Annals of Botany **82** (Supplement A): 45–55, 1998
Article No. bo980779

Repetitive DNA Elements as a Major Component of Plant Genomes

SYBILLE KUBIS*†, THOMAS SCHMIDT‡ and JOHN SEYMOUR (PAT) HESLOP-HARRISON*

** Karyobiology Group, John Innes Centre, Colney Lane, Norwich NR4 7UH, England, † Norman Borlaug Institute, DeMontfort University, Scraptoft, Leicester LE7 9SU, England and ‡ Institute of Crop Science and Plant Breeding, Christian-Albrechts University, Olshausenstr. 40, D-24118 Kiel, Germany*

Received: 18 August 1998 Returned for revision: 27 August 1998 Accepted: 4 September 1998

A major part of the nuclear genome of most plants is composed of different repetitive DNA elements. Studying these sequence elements is essential for our understanding of the nature and consequences of genome size variation between different species, and for studying the large-scale organization and evolution of plant genomes. Sugar beet (*Beta vulgaris* L.) is an important crop and a suitable model for such investigations: with a genome size of 0·8 pg 1C (760 Mbp) it contains significant amounts of all major groups of repetitive sequences among its nine chromosome pairs, but analysis is not complicated by polyploidy or the huge size of some genomes, and there are valuable genetic data, recombinant DNA libraries and wild relatives to complement studies of sequence contribution to genome size in sugar beet. A sophisticated understanding of the structure of the genome will provide valuable data about the major factors responsible for genome size variation, useful aids in the development of a molecular understanding of genome evolution, and perhaps indicate strategies for crop improvement. Using molecular and cytological approaches, we have characterized a range of differentially organized repetitive DNA sequence elements from the genomes of cultivated and wild beet species, leading to an extensive model of the repetitive DNA, its organization and evolution.

Key words: Genome evolution, *Beta*, retroelements, satellite DNA, microsatellites, nucleosomes.

INTRODUCTION

As reviewed in other papers in this volume, the amount of nuclear DNA varies extensively between plant species. A proportion of this variation is due to polyploidy, and it is assumed that 50% or more of angiosperms are polyploids. The size of the unreplicated haploid genome is characteristic for each species and expressed by the 1C value, which may be converted to the number of Mbp in the genome. Although all plants require approximately the same minimum number of genes and regulatory sequences for their development, including germination, growth, flowering and reproduction, the nuclear DNA amount still differs by several orders of magnitude in diploids, from about 0·2 pg (approx. 150 Mbp) 1C for several species including *Arabidopsis thaliana* L., up to nearly 90 pg (> 85000 Mbp) in *Fritillaria davisii* Turrill ($2n = 24$) (Bennett and Leitch, 1995). Species of one taxonomic family show similar morphology and, indeed, genes represented in all species can be often regarded as allelic variants. Furthermore, genes may be colinear (in the same genetic order) over large taxonomic distances. However, the nuclear DNA content of related species can vary widely; for example, the genomes of *Oryza sativa* L. (rice) and *Secale cereale* L. (rye) in the Gramineae differ by a factor of 16, with 1C values of 0·6 pg (580 Mbp) for rice and 9·5 pg (9300 Mbp) for rye (Bennett and Smith, 1976). Thomas (1971) described this phenomenon as the C-value paradox.

In total, the DNA sequences of low-copy genes and regulatory sequences make up a small proportion of the total amount of nuclear DNA in most plant species: the major fraction of most plant nuclear genomes is made up of repetitive DNA elements. Such DNA elements consist of sequence motifs ranging in size from dinucleotides to more than 10000 bp. Copy numbers of individual repetitive DNA motifs can vary from several hundred to hundreds of thousands, and single motifs may represent 10 or even 50% of a genome. Families of repetitive DNA sequences are differentiated by their degree of sequence homology, distribution among species and/or genomic and physical organization. Repetitive DNA elements can be divided into two major groups, distinguished by their genomic organization and localization on the chromosomes, although intermediate forms of organization can exist too. One group includes sequences showing an organization in tandem repeating units, where individual copies are arranged adjacently to each other forming tandem arrays of the monomeric unit. Such tandemly repeated DNAs are found preferentially at specific positions of the chromosomes, such as the pericentromeric, subtelomeric, telomeric or intercalary regions. DNA elements arranged in tandem arrays include different types of satellite DNAs, the telomeric repeat and the rDNA.

The other group of repetitive DNA sequences comprises elements with a dispersed organization. Dispersed repetitive DNA elements are scattered throughout the genome, interspersed with other sequences and distributed along the chromosomes, although regions of depletion or amplification can be found. Blocks of nested copies of such elements have been observed (SanMiguel *et al.*, 1996;

0305-7364/98/0A0045+11 $30.00/0

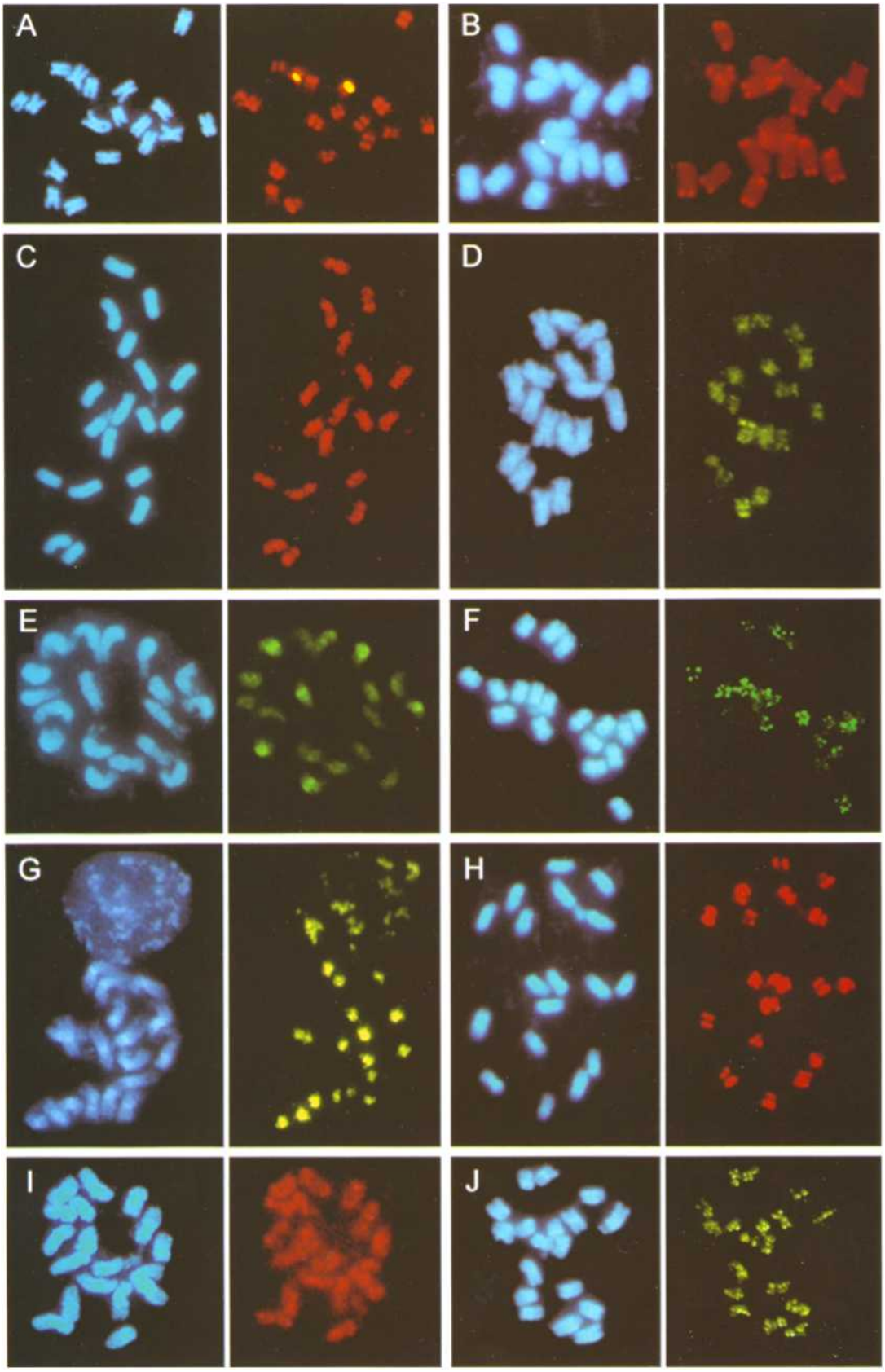

FIG. 1. For legend see facing page.

Higashiyama *et al.*, 1997). Dispersed DNA sequences include mobile elements, like DNA transposable elements and retroelements, their remnants and SINEs (short interspersed nuclear elements), and other dispersed repeats. DNA transposable elements move to new locations within the host genome via exclusively DNA intermediates. They code for a transposase responsible for integration and excision processes and are flanked by short terminal inverted repeats. During integration a short sequence of the host is duplicated, leading to direct repeats of 3–8 bp at each side of the transposon copy. The first transposon described in plants was the Ac-Ds control element of maize (*Zea mays* L.) discovered by Barbara McClintock (1951). Other well described examples are the En-Spm elements of maize and the Tam transposons of *Antirrhinum majus* L. (McClintock, 1961; Coen and Carpenter, 1986; Gierl *et al.*, 1988). Retroelements, the other major group of mobile elements, use RNA intermediates for transposition (see below). Retroelements are a major component of plant nuclear genomes, representing up to 50% of the nuclear DNA (Bennetzen, 1996; SanMiguel *et al.*, 1996; Pearce *et al.*, 1996; SanMiguel and Bennetzen, 1998), and copies of different types of retroelements were also detected in the mitochondrial genome of *A. thaliana* (Knoop *et al.*, 1996).

For cloning of highly repetitive DNA elements different strategies are feasible. One way is the shot-gun cloning of DNA fragments after digestion with a frequent cutting restriction enzyme followed by selection of highly repetitive clones through dot blot hybridization with total genomic DNA as a probe. A different approach is the cloning of distinct DNA fragments visible after gel electrophoresis of digested DNA. In the present paper, we aim to review the nature of these various elements as major contributors to the size of genomes in plants.

SUGAR BEET AS A MODEL GENOME

Sugar beet (*B. vulgaris* L.) is an important crop and a suitable model for investigations of the distribution of different classes of repetitive DNA and their contribution to genome size, and to investigate micro- and macro-evolution of the sequences in the species, its wild and more distant relatives. The genus *Beta* belongs to the widely distributed family Chenopodiaceae Vent. Wild species of the genus show a high genetic and phenotypic variability: the genus *Beta* is divided into four sections: *Beta*, *Corollinae*, *Nanae* and *Procumbentes* (Barocka, 1985). The basic chromosome number is $n = 9$; most *Beta* species are diploid ($2n = 2x = 18$), but tetra-, penta- and hexaploid plants are found too, and the triploid is grown commercially. The chromosomes are small and morphologically uniform (meta- to submetacentric; see Fig. 1). Cultivated beets, including sugar beet, fodder beet, beet root, leaf beet and Swiss chard, belong to the subspecies *B. vulgaris vulgaris*, with the 'maritima-complex' (*B. vulgaris maritima*) being close relatives to cultivated beets.

The genome of sugar beet is 0·8 pg 1C (760 Mbp) in size (Arumuganathan and Earle, 1991). About 63% of the genome is composed of repetitive DNA sequences (Flavell *et al.*, 1974). During the last 10 years sugar beet has increasingly become an object for studies using molecular biological methods. Genetic markers have been developed using different techniques, and genetic maps of sugar beet have been constructed including RFLP- (restriction fragment length polymorphism), RAPDs- (random amplified polymorphic DNAs) and AFLP- (amplified fragment length polymorphism) markers (Barzen *et al.*, 1992, 1995; Pillen *et al.*, 1992, 1993; Hallden *et al.*, 1996, 1993; Schondelmaier, Steinrücken and Jung, 1996). YAC (yeast artificial chromosome) libraries of sugar beet have been constructed (Eyers, Edwards and Schuch, 1992; Del-Favero *et al.*, 1994; Kleine *et al.*, 1995). The existence of the genetic maps together with the availability of the YAC libraries will be vital for the isolation and characterization of economically important traits or genes from sugar beet on the basis of their known map position by positional cloning or chromosome walking. In that respect the integration of physical and genetics maps is a major task, since the frequency of recombination between genes is not random over the genome. Low levels of recombination around the centromeres have been observed (Devos, Millan and Gale, 1993; Laurie *et al.*, 1993; Schondelmaier *et al.*, 1997). Therefore while the order of markers is the same, the genetic and physical distances between them show little connection (Heslop-Harrison, 1991). It is important to know the physical distance between genes to interpret genetic maps and to produce viable strategies for map based cloning. An understanding of repetitive DNA is critical for this study.

FUNCTIONAL TANDEM REPEATS

While most repetitive sequences have, at most, a controversial role in the genome—sometimes being regarded as

FIG. 1. Chromosomal localization of different repetitive DNA elements in sugar beet (*Beta vulgaris*). The left panel of each part shows metaphase chromosomes stained with DAPI (blue fluorescence). The right panel illustrates the signals of digoxygenin-labelled (detection with FITC – green-yellow fluorescence) and/or biotin-labelled (detection with Cy3 – red fluorescence) probes. A, Localization of the pEV1 satellite on both arms of all sugar beet chromosomes in different size clusters (red fluorescence). One pair of 5S rRNA gene clusters is present on the short arm of one chromosome (yellow fluorescence). B, Telomeric repeats (TTTAGGG) hybridize exclusively to chromosome ends (red fluorescence). C, Hybridization with the SSR motif $(GA)_8$ results in strong signals along all 18 chromosomes with exclusions from centromeric regions (red fluorescence). D, Hybridization with clone pDV3, containing a compound SSR, shows specific amplification on two chromosome arms (green fluorescence). E, The SSR motif $(GATA)_4$ is amplified at six sites, on three chromosome pairs (green fluorescence). F, The satellite pSV1 lies at multiple sites along intercalary domains of metaphase chromosomes (green fluorescence). G, The satellite repeat pBV1 hybridizes exclusively to the pericentromeric heterochromatin of all 18 chromosomes (yellow-green fluorescence). It contains the SSR motif $(AC)_8$, which itself gives strong hybridization signals at the centromere. H, The satellite pHC28 is amplified at the sites of intercalary heterochromatin of all 18 sugar beet chromosomes (red fluorescence), but shows different amplification patterns to pEV1 (Fig. 1A). I, The Tyl-*copia* retroelement Tbv1 is dispersed along all chromosomes (red fluorescence). Exclusion from the 18S-5.8S-25S rDNA sites and centromeric regions was observed. J, The LINE BNR1 shows organization in discrete clusters on all chromosomes (green fluorescence). It is excluded from centromeric regions and the 18S-5.8S-25S rDNA sites.

'junk DNA'—a few have important and well defined roles in stabilizing the chromosomes and encoding genes required in high copy numbers.

The ribosomal genes are highly repetitive and arranged in tandem at a small number of sites (loci) in the genome. The rDNA repeats include the 5S rRNA genes and the 18S-5·8S-25S rRNA genes, with intergenic spacers. The repeat unit of the latter may be 10 kbp long, and in *A. thaliana* the two pairs of sites have some 570 copies, representing 5% of the total genome size. The discrete numbers of sites, evolutionarily rapid change in copies at different loci, and easy assay by *in situ* hybridization, have made the rDNA loci valuable markers for investigating the evolution of chromosomes, particularly in the Triticeae (Leitch and Heslop-Harrison, 1993; Dubcovsky and Dvorak, 1995; Castilho and Heslop-Harrison, 1996). In sugar beet, the clusters of the 18S-5.8S-25S rRNA genes and intergenic spacers were localized at the secondary constriction at the end of the short arm of chromosome 1 (Schmidt, Schwarzacher and Heslop-Harrison, 1994) and the rDNA units have also been used in phylogenetic studies of the genus *Beta* (Santoni and Berville, 1992). Sugar beet contains one pair of 5S rRNA gene clusters near the centromere on the short arm of one chromosome (Fig. 1A) and the locus was mapped genetically to linkage group II using fluorescent *in situ* hybridization (Schondelmaier *et al.*, 1997).

Telomeric DNA consists in most plants of conserved 7 bp repeats (TTTAGGG); unlike all other nuclear DNA sequences, the terminal units are not replicated from pre-existing DNA by semi-conservative replication, but are added to the physical ends of the chromosomes by an enzyme, telomerase. This unusual enzyme is a reverse transcriptase (see retroelements below), incorporating an RNA template. The control of the number of copies of the sequence at chromosome ends is therefore under different constraints from those on other genomic DNA sequences (Schwarzacher and Heslop-Harrison, 1991), and the average length may vary both from cell to cell and within chromosome linkage groups, but is typically a few kilobases. Telomeric arrays can be visualized by *in situ* hybridization using a synthetic oligonucleotide probe complementary to the TTTAGGG motif. In sugar beet, signals are detected exclusively at the ends of chromosome arms of sugar beet (Schmidt *et al.*, 1998*b*; Fig. 1B). Intercalary arrays of the sequence are also known (Fuchs, Brandes and Schubert, 1995), but are presumably replicated by the usual mechanisms.

TANDEMLY REPEATED DNA

Simple sequence repeats and minisatellites

According to the size of the repeating unit, simple sequence repeats (SSRs or microsatellites) with motifs of 2–6 bp and minisatellites with monomeric units of 10–40 bp are distinguished from other satellite DNA families with larger repeat monomers (see below). Different SSRs are major components of the repeated DNA fraction in many species, and some motifs can be used to give 'fingerprint' patterns when probed to size-separated genomic DNA digests. In *Beta*, short exposures, indicating high genomic abundance, can differentiate subspecies and cultivars (Schmidt *et al.*, 1993). In beet, the chromosomal distribution pattern of di-, tri- and tetra-nucleotide microsatellites has been investigated by *in situ* hybridization (Schmidt and Heslop-Harrison, 1996*b*). Each microsatellite sequence shows a characteristic genomic distribution and motif-dependent dispersion, with site-specific amplification on one to seven pairs of centromeres or intercalary chromosomal regions (Fig. 1C–E). Several motifs revealed a weaker, dispersed hybridization along chromosomes (Fig. 1D). Microsatellites were excluded from 18S-5.8S-25S rDNA clusters and some motifs were further absent from centromeres and particular intercalary regions.

SSRs are ubiquitous in plants and evolve rapidly—hence they are valuable as molecular markers and for fingerprinting. Sometimes SSRs are components of large repeat motifs: pBV1 (Table 1 and Fig. 1G) is a sequence 327–328 bp long containing a $(AC)_8$ motif. Other SSRs may be present as much longer arrays not associated with other repetitive motifs. *In vitro* experiments suggest that slippage replication is the main mechanism responsible for the formation and expansion of microsatellite arrays (Schloetterer and Tautz, 1992).

Satellite DNA repeat families

Satellite DNA families or tandem DNA repeats are groups of identical or similar sequences, which are organized in blocks of tandemly repeated monomers. Satellite DNAs were first discovered in plants as additional bands of DNA beside the major fraction after ultracentrifugation in CsCl-density gradients due to their differing GC-content from the average genomic content of 40–45% (Hemleben, 1990). Alternatively, they are detectable as restriction satellite DNAs after gel electrophoretic separation of enzyme-digested DNA, where they are visible as distinct size fragments in the smear of restricted genomic DNA (Pech, Igo-Kemenes and Zachau, 1979). Many satellite DNA families have been identified in this way. Repeats from rye (*S. cereale*) were among the first satellite DNAs isolated, with one of them representing 6% of the rye genome (Bedbrook *et al.*, 1980). The genomic localization of such tandem DNA repeats was first demonstrated using tritium-labelled DNA probes. The development of fluorescent *in situ* hybridization enabled detailed studies of the physical organization of these repetitive DNA elements, and is now widely used for such analyses. Satellite DNA families were found to be localized in regions of the constitutive heterochromatin, present in pericentric, subtelomeric and, dependent on the plant species, distinct intercalary chromosomal regions (Traut, 1991). High resolution physical mapping on chromosomes in prophase and DNA fibres revealed that blocks of distinct satellite DNA families can alternate and that the repeats can be interspersed with other sequences, including retroelements (see below, Schmidt and Heslop-Harrison, 1996*a*; Brandes *et al.*, 1997).

The length of the monomeric unit of satellite DNA families is variable, but preferential sizes of 150–180 bp and 320–360 bp for the monomeric unit have been observed in dicotyledonous plant species, similar to the length of DNA

TABLE 1. *Tandemly repeated DNA sequences in the genus* Beta

Satellite	Enzyme	Isolated from	Repeat size (bp)	Chromosomal position
pBV1	*Bam*HI	*B. vulgaris* L.	327–328	Pericentric; Fig. 1G
pEV1	*Eco*RI	*B. vulgaris* L.	156–160	Intercalary; Fig. 1A
pSV1	*Sau3*AI	*B. vulgaris* L.	143	Intercalary; Fig. 1F
pHT30	*Hae*III	*B. trigyna* W. et K.	140–149	Not tested
pHT49	*Hae*III	*B. trigyna* W. et K.	162	Not tested
pHC28	*Hin*fI	*B. corolliflora* Zos.	149	Intercalary; Fig. 1H
pRN1	*Rsa*I	*B. nana* Bois. & Held.	209–233	Pericentric/intercalary
pTS5	*Sau3*AI	*B. procumbens* Chr. Sm.	153–160	Pericentric
pTS4.1	*Sau3*AI	*B. procumbens* Chr. Sm.	312	Pericentric/intercalary

Satellite	AT-content (%)	Distribution within the genus *Beta* Copy number in section: *Beta*	*Corollinae*	*Nanae*	*Procumbentes*	Reference
pBV1	69	High	n.d.	n.d.	n.d.	Schmidt and Metzlaff, 1991
pEV1	59	High	Low	n.d.	Middle	Schmidt *et al.*, 1991
pSV1	57	High	Middle	Middle	n.d.	Schmidt *et al.*, 1998*a*
pHT30	67	Low	High	Middle	n.d	Schmidt and Heslop-Harrison, 1993
pHT49	41	Low	High	Low	n.d.	
pHC28	43	High	High	Middle	Low	
pRN1	58	Low	Middle	High	n.d.	Kubis *et al.*, 1997
pTS5	70	n.d.	n.d.	Low	High	Schmidt and Heslop-Harrison, 1996*a*/*b*
pTS4.1	49	n.d.	n.d.	Low	High	

n.d., not detected by Southern hybridisation.

of mono- and di-nucleosomes (Hemleben, 1990; Traut, 1991). Satellite DNA repeats have been shown to interact with proteins and to be involved in nucleosomal phasing (Fischer *et al.*, 1994; Vershinin and Heslop-Harrison, 1998). Several examples of satellite DNAs with larger monomeric units have been found. Those monomers are often derived of complex rearrangements of smaller units and/or insertion and duplication events, as observed in *Anemone blanda* Schott & Ky. and Aveneae species (Hagemann, Scheer and Schweizer, 1993; Grebenstein *et al.*, 1996). Furthermore, higher order structures of monomeric repeats have been detected for many satellite DNA families (Vershinin, Schwarzacher and Heslop-Harrison, 1995; Schmidt *et al.*, 1998*a*). Some satellite monomers are composed of smaller units and may be derived from other classes of DNA sequence: for example, in *Brassica campestris* L., a centromeric tandem repeat is made up of three extensively diverged 60 bp units which are related to tRNA genes and SINEs (see below; Harrison and Heslop-Harrison, 1995).

Families of tandem repeats show varying levels of abundance and homology and distribution pattern between related species of a plant genus or family. They can exhibit species-, genome- and even chromosome-specificity (Zhao *et al.*, 1989; Preizner *et al.*, 1994; Wang *et al.*, 1995), and are therefore useful probes for studying taxonomic questions and phylogenetic relationships of plant species (Svitashev *et al.*, 1994; Nagaki *et al.*, 1995). Satellite DNA probes can be further helpful for the detection of hybrid species (Kamm *et al.*, 1995) and the selection of hybrid genomes, like somatic hybrids after protoplast fusion, chromosome-addition or translocation lines (Stadler *et al.*, 1995; Schmidt, Junghans and Metzlaff, 1990; Schmidt *et al.*, 1997).

Most tandem repeats are not routinely transcribed (Nagl and Schmitt, 1985), but occasionally transcripts are found: in rice, they have been shown to account for up to 3% of the total cellular RNA (Wu, Wang and Wu, 1994). We believe read-through of stop codons by RNA polymerase may account for this, and such transcription may be enhanced under stress conditions. Although the function of satellite DNA within the genome is unclear, the data can be interpreted to suggest that they play an important role in the stabilization and maintenance of chromosomal structures, and are involved in centromere formation and correct chromosome pairing during meiosis (Irick, 1994; Vig, 1994; Vershinin *et al.*, 1995; Csink and Henikoff, 1998). Tandem repeats are furthermore sites of recombination through crossing-over, which could lead to larger chromosomal rearrangements. Eberl, Duyf and Hilliker (1993) observed that a heterochromatic environment, constituted mainly by satellite DNA repeats, is essential for the full function of some genes in *Drosophila melanogaster*, and it is possible that such structures are present in plant species too.

Nine different satellite DNA families have been isolated from different *Beta* species, summarized in Table 1. Seven satellite repeats are abundant in sugar beet and related cultivars and wild beets of section *Beta* and four of them are highly amplified (Fig. 1A, F–H). It is unlikely that other highly amplified satellite DNA families are present in sugar beet, but additional families with low copy numbers could be detected. The satellite DNAs show variation in their

distribution, abundance, genomic organization and chromosomal localization between different sections of the genus *Beta*, reflecting differences in age and evolution of individual repeat families and phylogenetic relationships of *Beta* species. Some satellite DNA repeats can be used as species-specific probes. In sugar beet, the *Bam*HI satellite family is located at the centromeres of all 18 chromosomes as revealed by *in situ* hybridization (Schmidt *et al.*, 1998*b*). Other satellite DNAs (e.g. pEV1, pSV1; Fig. 1A and F) are clustered at intercalary sites, and chromosome specific variants of the size of repeat arrays have been observed.

DISPERSED DNA SEQUENCES

Retroelements

The term retroelement is used here as a short form of eukaryotic nonviral retroelement. The terminology of retroelements differs in the literature, but here the division of retroelements as defined by Hull and Covey (1996) will be followed. Retroelements are divided into three subgroups, depending on the presence and absence of different features. These subgroups are retrotransposons, retroposons and retrosequences; retroviruses are abundant in mammals, but not known in plants. The retrosequences comprise cDNAs and pseudogenes and cannot replicate autonomously. The major structural difference between retrotransposons and retroposons is the presence or absence of long terminal repeats (LTRs), respectively. Retrotransposons contain, like retroviruses, several open reading frames (ORFs), which code for specific proteins.

Dependent on the order of genes in the second ORF, two groups of retrotransposons are distinguished. One of them is the Ty3-*gypsy* retrotransposons, named after the first characterized elements of this group, Ty3 from *Saccharomyces cerevisiae* (Hansen, Chalker and Sandmeyer, 1988) and *gypsy* from *D. melanogaster* (Marlor, Parkhurst and Corces, 1986). In Ty3-*gypsy* elements the integrase domain is located downstream of reverse transcriptase and RNaseH, as found in retroviruses. Some elements feature a third ORF and therefore, Ty3-*gypsy* retrotransposons show the highest similarity to retroviruses. Several examples from plant species such as maize, *Sorghum bicolor* (L.) Moenchi, *Pinus radiata* and *Lilium henryi* have been described (for review see Bennetzen, 1996). Recently, Suoniemi, Tanskanen and Schulman (1998) illustrated the widespread presence of Ty3-*gypsy* elements in many plant species.

The other group is the Tyl-*copia* retrotransposons, named after the Tyl element from *S. cerevisiae* (Clare and Farabaugh, 1985) and *copia* from *D. melanogaster* (Mount and Rubin, 1985). Here the integrase domain is located upstream of the reverse transcriptase. Several full length elements have been described in plants (for review see Bennetzen, 1996), ranging in size from 4·8 kb (*Hopscotch* from *Zea mays*; White, Habera and Wessler, 1994) to more than 12 kb (BARE-1 from *Hordeum vulgare* L; Manninen and Schulmann, 1993). Studies based on a PCR assay for a part of the reverse transcriptase gene revealed their presence in all lineages of higher plants and green algae, showing them to be ubiquitous components of plant genomes (Flavell, Smith and Kumar, 1992*a*; Flavell *et al.*, 1992*b*; Voytas *et al.*, 1992; Hirochika and Hirochika, 1993; Lindauer *et al.*, 1993). Tyl-*copia* retroelements are mainly evenly distributed along plant chromosomes, but regions of depletion or amplification, differing between species, have been observed (Brandes *et al.*, 1997; Heslop-Harrison *et al.*, 1997). Members of Tyl-*copia* retrotransposons in plants feature a higher sequence divergence than observed in fungi or insects and many subfamilies of divergent elements are found (Flavell *et al.*, 1992*a*). Tyl-*copia* retroelements are amplified in plant genomes with copy numbers up to 1 million as found in *Vicia faba* L. (Pearce *et al.*, 1996). Most copies show stop codons and/or frame shifts, indicating that they are defective. The majority of plant Tyl-*copia* retrotransposons are transcriptionally inactive, but some examples of transcribed elements have been found (Manninen and Schulman, 1993; Suoniemi, Narvanto and Schulman, 1996; Grandbastien *et al.*, 1997). In tobacco and rice, actively amplifying elements have been identified (Grandbastien, Spielmann and Caboche, 1989; Hirochika *et al.*, 1996*a*, *b*).

Although a major component of the genome, retrotransposons can diverge between related species. In the hexaploid oat, *Avena sativa* L., Tyl-*copia* fragments isolated from the diploid progenitor-like species *A. strigosa* Schreb. and *A. clauda* Dur. and the tetraploid *A. vaviloviana* Malz. are able to distinguish the genomes when used as probes for *in situ* hybridization by their uniform labelling of chromosomes along their length (Katsiotis, Schmidt and Heslop-Harrison, 1996). Presumably all species had a common ancestor and the elements have diverged and amplified during evolution.

Retroposons are retroelements without LTRs and are exemplified by LINEs (long interspersed nuclear elements), best characterized in human and other mammals. In human there are an estimated 100000 copies of the LINE-family L1Hs and full length elements are about 7 kb in size (Hutchison *et al.*, 1989). LINEs can be frequently truncated at their 5′ ends as found for the majority of cin4 and L1Hs elements from maize and human (Schwarz-Sommer *et al.*, 1987, Hutchison *et al.*, 1989). LINE-like retroelements are well characterized in human, insects and other animal species. Plant LINEs include the cin4 element from *Z. mays*, *del2* from *Lilium speciosum* Thunb., BNR1 (partial sequence) from *Beta vulgaris*, Ta*11-1* from *A. thaliana* and Zepp from *Chlorella vulgaris* (Schwarz-Sommer *et al.*, 1987; Leeton and Smyth, 1993; Schmidt, Kubis and Heslop-Harrison, 1995; Wright *et al.*, 1996; Higashiyama *et al.*, 1997). Recent studies have investigated the distribution of the LINE-class of retroelements in plants, showing their presence in a wide range of plant species (Kubis *et al.*, 1998; Noma, Ohtsubo and Ohtsubo, 1998).

DNA sequences of the reverse transcriptase gene of two different groups of retroelements have been isolated from sugar beet by PCR (Schmidt *et al.*, 1995). Both LINEs and Tyl-*copia* retroelements are amplified in the sugar beet genome and show high levels of sequence divergence. The chromosomal distribution of the two types of retroelements is contrasting. Although both are excluded from centromeric and the 18S-5.8S-25S rDNA regions, the Tyl-*copia* retro-

transposons are distributed uniformly along chromosomes (Fig. 1F), whereas LINEs show an organization in discrete clusters (Fig. 1J). No analysis of LTR-retroelements of the Ty3-*gypsy* group have been performed to date in sugar beet.

Retroelements of different classes are abundant in plant genomes and show wide dispersion and a high variability indicating their contribution to the host genome organization, function and evolution (Bennetzen, 1996; Kumar, 1996; Flavell *et al.*, 1997). In maize, retrotransposons account for at least 50% of the DNA in large genomic regions and maybe the whole genome (SanMiguel *et al.*, 1996). In a 280 kb region around the maize *adh1* gene, retroelements were found in large blocks (> 50 kb) between single-copy gene sequences. Twenty families of retrotransposons were identified in the 280 kb region, with the five most abundant families comprising alone about 25% of the maize genome, some showing copy numbers of up to 30000 per haploid genome (SanMiguel *et al.*, 1996). In *Vicia faba*, Tyl-*copia*-like retroelements are estimated to account for at least 10% and possibly 40% of the genome (Pearce *et al.*, 1996; Kumar *et al.*, 1997). Similarly, the Tyl-*copia* *Bis*-1 family comprises about 5% of the wheat genome (Moore *et al.*, 1991) and the LINE *del2* about 4% of *Lilium speciosum* with about 250000 copies (Leeton and Smyth, 1993). These examples show that enormous copy numbers can be attained by retroelements, indicating that their amplification is one of the major factors for some very large plant genomes (Wessler, Bureau and White, 1995). Retroelements are not restricted to the nuclear genome but have also been found in the mitochondrial genome (Knoop *et al.*, 1996). Interestingly, the families of retrotransposons most abundant in the genome are not found within or next (within 25 to 50 kb) to genes (SanMiguel *et al.*, 1996), although retroelements or parts of them are frequently found in flanking regions of plant genes (White *et al.*, 1994).

SINEs (short interspersed nuclear elements) show some structural similarities to LINEs. They are 100–500 bp in length, exhibit homology to tRNAs or 7SL RNA at their 5′ end, but feature a polyA-tail like LINEs (Deininger, 1989; Smit, 1996). SINEs are very abundant in human, where the *Alu* family comprises about 500000 copies, representing 5% of the human genome. Studies in animal species revealed that the 3′ end of SINEs corresponds to the 3′ end of LINEs, but in plant species this relationship has not been investigated in detail (Okada and Hamada, 1997).

Non-retroelement dispersed repeats

The tandemly repeated DNA sequences and retroelements discussed above represent a major fraction of most plant genomes. However, there are also other repetitive DNA sequences which can be shown by *in situ* hybridization to be dispersed throughout the genome. For example, random repetitive clones in barley were found to localize over all chromosomes, although many were not homologous to retroelements (Busch *et al.*, 1995). Furthermore, the genome-specificity of the retroelements and dispersion of the tandemly repeated DNA classes together are not enough to account for the efficiency with which total genomic DNA—genomic *in situ* and Southern hybridization—can be used as a probe to distinguish closely related plant species by differential hybridization to chromosomes of different origins in hybrids or chromosomal addition and recombinant lines (see Schwarzacher *et al.*, 1992; Heslop-Harrison and Schwarzacher, 1996).

The pDRV sequence family in sugar beet is typical of non-retroelement repetitive DNA characterized by an interspersed genomic organization: the family is dispersed over all chromosomes of sugar beet with some regions of clustering and centromeric depletion (Schmidt *et al.*, 1998*a*). It is present in all sections of genus *Beta* with the highest amplification in sugar beet and other species of section *Beta*.

CONCLUSIONS

The question as to why particular repetitive DNA elements are amplified highly in some species but not in others (e.g. Table 1) remains unanswered. Earlier theories suggested that repetitive DNA elements are purely parasitic to the host genome and are able to amplify as long as they do not cause deleterious mutations or defects in essential gene sequences (Doolittle and Sapienza, 1980). But research over the last 15 years indicates that repetitive DNA elements, especially retroelements, have played a significant role in genome and species evolution, modulation of genes or gene expression, and in maintaining important structural features of chromosomes, such as the paracentromeric heterochromatin and telomeric/subtelomeric regions.

There is no reason to believe that plant genomes are now all larger than they ever have been. Indeed, some of the smaller genomes, such as that of *Arabidopsis thaliana* are probably reduced in size from their immediate ancestors. Nevertheless, where genome sizes are approximately the same in groups of related species, there are often considerable differences in the repetitive DNA sequences, leading to one sequence family being abundant in one species, but essentially absent in another (e.g. Table 1). As well as studies of the species distribution of individual repetitive elements, the experiments using genomic *in situ* hybridization show that the composition of most of the repetitive DNA may vary extensively between the species. One model might suggest that the common ancestor of each species group had a much smaller genome, and different sequences have been amplified during the speciation events. However, it would be surprising if no existing species in large groups such as the Triticeae tended to have such a 'primitive' character, so other mechanisms are more likely to account for the directional change in repetitive DNA composition of all chromosomes in individual species. We can note that slight differences in the average sizes of the DNA sequences coiled around nucleosomes can be detected between pairs of species such as wheat and rye (Vershinin and Heslop-Harrison, 1998). If the repeat length of one repetitive sequence were more stable than another in packing around a particular set of histones—including potentially species-specific modifications under genetic control such as acetylation—then amplification of one repeat class would be favoured over another. Where packing of the repetitive

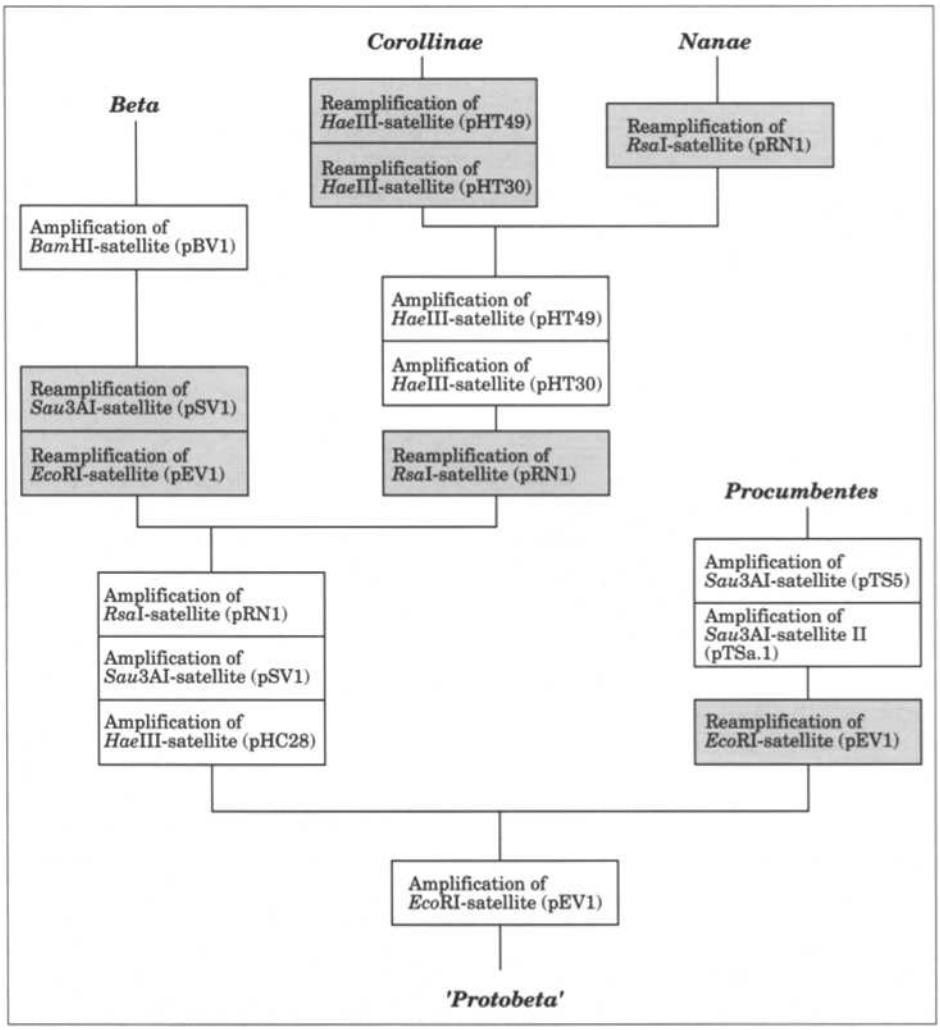

FIG. 2. Possible model of satellite repeat amplification based on the species-distribution analysis (Southern hybridization; Table 1) within the genus *Beta* and taking into account only amplification processes (no decrease or divergence of repeats). Shaded boxes indicate rounds of reamplification.

sequences were less stable, chromosome breakage would be slightly more frequent, leading to a selective disadvantage and having the consequence of directional turnover of repetitive sequence families. Interspecific hybrids often show increased chromosomal breakage and instability compared to species, and such unstable packing might contribute to this phenomenon.

Extensive studies on repetitive DNA elements have been performed on sugar beet and wild beet species. Elements belonging to both major groups of repetitive sequences (tandemly repeated and dispersed) have been identified. Using data on the species distribution of repetitive DNA families, we can examine which sequences have amplified in the different sections and suggest an order for amplification and re-amplification of the major sequence families (Fig. 2). In the future, it will be valuable to compare groups of related species with both contrasting genome sizes and fully characterized repetitive DNA sequence composition. The genus *Vicia*, where we have contrasts but no correlation between genome size and Tyl-*copia*-like retrotransposon copy number (Pearce *et al.*, 1996) may be a suitable candidate. The data for sugar beet, especially from *in situ* hybridization experiments (and including data presented here), led to the proposal of a generalized chromosome model showing the various repetitive elements along the chromosome arms, with clusters of genes between some of the regions of tandemly repeated DNA (Schmidt and Heslop-Harrison, 1998). The comparison of genome sizes and sequence distribution both within chromosomes and between species is already adding to our knowledge of the biological significance of genome size variation, and the importance of genome size in plant evolution and speciation.

ACKNOWLEDGEMENTS

S. Kubis acknowledges a studentship from the DeMontfort University, Leicester. T. Schmidt is supported by the Deutsche Forschungsgemeinschaft (Schm 1048/2-1 and Schm 1048/2-2) and by a grant of the Fonds der Chemischen Industrie (0657394). This work was supported by the ARC grant 313-ARC-XI-97/45 of the DAAD and the British Council. Retrotransposon work was supported by EU project B104-CT96-0508.

LITERATURE CITED

Arumuganathan K, Earle ED. 1991. Nuclear DNA content of some important plant species. *Plant Molecular Biology Reporter* **9**: 208–218.

Barocka KH. 1985. Zucker- und Futterrüben. In: Hoffmann W, Mudra A, Plarre W, eds. *Lehrbuch der Züchtung landwirtschaftlicher Kulturpflanzen*. Berlin and Hamburg: Parey, 245–287.

Barzen E, Mechelke W, Ritter E, Schulte-Kappert E, Salamini F. 1995. An extended map of the sugar beet genome containing RFLP and RAPD loci. *Theoretical and Applied Genetics* **90**: 189–193.

Barzen E, Mechelke W, Ritter E, Seitzer JF, Salamini F. 1992. RFLP markers for sugar beet breeding: chromosomal linkage maps and location of major genes for rhizomania resistance, monogermy and hypocotyl colour. *Plant Journal* **2**: 601–611.

Bedbrook JR, Jones J, O'Dell M, Thompson RD, Flavell RB, 1980. A molecular description of telomeric heterochromatin in *Secale* species. *Cell* **19**: 545–560.

Bennett MD, Leitch IJ. 1995. Nuclear DNA amounts in Angiosperms. *Annals of Botany* **76**: 113–176.

Bennett MD, Smith JB, 1976. Nuclear DNA amounts in angiosperms. *Philosophical Transaction of the Royal Society of London Series B* **274**: 227–274.

Bennetzen JL. 1996. The contribution of retroelements to plant genome organization, function and evolution. *Trends in Microbiology* **4**: 347–353.

Brandes A, Heslop-Harrison JS, Kamm A, Kubis S, Doudrick RL, Schmidt T. 1997. Comparative analysis of the chromosomal and genomic organization of *Tyl-copia*-like retrotransposons in pteridophytes, gymnosperms and angiosperms. *Plant Molecular Biology* **33**: 11–21.

Busch W, Martin R, Herrmann RG, Hohmann U. 1995. Repeated DNA sequences isolated by microdissection. I. Karyotyping of barley (*Hordeum vulgare* L.). *Genome* **38**: 1082–1090.

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

Clare J, Farabaugh P. 1985. Nucleotide sequence of a yeast *Ty* element: evidence for an unusual mechanism of gene expression. *Proceedings of the National Academy of Sciences of the USA* **82**: 2829–2833.

Coen ES, Carpenter R. 1986. Transposable elements in *Antirrhinum majus*: generators of genetic diversity. *Trends in Genetics* **2**: 292–296.

Csink AK, Henikoff S. 1998. Something from nothing: the evolution and utility of satellite repeats. *Trends in Genetics* **14**: 200–204.

Deininger PL. 1989. SINEs: Short interspersed repeated DNA elements in higher eucaryotes. In: Berg DE, Howe MM, eds. *Mobile DNA*. Washington DC: American Society for Microbiology, 619–636.

Del-Favero J, Vauterin M, Weyens G, Edwards KE, Jacobs M. 1994. Construction and characterisation of a yeast artificial chromosome library containing five haploid sugarbeet (*Beta vulgaris* L.) genome equivalents. *Theoretical and Applied Genetics* **88**: 449–453.

Devos KM, Millan T, Gale MD. 1993. Comparative RFLP maps of the homoeologous group-2 chromosomes of wheat, rye and barley. *Theoretical and Applied Genetics* **85**: 784–792.

Doolittle WF, Sapienza C. 1980. Selfish genes, the phenotype paradigm and genome evolution. *Nature* **284**: 601–603.

Dubcovsky J, Dvorak J. 1995. Ribosomal RNA multigene loci: Nomads of the triticeae genomes. *Genetics* **140**: 1367–1377.

Eberl DF, Duyf BJ, Hilliker AJ. 1993. The role of heterochromatin in the expression of a heterochromatic gene, the rolled locus of *Drosophila melanogaster*. *Genetics* **134**: 277–292.

Eyers M, Edwards K, Schuch W. 1992. Construction and characterisation of a yeast artificial chromosome library containing two haploid *Beta vulgaris* L. genome equivalents. *Gene* **121**: 195–201.

Fischer TC, Groner S, Zentgraf U, Hemleben V. 1994. Evidence for nucleosomal phasing and a novel protein specifically binding to cucumber satellite DNA. *Zeitschrift für Naturforschung* **49**: 79–86.

Flavell AJ, Smith DB, Kumar A. 1992*a*. Extreme heterogenity of Tyl-*copia* group retrotransposons in plants. *Molecular and General Genetics* **231**: 233–242.

Flavell AJ, Pearce SR, Heslop-Harrison JS, Kumar A. 1997. The evolution of Tyl-*copia* group retrotransposons in eukaryote genomes. *Genetica* **100**: 185–195.

Flavell AJ, Dunbar E, Anderson R, Pearce SR, Hartley R, Kumar A. 1992*b*. *Tyl-copia* group retrotransposons are ubiquitous and heterogeneous in higher plants. *Nucleic Acids Research* **20**: 3639–3644.

Flavell RB, Bennett MD, Smith JB, Smith DB. 1974. Genome size and the proportion of repeated nucleotide sequence DNA in plants. *Biochemical Genetics* **12**: 257–268.

Fuchs J, Brandes A, Schubert I. 1995. Telomere sequence localization and karyotype evolution in higher plants. *Plant Systematics and Evolution* **196**: 227–241.

Gierl A, Cuypes H, Lütticke S, Pereira A, Schwarz-Sommer Z. 1988. Structure and function of the En/Spm transposable element system of *Zea mays*: identification of the Suppressor component of En. In: Nelson, O., ed. *Proceedings of the international symposium on plant transposable elements*. New York: Plenum, 115–119.

Grandbastien M-A, Spielmann A, Caboche M. 1989. Tnt1, a mobile retroviral-like transposable element of tobacco isolated by plant cell genetics. *Nature* **337**: 376–380.

Grandbastien M-A, Lucas H, Morel J-B, Mhiri, C, Vernhettes S, Casacuberta JM. 1997. The expression of the tobacco Tnt1 retrotransposon is linked to plant defence responses. *Genetica* **100**: 241–252.

Grebenstein B, Grebenstein O, Sauer W, Hemleben V. 1996. Distribution and complex organization of satellite DNA sequences in Aveneae species. *Genome* **39**: 1045–1050.

Hagemann S, Scheer B, Schweizer D. 1993. Repetitive sequences in the genome of *Anemone blanda*: Identification of tandem arrays and of dispersed repeats. *Chromosoma* **102**: 312–324.

Hallden C, Hjerdin A, Rading IM, Säll T, Fridlundh B, Johannisdottir G, Tuvesson S, Akesson C, Nilsson N-O. 1996. A high density RFLP linkage map of sugar beet. *Genome* **39**: 634–645.

Hansen LJ, Chalker DL, Sandmeyer SB. 1988. Ty3, a yeast retrotransposon associated with tRNA genes, has homology to animal retroviruses. *Molecular Cell Biology* **8**: 5245–5256.

Harrison GE, Heslop-Harrison JS. 1995. Centromeric repetitive DNA in the genus *Brassica*. *Theoretical and Applied Genetics* **90**: 157–165.

Hemleben V. 1990. *Molekularbiologie der Pflanzen*. Stuttgart: Gustav Fischer Verlag.

Heslop-Harrison JS. 1991. The molecular cytogenetics of plants. *Journal of Cell Science* **100**: 15–21.

Heslop-Harrison JS, Schwarzacher T. 1996. Genomic Southern and *in situ* hybridization for plant genome analysis. In: Jauhar PP, ed. *Methods of genome analysis in plants*. Boca Raton: CRC, 163–179.

Heslop-Harrison JS, Brandes A, Taketa S, Schmidt T, Vershinin AV, Alkhimova EG, Kamm A, Doudrick RL, Schwarzacher T, Katsiotis A, Kubis S, Kumar A, Pearce SR, Flavell AJ, Harrison GE. 1997. The chromosomal distributions of Tyl-*copia* group retrotransposable elements in higher plants and their implications for genome evolution. *Genetica* **100**: 197–204.

Higashiyama T, Noutoshi Y, Fujie M, Yamada T. 1997. Zepp, a LINE-like retrotransposon accumulated in the *Chlorella* telomeric region. *EMBO Journal* **16**: 3715–3723.

Hirochika H, Hirochika R. 1993. *Tyl-copia* group retrotransposons as ubiquitous components of plant genomes. *Japanese Journal of Genetics* **68**: 35–46.

Hirochika H, Sugimoto K, Ossuki Y, Tsugaawa H, Kanda W. 1996*a*. Retrotransposons of rice involved in mutations induced by tissue culture. *Proceedings of the National Academy of Sciences of the USA* **93**: 7783–7788.

Hirochika H, Otsuki H, Yoshikawa M, Otsuki Y, Sugimoto K, Takeda S. 1996*b*. Autonomous transposition of the tobacco retrotransposon *Tto1* in rice. *Plant Cell* **8**: 725–734.

Hull R, Covey SN. 1996. Retroelements: propagation and adaptation. *Virus Genes* **11**: 105–118.

Hutchison CA, Hardies SC, Loeb DD, Shehee WR, Edgell MH. 1989. LINEs and related retrotransposons: long interspersed repeated sequences in the eukaryotic genome. In: Berg DE, Howe MM, eds. *Mobile DNA*. Washington DC: American Society for Microbiology, 593–617.

Irick H. 1994. A new function for heterochromatin. *Chromosoma* **103**: 1–3.

Kamm A, Galasso I, Schmidt T, Heslop-Harrison JS. 1995. Analysis of a repetitive DNA family from *Arabidopsis arenosa* and relationships between *Arabidopsis* species. *Plant Molecular Biology* **27**: 853–862.

Katsiotis A, Schmidt T, Heslop-Harrison JS. 1996. Chromosomal and genomic organization of Tyl-*copia*-like retrotransposon sequences in the genus *Avena*. *Genome* **39**: 410–417.

Kleine M, Cai D, Eibl C, Hermann RG, Jung C. 1995. Physical mapping and cloning of a translocation in sugar beet (*Beta vulgaris* L.) carrying a gene for nematode (*Heterodera schachtii*) resistance from *B. procumbens*. *Theoretical and Applied Genetics* **90**: 399–406.

Knoop V, Unseld M, Marienfeld J, Brandt P, Sünkel S, Ullrich H, Brennicke A. 1996. *copia*-, *gypsy*- and LINE-like retrotransposon fragments in the mitochondrial genome of *Arabidopsis thaliana*. *Genetics* **142**: 579–585.

Kubis S, Heslop-Harrison JS, Schmidt T. 1997. A family of differentially amplified repetitive DNA sequences in the genus *Beta* reveals genetic variation in *Beta vulgaris* subspecies and cultivars. *Journal of Molecular Evolution* **44**: 310–320.

Kubis SE, Heslop-Harrison JS, Desel C, Schmidt T. 1998. The genomic organization of non-LTR retrotransposons (LINEs) from three *Beta* species and five other angiosperms. *Plant Molecular Biology* **36**: 821–831.

Kumar A. 1996. The adventures of the Tyl-*copia* group of retrotransposons in plants. *Trends in Genetics* **12**: 41–43.

Kumar A, Pearce SR, McLean K, Harrison G, Heslop-Harrison JS, Waugh R, Flavell AJ. 1997. The Tyl-*copia* group of retrotransposons in plants: genomic organisation, evolution, and use as molecular markers. *Genetica* **100**: 205–217.

Laurie DA, Pratchett N, Devos KM, Leitch IJ, Gale MD. 1993. The distribution of RFLP markers on chromosome 2(2H) of barley in relation to the physical and genetic location of 5S rDNA. *Theoretical and Applied Genetics* **87**: 177–183.

Leeton PRJ, Smyth DR. 1993. An abundant LINE-like element amplified in the genome of *Lilium speciosum*. *Molecular and General Genetics* **237**: 97–104.

Leitch IJ, Heslop-Harrison JS. 1993. Physical mapping of four sites of 5S rDNA sequences and one site of the alpha-amylase-2 gene in barley (*Hordeum vulgare*). *Genome* **36**: 517–523.

Lindauer A, Fraser D, Brüderlein M, Schmitt R. 1993. Reverse transcriptase families and a *copia*-like retrotransposon, *Osser*, in the green alga *Volvox carteri*. *FEBS Letters* **319**: 261–266.

McClintock B. 1951. Chromosome organisation and genic expression. *Cold Spring Harbor Symposium on Quantitative Biology* **16**: 13–47.

McClintock B. 1961. Further studies on the *Suppressor-mutator* system of control of gene action in maize. *Carnegie Institute of Washington Yearbook* **60**: 469–476.

Manninen I, Schulman AH. 1993. *BARE-1*, a *copia*-like retroelement in barley (*Hordeum vulgare* L.) *Plant Molecular Biology* **22**: 829–846.

Marlor RL, Parkhurst SM, Corces VG. 1986. The *Drosophila melanogaster* gypsy transposable element encodes putative gene products homologous to retroviral protein. *Molecular and Cellular Biology* **6**: 1129–1134.

Moore G, Lucas H, Batty N, Flavell R. 1991. A family of retrotransposons and associated genomic variation in wheat. *Genomics* **10**: 461–468.

Mount SM, Rubin GM. 1985. Complete nucleotide sequence of the *Drosophila* transposable element *copia*: Homology between *copia* and retroviral proteins. *Molecular and Cellular Biology* **5**: 1630–1638.

Nagaki K, Tsujimoto H, Isono K, Sasakuma T. 1995. Molecular characterisation of a tandem repeat, Afa family, and its distribution among Triticeae. *Genome* **38**: 479–486.

Nagl W, Schmitt H-P. 1985. Transcription of repetitive DNA in condensed plant chromatin. *Molecular Biology Reporter* **10**: 143–146.

Noma K, Ohtsubo H, Ohtsubo E. 1998. Non-LTR retrotransposons (LINEs) as ubiquitous components among plant kingdom. *Plant and Animal Genome VI, San Diego, Abstracts guide*, 103.

Okada N, Hamada M. 1997. The 3′ ends of tRNA-derived SINEs originated from 3′ ends of LINEs: A new example from the bovine genome. *Journal of Molecular Evolution* **44**: S52–S56.

Pearce SR, Harrison G, Li D, Heslop-Harrison JS, Flavell A, Kumar A. 1996. The *Tyl-copia* group retrotransposons in *Vicia* species: copy number, sequence heterogeneity and chromosomal localization. *Molecular and General Genetics* **250**: 305–315.

Pech M, Igo-Kemenes T, Zachau HG. 1979. Nucleotide sequence of a highly repetitive component of rat DNA. *Nucleic Acids Research* **7**: 417–432.

Pillen K, Steinrücken G, Wricke G, Herrmann RG, Jung C. 1992. A linkage map of sugar beet (*Beta vulgaris* L.). *Theoretical and Applied Genetics* **84**: 129–135.

Pillen K, Steinrücken G, Wricke G, Herrmann RG, Jung C. 1993. An extended linkage map of sugar beet (*Beta vulgaris* L.) including nine putative lethal genes and the restorer gene X. *Plant Breeding* **111**: 265–272.

Preizner J, Takács I, Bilgin M, Györgyey J, Dudits D, Fehér A. 1994. Organization of a *Solanum brevidens* repetitive sequence related to the TGRI subtelomeric repeats of *Lycopersicon esculentum*. *Theoretical and Applied Genetics* **89**: 1–8.

SanMiguel P, Bennetzen JL. 1998. Evidence that a recent increase in maize genome size was caused by the massive amplification of intergene retrotransposons. *Annals of Botany* **82** (Supp. A.): 37–44.

SanMiguel P, Tikhonov A, Jin Y-K, Motchoulskaia N, Zakharov D, Melake-Berhan A, Springer PS, Edwards KJ, Lee M, Avramova Z, Bennetzen JL. 1996. Nested retrotransposons in the intergenic regions of the maize genome. *Science* **274**: 765–768.

Santoni S, Berville A. 1992. Characterization of the nuclear ribosomal DNA units and phylogeny of *Beta* L. wild forms and cultivated beets. *Theoretical and Applied Genetics* **83**: 533–542.

Schlötterer C, Tautz D. 1992. Slippage synthesis of simple sequence DNA. *Nucleic Acids Research* **20**: 211–215.

Schmidt T, Heslop-Harrison JS. 1993. Variability and evolution of highly repeated DNA sequences in the genus *Beta*. *Genome* **36**: 1074–1079.

Schmidt T, Heslop-Harrison JS. 1996*a*. High-resolution mapping of repetitive DNA by *in situ* hybridization: molecular and chromosomal features of prominent dispersed and discretely localized DNA families from the wild beet species *Beta procumbens*. *Plant Molecular Biology* **30**: 1099–1114.

Schmidt T, Heslop-Harrison JS. 1996*b*. The physical and genomic organization of microsatellites in sugar beet. *Proceedings of the National Academy of Sciences of the USA* **93**: 8761–8765.

Schmidt T, Heslop-Harrison JS. 1998. Genomes, genes and junk: the large-scale organization of plant genomes. *Trends in Plant Science* **3**: 195–199.

Schmidt T, Metzlaff M. 1991. Cloning and characterization of a *Beta vulgaris* satellite DNA family. *Gene* **101**: 247–250.

Schmidt T, Jung C, Metzlaff M. 1991. Distribution and evolution of two satellite DNAs in the genus *Beta*. *Theoretical and Applied Genetics* **82**: 793–799.

Schmidt T, Junghans H, Metzlaff M. 1990. Construction of *Beta procumbens*-specific DNA probes and their application for the screening of *B. vulgaris* × *B. procumbens* ($2n = 19$) addition lines. *Theoretical and Applied Genetics* **79**: 177–181.

Schmidt T, Kubis S, Heslop-Harrison JS. 1995. Analysis and chromosomal localization of retrotransposons in sugar beet (*Beta vulgaris* L.): LINEs and *Tyl-copia*-like elements as major components of the genome. *Chromosome Research* **3**: 335–345.

Schmidt T, Schwarzacher T, Heslop-Harrison JS. 1994. Physical mapping of rRNA genes by fluorescent *in situ* hybridization and structural analysis of 5S rRNA genes and intergenic spacer sequences in sugar beet (*Beta vulgaris*). *Theoretical and Applied Genetics* **88**: 629–636.

Schmidt T, Jung C, Heslop-Harrison JS, 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 Research* **5**: 186–193.

Schmidt T, Kubis S, Katsiotis A, Jung C, Heslop-Harrison JS. 1998*a*. Molecular and chromosomal organization of two repetitive DNA

sequences with intercalary locations in sugar beet and other *Beta* species. *Theoretical and Applied Genetics* (in press).

Schmidt T, Boblenz K, Metzlaff M, Kaemmer D, Weising K, Kahl G. 1993. DNA fingerprinting in sugar beet (*Beta vulgaris*) – identification of double-haploid breeding lines. *Theoretical and Applied Genetics* **85**: 653–657.

Schmidt T, Kubis S, Doudrick RL, Kamm A, Brandes A, Desel C, Heslop-Harrison JS. 1998*b*. FISHing DNA sequences – fluorescent *in situ* hybridisation in plant genome analysis. *Scanning Microscopy* (in press).

Schondelmaier J, Steinrücken G, Jung C. 1996. Integration of AFLP markers into a linkage map of sugar beet (*Beta vulgaris* L.). *Plant Breeding* **115**: 231–237.

Schondelmaier J, Schmidt T, Heslop-Harrison JS, Jung C. 1997. Genetic and chromosomal localization of the 5S rDNA locus in sugar beet (*Beta vulgaris* L.). *Genome* **40**: 171–175.

Schwarzacher T, Heslop-Harrison JS. 1991. *In situ* hybridization to plant telomeres using synthetic oligomers. *Genome* **34**: 317–33.

Schwarzacher T, Anamthawat-Jónsson K, Harrison GE, Islam AKMR, Jia JZ, King IP, Leitch AR, Miller TE, Reader SM, Rogers WJ, Shi M, Heslop-Harrison JS. 1992. Genomic *in situ* hybridization to identify alien chromosomes and chromosome segments in wheat. *Theoretical and Applied Genetics* **84**: 778–786.

Schwarz-Sommer Z, Leclercq L, Goebel E, Saedler H. 1987. Cin4, an insert altering the structure of the *A1* gene in *Zea mays*, exhibits properties of nonviral retrotransposons. *EMBO Journal* **13**: 3873–3880.

Smit AFA. 1996. The origin of interspersed repeats in the human genome. *Current Opinion in Genetics and Development* **6**: 743–748.

Stadler M, Stelzer T, Borisjuk N, Zanke C, Schilde-Rentschler L, Hemleben V. 1995. Distribution of novel and known repeated elements of *Solanum* and application for the identification of somatic hybrids among *Solanum* species. *Theoretical and Applied Genetics* **91**: 1271–1278.

Suoniemi A, Narvanto A, Schulman AH. 1996. The *BARE-1* retrotransposon is transcribed in barley from an LTR promoter active in transient assays. *Plant Molecular Biology* **31**: 295–306.

Suoniemi A, Tanskanen J, Schulman AH. 1998. Gypsy-like retrotransposons are widespread in the plant kingdom. *Plant Journal* **13**: 699–705.

Svitashev S, Bryngelsson T, Vershinin A, Pedersen C, Säll T, Bothmer R von. 1994. Phylogenetic analysis of the genus *Hordeum* using repetitive DNA sequences. *Theoretical and Applied Genetics* **89**: 801–810.

Thomas CA. 1971. The genetic organisation of chromosomes. *Annual Review of Genetics* **5**: 237–256.

Traut W. 1991. *Chromosomen. Klassische und molekulare Zytogenetik*. Berlin, Heidelberg: Springer-Verlag.

Vershinin AV, Heslop-Harrison JS. 1998. Comparative analysis of the nucleosomal structure of rye, wheat and their relatives. *Plant Molecular Biology* **36**: 149–161.

Vershinin AV, Schwarzacher T, Heslop-Harrison JS. 1995. The large-scale genomic organization of repetitive DNA families at the telomeres of rye chromosomes. *Plant Cell* **7**: 1823–1833.

Vig BK. 1994. Do specific nucleotide bases constitute the centromere? *Mutation Research* **309**: 1–10.

Voytas DF, Cummings MP, Konieczny A, Ausubel FM, Rodermel SR. 1992. *Copia*-like retrotransposons are ubiquitous among plants. *Proceedings of the National Academy of Sciences of the USA* **89**: 7124–7128.

Wang ZX, Kurata N, Saji S, Katayose Y, Minobe Y. 1995. A chromosome 5-specific repetitive DNA-sequence in rice (*Oryza sativa* L.) *Theoretical and Applied Genetics* **90**: 907–913.

Wessler SR, Bureau TE, White SE. 1995. LTR-retrotransposons and MITEs: important players in the evolution of plant genomes. *Current Opinion in Genetics and Development* **5**: 814–821.

White SE, Habera LF, Wessler SR. 1994. Retrotransposons in the flanking regions of normal plant genes: A role for *copia*-like elements in the evolution of gene structure and expression. *Proceedings of the National Academy of Sciences of the USA* **91**: 11792–11796.

Wright DA, Ke N, Smalle J, Hauge BM, Goodman HM, Voytas DF. 1996. Multiple non-LTR retrotransposons in the genome of *Arabidopsis thaliana*. *Genetics* **142**: 569–578.

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

Zhao X, Wu T Xie Y, Wu R. 1989. Genome-specific repetitive sequences in the genus *Oryza*. *Theoretical and Applied Genetics* **78**: 201–209.