditions. Trends Plant Sci. **3,** 181–187.

Gruenbaum, Y., Naveh-Many, T., Cedar, H., and Razin, A. (1981). Sequence specificity of methylation in higher plant DNA. Nature **292,** 860–862.

Harper, G., Osuji, J.O., Heslop-Harrison, J.S., and Hull, R. (1999). Integration of banana streak badnavirus into the *Musa* genome: Molecular and cytogenetic evidence. Virology **255,** 207–213.

Harrington, J.J., Van Bokkelen, G., Mays, R.W., Gustashaw, K., and Willard, H.F. (1997). Formation of *de novo* centromeres and construction of first-generation human artificial microchromosomes. Nat. Genet. **15,** 345–355.

Henikoff, S., and Comai, L. (1998). A DNA methyltransferase homolog with a chromodomain exists in multiple polymorphic forms in *Arabidopsis*. Genetics **149,** 307–318.

Henikoff, S., Loughney, K., and Dreesen, T.D. (1993). The enigma of dominant position-effect variegation in *Drosophila*. In The Chromosome, J.S. Heslop-Harrison and R.B. Flavell, eds (Oxford, UK: Bios), pp. 193–206.

Heslop-Harrison, J.S. (2000). RNA, genes, genomes and chromosomes: Repetitive DNA sequences in plants. Chromosomes Today **14,** in press.

Heslop-Harrison, J.S., and Bennett, M.D. (1990). Nuclear architecture in plants. Trends Genet. **6,** 401–405.

Heslop-Harrison, J.S., Huelskamp, M., Wendroth, S., Atkinson, M.D., Leitch, A.R., Bennett, M.D. (1988). Chromatin and centromeric structures in interphase nuclei. In Kew Chromosome Conference III, P.E. Brandham, ed (London: Allen and Unwin), pp. 209–217.

Heslop-Harrison, J.S., Leitch, A.R., and Schwarzacher, T. (1993). The physical organization of interphase nuclei. In The Chromosome, J.S. Heslop-Harrison and R.B. Flavell, eds (Oxford, UK: Bios), pp. 221–232.

Heslop-Harrison, J.S., et al. (1997). The chromosomal distributions of Ty1–*copia* group retrotransposable elements in higher plants and their implications for genome evolution. Genetica **100,** 197–204.

Heslop-Harrison, J.S., Murata, M., Ogura, Y., Schwarzacher, T., and Motoyoshi, F. (1999). Polymorphisms and genomic organization of repetitive DNA from centromeric regions of *Arabidopsis thaliana* chromosomes. Plant Cell **11,** 31–42.

Hirochika, H., Sugimoto, K., Otsuki, Y., Tsugawa, H., and Kanda, M. (1996). Retrotransposons of rice involved in mutations induced by tissue culture. Proc. Natl. Acad. Sci. USA **93,** 7783–7788.

Holmquist, G.P. (1992). Chromosome bands, their chromatin flavors, and their functional features. Am. J. Hum. Genet. **51,** 17–37.

Hull, R., and Covey, S.N. (1996). Retroelements: Propagation and adaptation. Virus Genes **11,** 105–118.

Jackson, D.A., and Cook, P.R. (1995). The structural basis of nuclear function. Int. Rev. Cytol. **162A,** 125–149.

Jacobsen, S.E., and Meyerowitz, E.M. (1997). Hypermethylated SUPERMAN epigenetic alleles in *Arabidopsis*. Science **277,** 1100–1103.

Jahner, D., and Jaenisch, R. (1984). DNA methylation in early mammalian development. In DNA Methylation: Biochemistry and Biological Significance, A. Razin, H. Cedar, and A. Riggs, eds (New York: Springer-Verlag), pp. 189–219.

Jakowitsch, J., Metter, M.F., van der Winden, J., Matzke, M.A., and Matzke, A.J.M. (1999). Integrated pararetroviral sequences define a unique class of dispersed repetitive DNA in plants. Proc. Natl. Acad. Sci. USA **66,** 13241–13246.

Jeddeloh, J.A., Bender, J., and Richards, E.J. (1998). The DNA methylation locus DDM1 is required for maintenance of gene silencing in *Arabidopsis*. Genes Dev. **12,** 1714–1725.

Jeddeloh, J.A., Stokes, T.L., and Richards, E.J. (1999). Maintenance of genomic methylation requires a SWI2/SNF2-like protein. Nat. Genet. **22,** 94–97.

Jiang, J., Nasuda, S., Dong, F., Scherrer, C.W., Woo, S.-S., Wing, R.A., Gill, B.S., and Ward, D.C. (1996). A conserved repetitive

DNA element located in the centromeres of cereal chromosomes. Proc. Natl. Acad. Sci. USA **93**, 14210–14213.

Johnson, C.A., Pradham, S., and Adams, R.L.P. (1995). The effect of histone H1 and DNA methylation on transcription. Biochem. J. **305**, 791–798.

Kakutani, T., Jeddeloh, J.A., and Richards, E.J. (1995). Characterization of an *Arabidopsis thaliana* DNA hypomethylation mutant. Nucleic Acids Res. **23**, 130–137.

Kakutani, T., Jeddeloh, J.A., Flowers, S.K., Munakata, K., and Richards, E.J. (1996). Developmental abnormalities and epimutations associated with DNA hypomethylation mutations. Proc. Natl. Acad. Sci. USA **93**, 12406–12411.

Kamm, A., Schmidt, T., and Heslop-Harrison, J.S. (1994). Molecular and physical organization of highly repetitive undermethylated DNA from *Pennisetum glaucum*. Mol. Gen. Genet. **244**, 420–425.

Kamm, A., Doudrick, R.L., Heslop-Harrison, J.S., and Schmidt, T. (1996). The genomic and physical organization of *Ty1-copia*–like sequences as a component of large genomes in *Pinus elliottii* var. *elliottii* and other gymnosperms. Proc. Natl. Acad. Sci. USA **93**, 2708–2713.

Karpen, G.H., Le, M.H., and Le, H. (1996). Centric heterochromatin and the efficiency of achiamaate disjunction in *Drosophila* female meiosis. Science **273**, 118–122.

Kass, S.U., and Adams, R.L.P. (1993). Inactive chromatin spreads from a focus of methylation. Mol. Cell. Biol. **13**, 7372–7379.

Kass, S.U., Landsberger, N., and Wolffe, A.P. (1997). DNA methylation directs a time-dependent repression of transcription initiation. Curr. Biol. **7**, 157–165.

Katsiotis, A., Schmidt, T., and Heslop-Harrison, J.S. (1996). Chromosomal and genomic organization of *Ty1-copia*-like retrotransposon sequences in the genus *Avena*. Genome **39**, 410–417.

Kihara, H. (1924). Zytologische und genetische Studien bei wichtigen Getreidearten mit besonderer Rücksicht auf das Verhalten der Chromosomen und die Sterilität in den Bastarden. Mem. Coll. Sci. Kyoto Imp. Univ. **B**, 1–200.

Kilby, N.J., Leyser, H.M.O., and Furner, I.J. (1992). Promoter methylation and progressive transgene inactivation in *Arabidopsis*. Plant Mol. Biol. **20**, 103–112.

Kilian, A., Stiff, C., and Kleinhofs, A. (1995). Barley telomeres shorten during differentiation but grow in callus culture. Proc. Natl. Acad. Sci. USA **92**, 9555–9559.

Kipling, D., and Warburton, P.E. (1997). Centromeres, CENP-B, and *Tigger* too. Trends Genet. **13**, 141–145.

Koukalova, B., Kuhrova, V., Vyskot, B., Siroky, J., and Bezdek, M. (1994). Maintenance of the induced hypomethylation state of tobacco nuclear repetitive DNA sequences in the course of protoplast and plant regeneration. Planta **194**, 306–310.

Kubis, S., Heslop-Harrison, J.S., and Schmidt, T. (1997). A family of differentially amplified DNA sequences in the genus *Beta* reveals genetic variation in *Beta vulgaris* subspecies and cultivars. J. Mol. Evol. **44**, 310–320.

Kubis, S.E., Schmidt, T., and Heslop-Harrison, J.S. (1998a). Repetitive DNA elements as a major component of plant genomes. Ann. Bot. **82S**, 45–55.

Kubis, S.E., Heslop-Harrison, J.S., Desel, C., and Schmidt, T. (1998b). The genomic organization of non-LTR retrotransposons (LINEs) from three *Beta* species and five other angiosperms. Plant Mol. Biol. **36**, 821–831.

Kumar, A. (1998). The evolution of plant retroviruses: Moving to green pastures. Trends Plant Sci. **3**, 371–374.

Kumar, A., and Bennetzen, J.F. (1999). Plant retrotransposons. Annu. Rev. Genet. **33**, 497–532.

Lin, X.Y., et al. (1999). Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*. Nature **402**, 761–768.

Lucas, H., Feuerbach, F., Kunert, K., Grandbastien, M.-A., and Caboche, M. (1995). RNA-mediated transposition of the tobacco retrotransposon Tnt1 in *Arabidopsis thaliana*. EMBO J. **14**, 2364–2373.

Maluszynska, J., and Heslop-Harrison, J.S. (1991). Localization of tandemly repeated DNA sequences in *Arabidopsis thaliana*. Plant J. **1**, 159–166.

Manuelidis, L., and Chen, T.L. (1990). A unified model of eukaryotic chromosomes. Cytometry **11**, 8–25.

Martienssen, R. (1998). Transposons, DNA methylation, and gene control. Trends Genet. **14**, 263–264.

Martienssen, R., and Henikoff, S. (1999). The *House and Garden* guide to chromatin remodelling. Nat. Genet. **22**, 6–7.

Martin, W., and Herrmann, R.G. (1998). Gene transfer from organelles to the nucleus: How much, what happens, and why? Plant Physiol. **118**, 9–17.

Matzke, M.A., Primig, M., Trnovsky, J., and Matzke, A.J.M. (1989). Reversible methylation and inactivation of marker genes in sequentially transformed tobacco plants. EMBO J. **8**, 643–649.

Matzke, A.J.M., Neuhuber, F., Park, Y.-D., Ambros, P.F., and Matzke, M.A. (1994). Homology-dependent gene silencing in transgenic plants: Epistatic silencing loci contain multiple copies of methylated transgenes. Mol. Gen. Genet. **244**, 219–229.

Mayer, K., Schuller, C., European Union Arabidopsis Sequencing Consortium, and the Cold Spring Harbor, Washington University in St. Louis, and PE Biosystems Arabidopsis Sequencing Consortium (1999). Sequence and analysis of chromosome 4 of the plant *Arabidopsis thaliana*. Nature **402**, 769–777.

McAllister, B.F., and Werren, J.H. (1999). Evolution of tandemly repeated sequences: What happens at the end of an array? J. Mol. Evol. **48**, 469–481.

Miller, J.T., Dong, F., Jackson, S.A., Song, J., and Jiang, J. (1998). Retrotransposon-related DNA sequences in the centromeres of grass chromosomes. Genetics **150**, 1615–1623.

Moran, J.V., DeBerardinis, R.J., and Kazazzian, H.H., Jr. (1999). Exon shuffling by L1 retrotransposition. Science **283**, 1530–1534.

Muller, A., Philipps, G., and Gigot, C. (1980). Properties of condensed chromatin in barley nuclei. Planta **149**, 69–77.

Murata, M., Ogura, Y., and Motoyoshi, F. (1994). Centromeric repetitive sequences in *Arabidopsis thaliana*. Jpn. J. Genet. **69**, 361–371.

Neves, N., Heslop-Harrison, J.S., and Viegas, W. (1995). rRNA gene activity and control of expression mediated by methylation and imprinting during embryo development in wheat × rye hybrids. Theor. Appl. Genet. **91**, 529–533.

Neves, N., Silva, M., Heslop-Harrison, J.S., and Viegas, W. (1997). Nucleolar dominance in triticale: Control by unlinked genes. Chromosome Res. **5**, 125–131.

Nitta, I., Kamada, Y., Noda, H., Ueda, T., and Watanabe, K. (1998). Reconstitution of peptide bond formation with *Escherichia coli* 23S ribosomal RNA domains. Science **281,** 666–669.

Nonomura, K.I., and Kurata, N. (1999). Organization of the 1.9-kb repeat unit RCE1 in the centromeric region of rice chromosomes. Mol. Gen. Genet. **261,** 1–10.

Oakeley, E.J., Podesta, A., and Jost, J.-P. (1997). Developmental changes in DNA methylation of the two tobacco pollen nuclei during maturation. Proc. Natl. Acad. Sci. USA **94,** 11721–11725.

Ohta, K., Shibata, T., and Nicolas, A. (1994). Changes in chromatin structure at recombination initiation sites during yeast meiosis. EMBO J. **13,** 5754–5763.

Paques, F., Leuyn, W.-Y., and Haber, J.E. (1998). Expansions and contractions in a tandem repeat induced by double-strand break repair. Mol. Cell. Biol. **18,** 2045–2054.

Pearce, S.R., Harrison, G., Li, D., Heslop-Harrison, J.S., Kumar, A., and Flavell, A.J. (1996). The Ty1–*copia* group retrotransposons in *Vicia* species: Copy number, sequence heterogeneity, and chromosomal localisation. Mol. Gen. Genet. **250,** 305–315.

Pedersen, C., and Linde-Laursen, I. (1994). Chromosomal locations of four minor rDNA loci and a marker microsatellite sequence in barley. Chromosome Res. **2,** 65–71.

Pelissier, T., Thalmeir, S., Kempe, D., Sänger, H.-L., and Wassenegger, M. (1999). Heavy de novo methylation at symmetrical and non-symmetrical sites is a hallmark of RNA-directed DNA methylation. Nulceic Acids Res. **27,** 1625–1634.

Plasterck, R. (1998). Ragtime jumping. Nature **394,** 718–719.

Pluta, A., Cooke, F.C.A., and Earnshaw, W.C. (1990). Structure of the human centromere at metaphase. Trends Biol. Sci. **15,** 181–185.

Pradhan, S., and Adams, R.L.P. (1995). Distinct CG and CNG methyltransferases in *Pisum sativum*. Plant J. **7,** 471–481.

Pradhan, S., Houlston, C., and Adams, R.L.P. (1995). CG and CNG methyltransferases in plants. Gene **157,** 289–291.

Presting, G.G., Malysheva, L., Fuchs, J., and Schubert, I. (1998). A TY3/GYPSY retrotransposon-like sequence localizes to the centromeric regions of cereal chromosomes. Plant J. **16,** 721–728.

Pruitt, R.E., and Meyerowitz, E.M. (1986). Characterization of the genome of *Arabidopsis thaliana*. J. Mol. Biol. **187,** 169–183.

Prymakowska-Bosak, M., Przewloka, M.R., Iwkiewicz, J., Egierszdorff, S., Kuras, M., Chaubet, N., Gigot, C., Spiker, S., and Jerzmanowski, A. (1996). Histone H1 overexpressed to high level in tobacco affects certain developmental programs but has limited effect on basal cellular functions. Proc. Natl. Acad. Sci. USA **93,** 10250–10255.

Rabl, C. (1885). Über Zelltheilung. Morphol. Jahrb. **10,** 214–330.

Radic, M.Z., Lundgren, K., and Hamkalo, B.A. (1987). Curvature of mouse satellite DNA and condensation of heterochromatin. Cell **50,** 1101–1108.

Raman, R., and Nanda, I. (1986). Mammalian sex chromosomes. I. Cytological changes in the chiasmatic sex chromosomes of the male musk shrew, *Suncus murinus*. Chromosoma **93,** 367–374.

Rawlins, D.J., Highett, M.I., and Shaw, P.J. (1991). Localization of telomeres in plant interphase nuclei by *in situ* hybridization and 3D confocal microscopy. Chromosoma **100,** 424–431.

Razin, A., and Cedar, H. (1993). DNA methylation and embryogenesis. In DNA Methylation: Molecular Biology and Biological Significance, J.P. Jost and H.P. Salus, eds (Basel, Switzerland: Birkhauser), pp. 523–568.

Richards, E.J., and Ausubel, F.M. (1988). Isolation of a higher eukaryotic telomere from *Arabidopsis thaliana*. Cell **53,** 127–136.

Ronemus, M.J., Galbiati, M., Ticknor, C., Chen, J., and Dellaporta, S.L. (1996). Demethylation-induced developmental pleiotropy in *Arabidopsis*. Science **273,** 654–657.

Sabl, J.F., and Henikoff, S. (1996). Copy number and orientation determine the susceptibility of a gene to silencing by nearby heterochromatin in *Drosophila*. Genetics **142,** 447–458.

Saccone, S., De Sario, A., Wiegant, J., Raap, A.K., Valle, G.D., and Bernardi, G. (1993). Correlations between isochores and chromosomal bands in the human genome. Proc. Natl. Acad. Sci. USA **90,** 11929–11933.

Sadoni, N., Lange, S., Fauth, C., Bernardi, G., Cremer, T., and Turner, B.M. (1999). Nuclear organization of mammalian genomes: Polar chromosome territories build up functionally distinct higher compartments. J. Cell Biol. **146,** 1211–1226.

SanMiguel, P., Tikhonov, A., Jin, Y.-K., Motchoulskaia, N., Zakharov, D., Melake-Berhan, A., Springer, P.S., Edwards, K.J., Lee, M., Avramova, Z., and Bennetzen, J.L. (1996). Nested retrotransposons in the intergenic regions of the maize genome. Science **274,** 765–768.

Sardana, R., O'Dell, M., and Flavell, R.B. (1993). Correlation between the size of the intergenic regulatory region, the status of cytosine methylation of rRNA genes, and nucleolar expression in wheat. Mol. Gen. Genet. **236,** 155–162.

Schlotterer, C., and Tautz, D. (1992). Slippage synthesis of simple sequence DNA. Nucleic Acids Res. **20,** 211–215.

Schmidt, T. (1999). LINEs, SINEs, and repetitive DNA: Non-LTR retrotransposons in plant genomes. Plant Mol. Biol. **40,** 903–910.

Schmidt, T., and Heslop-Harrison, J.S. (1994). Variability and evolution of highly repeated DNA sequences in the genus *Beta*. Genome **36,** 1074–1079.

Schmidt, T., and Heslop-Harrison, J.S. (1996). The physical and genomic organization of microsatellites in sugar beet. Proc. Natl. Acad. Sci. USA **93,** 8761–8765.

Schmidt, T., and Heslop-Harrison, J.S. (1998). Genomes, genes, and junk: The large-scale organization of plant chromosomes. Trends Plant Sci. **3,** 195–199.

Schmidt, T., Jung, C., Heslop-Harrison, J.S., and Kleine, M. (1997). Detection of alien chromatin conferring resistance to the beet cyst nematode (*Heterodera schachtii* Schm.) in cultivated beet (*Beta vulgaris* L.) using *in situ* hybridization. Chromosome Res. **5,** 186–193.

Schwarzacher, T. (1997). Three stages of meiotic homologous chromosome pairing in wheat: Cognition, alignment, and synapsis. Sex. Plant Reprod. **10,** 324–331.

Schwarzacher, T., and Heslop-Harrison, J.S. (1990). *In situ* hybridization to plant telomeres using synthetic oligomers. Genome **34,** 317–323.

Schwarzacher, T., and Heslop-Harrison, J.S. (2000). Practical *in situ* Hybridization. (Oxford, UK: Bios).

Sears, E.R. (1941). Chromosome pairing and fertility in hybrids and amphidiploids in the *Triticinae*. Mo. Agric. Exp. Stn. Res. Bull. **337,** 1–20.

Silva, M., Queiroz, A., Neves, N., Barao, A., Castilho, A., Morais, L., and Viegas, W. (1995). Reprogramming of rye rDNA in triticale during microsporogenesis. Chromosome Res. **3**, 492–496.

Siroky, J., Ruffini Castiglione, M., and Vyskot, B. (1998). DNA methylation and replication patterns of *Melandrium album* chromosomes. Chromosome Res. **6**, 441–446.

Smit, A.F.A. (1996). The origin of interspersed repeats in the human genome. Curr. Opin. Genet. Dev. **6**, 743–748.

Smith, S.S. (1998). Stalling of DNA methyltransferase in chromosome stability and chromosome remodelling (review). Int. J. Mol. Med. **1**, 147–156.

Taketa, S., Harrison, G., and Heslop-Harrison, J.S. (1999). Comparative physical mapping of the 5S and 18S–25S rDNA in nine wild *Hordeum* species and cytotypes. Theor. Appl. Genet. **98**, 1–9.

Tate, P.H., and Bird, A.P. (1993). Effects of DNA methylation on DNA-binding proteins and gene expression. Curr. Opin. Genet. Dev. **3**, 226–231.

Tatout, C., Lavie, L., and Deragon, J.-M. (1998). Similar target site selection occurs in integration of plant and mammalian retroposons. J. Mol. Ecol. **47**, 463–470.

Tautz, D., and Renz, M. (1984). Simple sequences are ubiquitous components of eukaryotic genomes. Nucleic Acids Res. **12**, 4127–4138.

Tazi, J., and Bird, A. (1990). Alternative chromatin structure at CpG islands. Cell **60**, 909–920.

Tiersch, T.R., and Wachtel, S.S. (1991). On the evolution of genome size of birds. J. Hered. **82**, 363–368.

Tikhonov, A.P., SanMiguel, P.J., Nakajima, Y., Gorenstein, N.M., Bennetzen, J.L., and Avramova, Z. (1999). Colinearity and its exceptions in orthologous adh regions of maize and sorghum. Proc. Natl. Acad. Sci. USA **96**, 7409–7414.

Turner, B.M. (1991). Histone acetylation and control of gene expression. J. Cell Sci. **99**, 13–20.

Tyler-Smith, C., Corish, P., and Burns, E. (1998). Neocentromeres, the Y chromosome, and centromere evolution. Chromosome Res. **6**, 65–71.

Vaughan, H.E., Heslop-Harrison, J.S., and Hewitt, G.M. (1999). The localisation of mitochondrial sequences to chromosomal DNA in Orthopterans. Genome **42**, 874–880.

Vernhettes, S., Grandbastien, M.-A., and Casacuberta, J.M. (1998). The evolutionary analysis of the Tnt1 retrotransposon in *Nicotiana* species reveals the high plasticity of its regulatory sequences. Mol. Biol. Evol. **15**, 827–836.

Vershinin, A.V., and Heslop-Harrison, J.S. (1998). Comparative analysis of the nucleosomal structure of rye, wheat, and their relatives. Plant Mol. Biol. **36**, 149–161.

Vershinin, A.V., Schwarzacher, T., and Heslop-Harrison, J.S. (1995). The large-scale genomic organization of repetitive DNA families at the telomeres of rye chromosomes. Plant Cell **7**, 1823–1833.

Vongs, A., Kakutani, T., Martienssen, R.A., and Richards, E.J. (1993). *Arabidopsis thaliana* DNA methylation mutants. Science **260**, 1926–1928.

Wade, P.A., Pruss, D., and Wolffe, A.P. (1997). Histone acetylation: Chromatin in action. Trends Biochem. Sci. **22**, 128–132.

Walbot, V. (1999). UV-B damage amplified by transposons in maize. Nature **397**, 398–399.

Wang, S., Lapitan, N.L.V., and Tsuchiya, T. (1992). Characterization of telomeres in *Hordeum vulgare* chromosomes by in situ hybridization. II. Healed broken chromosomes in telotrisomic 4L and acrotrisomic 4L4S lines. Genome **35**, 975–980.

Wassenegger, M., Heimes, S., Riedel, L., and Sänger, H.L. (1994). RNA-directed *de novo* methylation of genomic sequences in plants. Cell **76**, 567–576.

Weising, K., Bayermann, B., Ramser, J., and Kahl, G. (1991). Plant DNA fingerprinting with radioactive and digoxigenated oligonucleotide probes complementary to simple repetitive DNA sequences. Electrophoresis **12**, 159–169.

Wendel, J.F., Schnabel, A., and Seelanan, T. (1995). Bidirectional interlocus concerted evolution following allopolyploid speciation in cotton (*Gossypium*). Proc. Natl. Acad. Sci. USA **92**, 280–284.

Wessler, S.R. (1996). Plant retrotransposons: Turned on by stress. Curr. Biol. **6**, 959–961.

Wessler, S.R., Bureau, T.E., and White, S.E. (1995). LTR-retrotransposons and MITEs: Important players in the evolution of plant genomes. Curr. Opin. Genet. Dev. **5**, 814–821.

Wolffe, A.P. (1995). Chromatin: Structure and Function. (San Diego, CA: Academic Press).

Wolffe, A.P., and Pruss, D. (1996). Deviant nucleosomes: The functional specialization of chromatin. Trends Genet. **12**, 58–62.

Wu, T.-C., and Lichten, M. (1994). Meiosis-induced double-strand break sites determined by yeast chromatin structure. Science **263**, 515–518.

Wu, T., Wang, Y., and Wu, R. (1994). Transcribed repetitive DNA sequences in telomeric regions of rice (*Oryza sativa*). Plant Mol. Biol. **26**, 363–375.

Yoder, J.A., Walsh, C.P., and Bestor, T.H. (1997). Cytosine methylation and the ecology of intragenomic parasites. Trends Genet. **13**, 335–340.

Zakian, V. (1995). Telomeres: Beginning to understand the end. Science **270**, 1601–1607.

Zhong, X.B., Fransz, P.F., Wennekes-van Eden, J., Ramanna, M.S., van Kammen, A., Zabel, P., and de Jong, J.H. (1998). FISH studies reveal the molecular and chromosomal organization of individual telomere domains in tomato. Plant J. **13**, 507–517.

Zink, D., Cremer, T., Saffrich, R., Fischer, R., Trendelenburg, M.F., Ansorge, W., and Stelzer, E.H.K. (1998). Structure and dynamics of human interphase chromosome territories *in vivo*. Hum. Genet. **102**, 241–251.