PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Repetitive DNA and Chromosome Evolution in Plants

R. B. Flavell

Phil. Trans. R. Soc. Lond. B 1986 **312**, 227-242
doi: 10.1098/rstb.1986.0004

Phil. Trans. R. Soc. Lond. B **312**, 227–242 (1986)
Printed in Great Britain

Repetitive DNA and chromosome evolution in plants

By R. B. Flavell
Plant Breeding Institute, Trumpington, Cambridge CB2 2LQ, U.K.

Most higher plant genomes contain a high proportion of repeated sequences. Thus repetitive DNA is a major contributor to plant chromosome structure. The variation in total DNA content between species is due mostly to variation in repeated DNA content. Some repeats of the same family are arranged in tandem arrays, at the sites of heterochromatin. Examples from the *Secale* genus are described. Arrays of the same sequence are often present at many chromosomal sites. Heterochromatin often contains arrays of several unrelated sequences. The evolution of such arrays in populations is discussed. Other repeats are dispersed at many locations in the chromosomes. Many are likely to be or have evolved from transposable elements. The structures of some plant transposable elements, in particular the sequences of the terminal inverted repeats, are described. Some elements in soybean, antirrhinum and maize have the same inverted terminal repeat sequences. Other elements of maize and wheat share terminal homology with elements from yeast, *Drosophila*, man and mouse. The evolution of transposable elements in plant populations is discussed. The amplification, deletion and transposition of different repeated DNA sequences and the spread of the mutations in populations produces a turnover of repetitive DNA during evolution. This turnover process and the molecular mechanisms involved are discussed and shown to be responsible for divergence of chromosome structure between species. Turnover of repeated genes also occurs.

The molecular processes affecting repeats imply that the older a repetitive DNA family the more likely it is to exist in different forms and in many locations within a species. Examples to support this hypothesis are provided from the *Secale* genus.

Introduction

Chromosomes are the vehicles in which the genes are replicated and moved during cell division to ensure that each cell of an individual has the same complement of genes. At meiosis, the chromosomes undergo pairing and recombination to enable new gene combinations to be created and to facilitate regular disjunction. The chromosomes therefore must retain, during evolution, the ability to be replicated faithfully, to condense, undergo mitosis, decondense, and to recombine. The gene functions of chromosomes must also be conserved during evolution. The amount of DNA specifying all these properties of chromosomes has been estimated to occupy about 10^7–10^8 base pairs (Flavell 1980). Whatever the precise value it is relatively small compared with the size of most plant genomes. The anomaly between the minimal essential DNA content and that commonly accumulated in plant genomes is further highlighted by the fact that DNA contents range from 5×10^7 to over 8×10^{10} base pairs among flowering plants (Bennett & Smith 1976). Some of this variation is due to polyploidy but most of the DNA in each haploid chromosomal set appears to be in excess of the minimum required. The sequences of this 'excess' or 'secondary' DNA (Hinegardner 1976) appear not to be highly conserved during evolution, in contrast to the genes, as might be expected if they are not intimately involved in *determining* specific chromosomal functions. Much of the secondary DNA consists

of sequences highly repeated in the genome or has evolved from them (Flavell 1980, 1982*b*; Thompson & Murray 1981). The fraction of repetitive DNA is so high in many plant genomes (it has been estimated to be over 95% of the total DNA in pea for example (Thompson & Murray 1980)) that it clearly plays a dominant role in determining chromosome size and structure. Changes in repeated DNA are responsible for most of the changes in chromosome size and structure during the evolution and divergence of species (Flavell 1983).

The large fraction of highly repeated DNA in plant genomes emphasizes that amplification processes are responsible for creating much of the nuclear DNA. These processes together with the ways in which the amplified sequences spread and become fixed in the chromosomes have a major influence on chromosome biology. Repeated sequences belonging to the same family are scattered throughout plant chromosomes (Flavell 1980) and it is a major scientific challenge to understand the mechanisms that are responsible for the vast number of DNA amplifications and transpositions that take place during evolution. One such mechanism, emerging from studies on genetically defined transposable elements is discussed later. However, first I discuss repeats that are arranged in tandem arrays, each array often containing tens of thousands of copies.

Amplification and tandem arrays of repeats

Tandem arrays of repeats are almost certainly present in all plant genomes. The unit sequence that is amplified can be from a few base pairs, for example, GAAGAA/G which is amplified in wheat and related species, to many thousands of base pairs (Dennis *et al.* 1980; Bedbrook *et al.* 1980*b*). The precise mechanisms of initial amplification in specific cases are unknown but amplification by many rounds of unequal crossing over between duplicated sequences (Smith 1976), a 'rolling circle' mechanism (Hourcade *et al.* 1973), 'slippage replication' (Tautz & Renz 1984) or 'aberrant *in situ* replication' (Schimke 1982) have been proposed. All are capable of producing a tandem array of a single sequence. For the rapid production of large tandem arrays the latter three mechanisms are more appealing and the studies on the amplification of sequences associated with drug resistance in animal cells confirm that rapid massive amplification can occur in a single or a few cell generations and this kind of mutation may be very common (Bostock & Tyler-Smith 1982; Schimke 1982).

Tandem arrays of essentially the same sequence are often found on many, if not all, chromosomes of a species as illustrated for *Secale cereale* in figure 1. This has been shown either by hybridization of the purified sequence to metaphase chromosomes (for example, Bedbrook *et al.* 1980*a*; Jones & Flavell 1982*a*, *b*; Deumling & Greilhuber 1982; Gerlach & Peacock 1980) or by hybridization of the sequence to a series of DNAs isolated from plants containing only a single chromosome of one species in addition to the full complement of chromosomes from another species (Bedbrook *et al.* 1980*a*, *b*; Jones & Flavell, 1982*a*). The dispersed distribution of arrays of repeats illustrates that arrays are duplicated or divided and translocated between chromosomes during evolution. Such translocations may have played a vital role in the fixation of arrays of repeats in a species. If an array, or segments of it, in an individual is frequently translocated to another chromosome and the new genotypes are not eliminated by selection, then the arrays of the sequence will increase in the population relatively rapidly. This is due to the meiotic segregation of homologues and non-homologues bearing the new sequences to different individuals in each generation (Dover 1982; Ohta & Dover 1983). Ohta & Dover

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

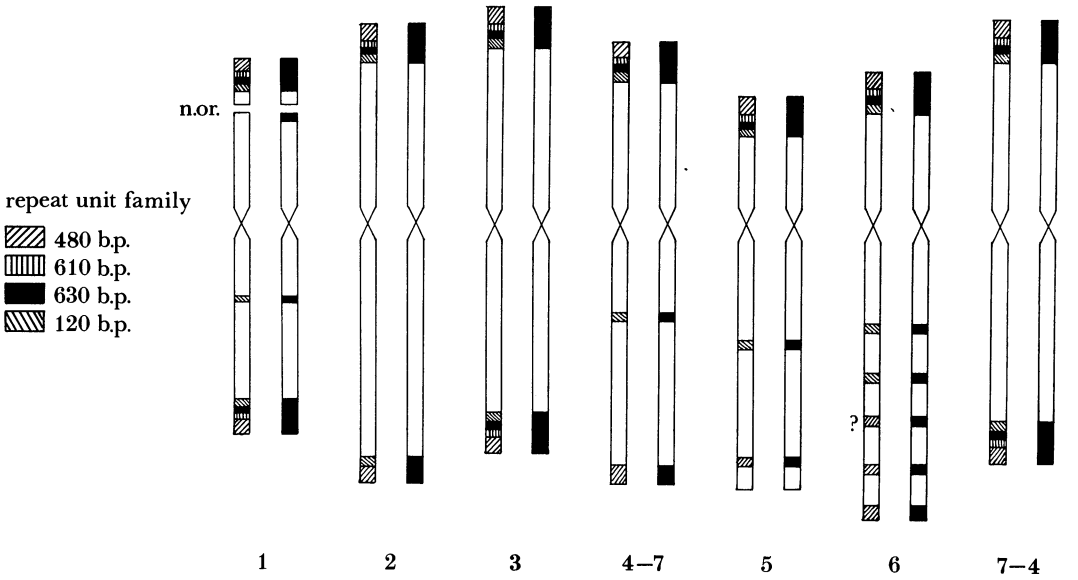


FIGURE 1. Schematic summary of the localization of the major tandem arrays of repeats in the chromosomes of *Secale cereale* cultivar King II. The results were gained by hybridizing representative repeats of the four families, purified by molecular cloning, to metaphase chromosomes as described in Jones & Flavell (1982*a*). Individual chromosomes were identified by hybridization to wheat lines containing single rye chromosomes (Jones & Flavell 1982*a*). In the right-hand member of each chromosome pair ■ denotes C bands revealed by Giemsa staining. The arrangement of the arrays within each telomeric block may differ from that illustrated (see Jones & Flavell 1982*a*). ? implies that this hybridization site differs between cultivars. n.or., Nucleolus organizer.

(1984) have pointed out that in these circumstances the population can change with respect to the repeat family in a 'cohesive' manner.

Some tandem arrays of repeats may be predisposed to translocation because of the sequence of the repeat unit. Alternatively, the tandem arrangement may facilitate the deletion of circular arrays by intrachromosomal recombination and these arrays could insert elsewhere into the genome (Flavell 1985). A third possibility is that sequence translocations occur often by chance, not as a result of the sequence or its arrangement, and because the event is not lethal the arrays spread in the population rapidly. Whatever mechanisms are behind the appearance of arrays of repeats on all chromosomes it is probable that the multiple locations in individuals are connected with the relatively rapid spread of these arrays in populations during species divergence. These lines of argument ignore the possibility that selection has played a major role in the fixation of tandem arrays. To evaluate this possibility it would be helpful to know if individual arrays have a function or an effect on the individuals that carry them. Like all DNA sequences, these repeats contribute to genome and chromosome size (see later) but rarely are they transcribed.

The major tandem arrays are often localized in a similar position on many of the chromosomes of a species, for example, at the telomeres or around the centromere (see figure 1). In wheat and rye (see figure 1) the major blocks of repeats have been shown to contain several arrays of unrelated repeats (Gerlach & Peacock 1980; Jones & Flavell 1982*a*). These observations suggest that there are specific positions where arrays of repeats are favoured or tolerated. Alternatively, they may accumulate preferentially at the sites where non-homologous and homologous chromosomes interact physically: an interaction that could facilitate the

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

recombination necessary for transposition (Flavell 1983). Some evidence in favour of this is the observation in species of *Scilla* that when chromosomes are ordered, such that chromosome arms of similar length lie adjacent to one another, a model proposed by Bennett (1982) and for which experimental support has been gained, then arrays of repeats in heterochromatin lie at similar positions on adjacent chromosomes (Greilhuber & Loidl 1983). Although in *Secale cereale* major arrays of 480 b.p. repeats lie at almost all telomeres (figure 1), this same sequence shows a much less uniform distribution between chromosomes in *Secale montanum* in which some accessions have a major array only on one or two pairs of chromosomes (Jones & Flavell 1982*b*). This suggests that selection may be actively determining the distribution of this particular array in the *Secale* genus.

From studies in *Drosophila*, Yamamoto & Miklos (1978) concluded that blocks of tandem arrays of repeats localized in heterochromatin suppressed recombination, causing chiasmata to occur much more frequently at locations distal to the repeats. The manipulation of the position of recombination clearly could greatly affect the extent to which linkage is maintained between alleles in a population and thus could greatly influence the population genetics of the species. In a recent study of recombination on the short arm of chromosome 1B of wheat (Snape *et al.* 1985) it was found that recombination occurred predominantly in the distal third of the chromosome and rarely in the proximal one-third adjacent to the centromere. This latter segment contains large blocks of tandem arrays of repeats. It is possible, therefore, that as in *Drosophila* and *Atractomorpha* (John & Miklos 1979), the tandem arrays suppress recombination, making it more distal. However, an alternative explanation to be investigated for this wheat chromosome is that sequences necessary to initiate recombination are in much lower concentration in the proximal region of the chromosomal arm. Whether or not the tandem arrays of repeats affect recombination in plants the idea serves as a good example to illustrate how tandem arrays of repeats might affect chromosome biology and so be maintained or deleted by natural selection.

Tandem arrays of repeats appear to be the molecular basis of heterochromatin (John & Miklos 1979; Flavell 1980). Heterochromatin is often late-replicating in the cell cycle, thus influencing the cell cycle duration via this specific stage. Bennett (1977) has provided data to show that in wheat × rye hybrids (*triticale*) endosperm development, which includes a series of extremely rapid nuclear divisions, can be severely affected by the heterochromatin on the ends of the rye chromosomes. In this example, the arrays of repeats interfere with grain development, not by being transcribed, but by affecting chromosome replication and division at a developmental stage when these are programmed to occur extremely rapidly.

Sequences are amplified into tandem arrays and fixed in species sufficiently often that most species can be distinguished qualitatively or quantitatively from closely related species by at least one major family of repeats. This is illustrated for members of the *Secale* genus in table 1. It is interesting to note that in *S. silvestre*, three of the four major *S. cereale* repeats are in low copy number or absent and this is correlated with a low amount of heterochromatin (Bedbrook *et al.* 1980*a*; Jones & Flavell 1982*b*). Thus in this species, not only have these particular repeats failed to be amplified and spread but no other arrays have accumulated to establish large blocks of terminal heterochromatin. This suggests that there is selection against the accumulation of major arrays of repeats in *S. silvestre* but not in the other species.

Tandem arrays frequently demonstrate variation in the number of repeats in an array between individuals (Jones & Flavell 1982*a*, *b*). This is probably the result of unequal crossing

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

TABLE 1. LEVELS OF SPECIFIC SEQUENCES IN MAJOR TANDEM ARRAYS IN *SECALE* SPECIES

repeat unit base pairs	*S. cereale*	*S. vavilovii*	*S. iranicum*	*S. montanum*	*S. africanum*	*S. silvestre*
480	6.1	2–5	2–5	1–5	—	—
610	2.7	—	—	0.5	0.04	—
120	2.4	1.5–3.0	1.5–3.0	1.5–3.0	2.4	2.4
630	0.6	0.01–0.04	0.01–0.04	0.16	—	—

Figures are percentage of total DNA. As described in Jones & Flavell (1982*b*), there are substantial errors in these values but they illustrate major quantitative differences between related species for the amounts of specific repeats. —, Below the level of detection and not necessarily that the sequence is absent from the species.

over or deletion due to intrastrand recombination, that is, due to mechanisms that are highly likely to be active with this kind of sequence arrangement.

TRANSPOSITION AND TRANSPOSABLE ELEMENTS

A vast number of closely related short sequences are found scattered throughout plant chromosomes. These 'dispersed' repeats can occupy more than 50% of the total DNA. Their presence implies that transposition has played a major role in chromosome evolution. The transposition of DNA from one chromosomal site to another was first postulated by McClintock (1951) from genetic studies on maize. She concluded that unstable alleles were due to the presence of an unstable element in a gene which was frequently excised to restore gene activity. In the last three years a number of such genetically defined transposable elements have been isolated by molecular cloning (see, for example, Federoff *et al.* 1983; Doring *et al.* 1984; Schwarz-Sommer *et al.* 1985; Bonas *et al.* 1984). The structure of these elements conforms to the generalized structure of transposable elements in other organisms in having inverted terminal repeats. This sequence structure would permit the formation of cruciform-type DNA secondary structures which probably facilitate the excision from and insertion of the element into the chromosomes. A model for this has recently been proposed which takes into account these and other features of plant transposable elements (Saedler & Nevers 1985).

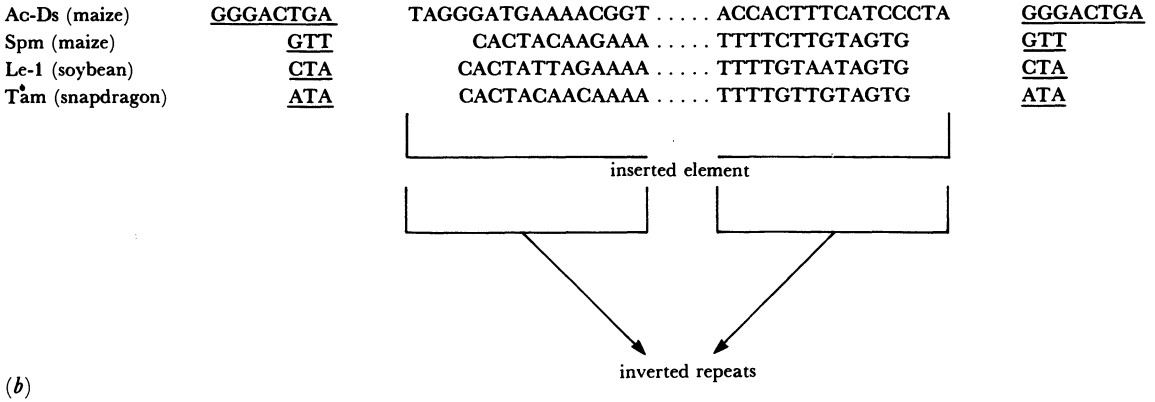
One of the remarkable findings is that there is considerable homology between the inverted repeats found on some (but not all!) transposable elements in different plant species and also mouse and human retroviruses, as well as sequences in *Drosophila*, yeast and plants. In figure 2*a*, the similarity between the inverted repeats of the Spm (maize) Le-1 (soybean) and Tam elements (antirrhinum) is illustrated and in figure 2*b* homologies between the animal, fungal and plant species are illustrated. The terminal six nucleotides are TGTTGG or a related sequence. T-rich and A-rich regions are also common to the six elements analysed. This homology implies either that excision and integration sequences of elements have been highly conserved over extraordinarily long time periods or that transposable elements have moved between kingdoms in more recent times. This latter idea is speculative but should be considered.

Upon insertion of an element a short chromosomal sequence at the site of insertion is duplicated and subsequently lies on either side of the element (see figures 2 and 3). The duplicated sequence differs of course with each insertion event. When an element is excised from the chromosomes the duplication of plant DNA remains, but the excision process often

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

(*a*)



(*b*)

Cin1 (maize)	TGTTGG----C---C-C------A--T----AA-A------T-A--AT-TTTTT---T-T---AA-AA
Ty1 (yeast)	TGTTGGAA-A-A-----A-----A--TA---A--A--A-T-A-A--A-T--T-TAT--------A---
IAP (mouse)	TGTTGG-A--C---C-CA-AT----------AA-AT-----T-A----T-T-TT-T---T---AA-AA
Copia (*Drosophila*)	TGTTGGAA-A-A-------A--TA---A---AA--T-A---A-------TTT-TAT-T------AA--
B104 (*Drosophila*)	TGTT---A-A----C-C--AT-TA--TA---A-----A-T-T-----T---TT-AT-T-T-----AAA
MMSV (mouse)	TG---GA----A--C--A--T-T----A---A---T-A-T-A----ATTTT---A----T---AAAAA

FIGURE 2. (*a*) The inverted repeats and duplicated sequences of some plant transposable elements. The sequences underlined represent the bases duplicated in a particular insertion event. Data taken from Sachs *et al.* (1983) (Ac-Ds); Schwarz-Sommer *et al.* (1985) (Spm); Vodkin *et al.* (1983) (Le-1) and Bonas *et al.* (1984) (Tam-1). (*b*) Terminal sequences of different elements capable of 'transposition'. The nucleotides included are those held in common between three or more of the elements; –, represents a non-conserved nucleotide. Cin-1 is a dispersed repeat from maize (Shepherd *et al.* 1984). Ty1 is a transposable element of *Saccharomyces cerevisiae* (Farabaugh & Fink 1980). IAP is the intracisternal A-particle gene of mouse which belongs to a group of endogenous proretroviral-like elements with long terminal repeats (Kuff *et al.* 1983). Copia is a transposable element of *Drosophila* comprising up to 5% of the *Drosophila* genome (Levis *et al.* 1980). B104 is a dispersed repeat of *Drosophila* (Scherer *et al.* 1982). MMSV is the Moloney Sarcoma Virus of mouse which has long terminal repeats (Dhar *et al.* 1980).

seems to remove one or a few bases of the duplicated sequences (see figure 3 and Sutton *et al.* 1984). These observations lead to the important evolutionary conclusions that the integration and excision of transposable elements into genes modifies gene structure, because of the duplication, and so is a source of novelty for evolutionary forces to test (Schwarz-Sommer *et al.* 1985). Searches are now underway among genes whose sequences are known for the duplications ('transposon footprints') which suggest an element has at one time resided at this position. It remains to be seen how significant transposable elements have been in the evolution of plant gene structure.

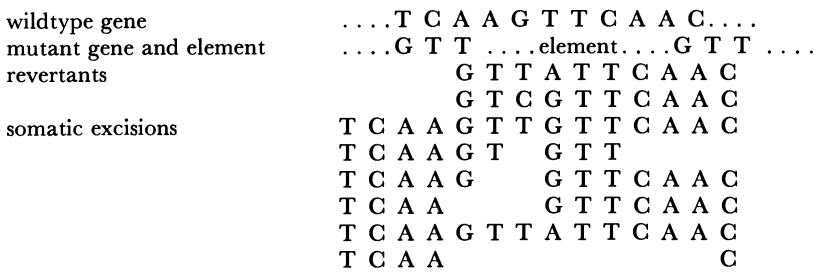


FIGURE 3. Examples of sequence modifications due to excision of Spm transposable element from the waxy gene in maize. The data are taken from Schwarz-Sommer *et al.* (1985). Note that upon insertion of the element into the waxy gene the sequence GTT was duplicated. The gaps in the sequences after excision of the element in somatic cells represent bases deleted.

How many copies of these identified transposable elements are in plant species? Federoff *et al.* (1982) estimated a maximum of around 6–10 copies of the element activator, 'Ac' in maize genomes tested, based upon segments of the element which are present in fewest copies per haploid genome. However, other segments of the element are present in many more copies (for example, greater than 40). This suggests that many of the dispersed repeats are defective transposable elements and were originally translocated via the same mechanisms as complete elements. It is well known from the original genetic studies on maize (McClintock 1951) as well as recent molecular sequence data (Federoff *et al.* 1982) that defective elements can be derived from active elements and mobilized by active elements. Are many dispersed repeats fossilized or defective transposable elements or have been transposed by being inserted into or between active elements? It appears likely (Flavell 1984). One dispersed repeat from maize has been sequenced and found to show characteristics of a transposable element and the long terminal repeat of mammalian retroviruses (Shepherd *et al.* 1984). Another in wheat (Flavell *et al.* 1981) being studied in this laboratory by N. Harris has a terminal sequence that is identical to the conserved inverted terminal repeat shown in figure 2*b*.

Transposable elements, as noted above, have inverted repeats at their ends. Where these are sufficiently long to form duplexes under stringent hybridization conditions, then the resulting hairpin structures can be observed in the electron microscope. They can also be isolated and their frequency in the genome estimated. Such studies in wheat (Bazetoux *et al.* 1978; R. B. Flavell, unpublished) have suggested that there are over 10^6 different chromosomal fragments containing inverted repeats lying close together and sufficiently long to form stable DNA duplexes *in vitro*. When all the fragments with shorter inverted repeats are added to these then perhaps up to 10% of nuclear DNA consists of sequences with this structure. They are not all likely to be related to sequences that have moved by the mechanisms used by transposable elements but the frequency of this kind of structure is consistent with the hypothesis

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

that many of the dispersed repeats have evolved from or been mobilized by transposable elements.

Transposable elements, by carrying genes that appear to code for enzymes that carry out the excision and integration functions, are units of DNA that can propagate themselves through populations if they are also occasionally duplicated. They are thus examples of what has been termed 'selfish DNA' (Doolittle & Sapienza 1980). Unless there is very strong selection pressure against them they will readily accumulate in populations (Ohta 1983*b*; Hickey 1982), although it is likely that there will be coselection for mechanisms, which could include methylation of regulatory sequences on the element, which suppress or control movement to acceptably low levels. The spread of transposable elements as a consequence of their DNA structure and properties provides a satisfying explanation for how many dispersed repeats have been fixed in plant populations. It is difficult to believe that the plethora of dispersed repeat families have all been spread by genetic drift or selection of individuals carrying them. Thus as discussed for tandem arrays, it is likely that the mechanisms that have moved them between chromosomes in individuals have also ensured their rapid spread in populations.

Transposition of DNA sequences has taken place not only within and between nuclear chromosomes but also between chloroplasts and mitochondrial chromosomes, mitochondrial and nuclear chromosomes and chloroplast and nuclear chromosomes. The transfer of genes from organelle genomes to nuclear chromosomes is consistent with the endosymbiont origin for organelles (Gray & Doolittle 1982) and the subsequent transfer of many of the endosymbiont's genes to the nucleus. Recent data have shown that many of the chloroplast DNA sequences are also present in the nucleus of plants (Timmis & Steele Scott 1983; Steele Scott & Timmis 1984) and perhaps more surprisingly that many chloroplast DNA sequences, but different sequences in different species, are to be found in mitochondrial genomes (Stern & Lonsdale 1982; Lonsdale *et al.* 1983). It is not known if any of these are 'functional' in the mitochondrion or are neutral passengers. However, how they became fixed in the mitochondrial population of each species, bearing in mind how many copies of each organelle genome exists in a plant cell, is an intriguing problem (Avise, this symposium; Birky 1983). It remains to be elucidated if the mechanisms of transfer involve transposable elements carrying genes that facilitate the transposition. Mitochondrial genomes are very variable in size between species, like nuclear genomes, and appear to tolerate the accumulation of considerable amounts of non-coding DNA (Ward *et al.* 1981). The interspecies variation between chloroplast genomes in size and sequence is very much less. This, of course, does not indicate that mutational processes are any less active in chloroplast DNA but that new variants are rarely fixed, presumably due to the need for conservation of genome form and sequence.

Repeated DNA turnover during evolution

In the foregoing the extent to which, during evolution, sequences become amplified and transposed and the resulting 'mutations' spread in populations has been emphasized. Within these repeated sequences many mutations, such as base changes, small deletions and insertions, also accumulate. Larger deletions also occur. It was stated earlier that deletions are commonly detected in tandem arrays, and they probably arise frequently because of the sequence arrangement. Deletion of dispensable but more complex chromosome segments may also occur by intrastrand homologous recombination where homologous dispersed repeats reside close to

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

one another. The specific deletion of families of dispersed repeats from a population or species is more difficult to understand unless strong selective forces against the presence of the particular sequence suddenly arise. However, inactive transposable elements could be lost by random drift, especially if the enzymes to excise the particular elements, but not reinsert them into chromosomes, were retained in a population. Excision of transposable elements is usually more frequent than reinsertion. 'Passive' loss of inactive elements could also occur by intrastrand recombination between the duplicated sequences at either end of the element.

Where mechanisms that lead to the fixation of additional DNA sequences are active, deletions may be selected because of the need to limit total DNA content and chromosome size. Selection against ever-increasing DNA contents would be expected because:

(i) Total DNA content is related to minimum cell cycle times and maximum developmental rates (Bennett 1972, 1973). If DNA contents increase, maximum development rates decrease. This can affect minimum generation time and the survival of species whose developmental rates and life cycle times are close to being limited by total DNA content.

(ii) Chromosome arm length may determine the position of the chromosome in the nucleus relative to other chromosomes and position may affect many features of the biology of the nucleus, as yet not understood (Bennett 1982; Heslop-Harrison & Bennett 1984). Therefore major changes in DNA content in localized positions may affect the position of that chromosome arm relative to others. In relation to this it is interesting to note that variation in total DNA content between closely related species is often found to be distributed rather uniformly between chromosomes and the relative chromosome sizes are preserved (Rees & Hazarika 1969).

The amplification and fixation of *new* sequences and deletion of pre-existing sequences leads to turnover of DNA sequences during evolution. Turnover models have been discussed extensively elsewhere (Flavell 1980; Thompson & Murray 1980). The turnover may be very rapid and driven by fixation mechanisms for new sequences and selection against increasing DNA content. However, these are not the only manifestations of turnover. Existing repeats are often 'replaced' during evolution with the same or a related variant by mechanisms that conserve the sequence at the same position, in contrast to mechanisms that lead to deletion first and replacement at the same site at some later date in the evolution of the population. Within tandem arrays recurrent unequal crossing over (Smith 1976) and also gene conversion type events can lead to replacement of the repeats, often by identical sequences but occasionally by a new variant (Dover 1982). These processes also lead to homogenization of an array, that is, maintenance, of the similarity of the repeats. This 'concerted evolution' of repeats is illustrated by many tandem arrays of repeats (Arnheim 1983) but particularly by the ribosomal RNA (rRNA) genes which have been studied in several plant species (Flavell 1983; Appels & Dvorak 1982; Yakura *et al.* 1983; Yakura & Tanifuji 1983).

The rRNA gene repeat unit in many plant species contains, in addition to the transcribed sequences, a series of tandemly arrayed repeats upstream from the promoter. The number of these repeats differs considerably between arrays within the species but rarely within an array (Appels & Dvorak 1982; Flavell 1983). The same is true for several point mutations which have been mapped within the spacer sequence (R. B. Flavell, unpublished results). The homology within an array illustrates that the fixation processes within an array are efficient compared with the mutations that create variation within an array.

Dispersed repeat families also appear to evolve 'in concert'. This has been concluded from several observations in plant genomes including the finding that for many of the repeats present

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

in wheat, barley and rye the sequences are much more closely related within a species than between species (Flavell *et al.* 1977). Gene conversion processes, if they were biased in favour of a particular variant, could account for the concerted evolution of some dispersed repeats (Flavell 1985; Dover & Tautz, this symposium). However, it is difficult to understand how this process could occur frequently enough between homologous repeats in a consistently biased way. Furthermore, frequent homologous recombination would be disastrous for plant chromosomes with so many repeats. Much of the greater similarity of dispersed repeats within species than between species is probably the result of amplification of new variants and deletion of old ones or the spreading of new variant transposable elements with deletion of old elements as discussed above. Illustrations of the reamplification of repeats, often in a new variant form or in combination with a sequence with which they had not been amplified previously have been published previously from this laboratory (Bedbrook *et al.* 1980*a*, *b*, 1981).

RNA transcripts may also play an important role in the concerted evolution of dispersed repeats via a gene conversion type process. Dispersed repeats of the same family will usually differ in their transcriptional activity due to mutation and their position in the genome. This variation in transcriptional activity would lead to the transcripts being predominantly of a small subset of the dispersed repeat family. If the RNA transcripts occasionally formed a DNA–RNA duplex in the chromosome and the DNA was converted to the sequence of the RNA strand, then the dispersed repeats would gradually be homogenized to the sequence of the predominant transcripts. A similar situation could result from a transcript being the source of new DNA copies via reverse transcription. The structure of the predominant transcripts could obviously vary over evolutionary time and between different populations or species.

During evolution then, amplifications, deletion, transposition, homogenization and replacement processes result in turnover of repeated sequences in a genome. These processes are also often involved, with natural selection and genetic drift, in the fixation of repeated DNA in populations. The fixation by the stochastic non-reciprocal mechanisms, as opposed to fixation by selection or drift has been called molecular drive by Dover (1982) and is also discussed in this volume (Dover & Tautz, this symposium). As populations diverge, then different sequences and sequence variants spread through each population. Each population is therefore part of a different turnover cycle. This produces major differences in the repeated DNA between even closely related species as illustrated by results on *Osmunda* (Stein *et al.* 1979), Cichorieae (Bachmann & Price 1977), *Vicia* (Straus 1972) and some cereal species (Rimpau *et al.* 1978, 1980; Flavell 1982).

The rate of turnover will depend on many factors, including genome size. In larger genomes, there is a greater probability of amplifications or deletions occurring because more DNA has to be replicated, etc. If more amplification events are tolerated in larger genomes because selection against small changes in genome size is less, then amplification rates will appear greater and the turnover rate will probably also be greater. Thompson & Murray (1980) and Preisler & Thompson (1981*a*, *b*) have provided some experimental evidence consistent with the hypothesis that the amplification rate is greater in species with larger genomes.

Another example comes from species in the genus *Lathyrus* (Narayan & Rees 1976, 1977). There is a threefold variation in nuclear DNA content between species in the genus but all species have the same number of chromosomes. This variation is mostly but not entirely due to repetitive DNA. The variation is also highly correlated with the amount of heterochromatin, a conclusion which is consistent with heterochromatin consisting predominantly of repetitive

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

DNA (Bedbrook *et al.* 1980*a*). The DNA that is amplified in the genus differs from species to species as illustrated by the percentage of repetitive DNA from one species that hybridizes to the DNA of another (table 2). For example only 14% of the repetitive DNA of *L. hirsutus* hybridized to the repetitive DNA of *L. clymenum*, which has a genome size 67% that of *L. hirsutus*, while 44% hybridized to the repetitive DNA of *L. articulatus*, which has a genome size 61% that of *L. hirsutus*.

TABLE 2. REPEATED SEQUENCE DNA HOMOLOGIES BETWEEN DIFFERENT *LATHYRUS* SPECIES

repeated DNAs hybridized together	*L. hirsutus* repeated DNA hybridized (%)	ΔT_m of hybrids
L. hirsutus × *L. hirsutus* (20.3)	100	0.0
L. hirsutus × *L. tingitanus* (17.9)	50	1.25
L. hirsutus × *L. odoratus* (17.2)	62	2.25
L. hirsutus × *L. sphaericus* (14.2)	17	4.0
L. hirsutus × *L. clymenum* (13.8)	14	4.5
L. hirsutus × *L. articulatus* (12.5)	44	3.0
L. hirsutus × *L. angulatus* (9.2)	21	3.5

Data taken from Narayan & Rees (1977).

The percentage hybridization values and stabilities (ΔT_m) of the DNA–DNA hybrids have been normalized to the values obtained for the self-hybridization of *L. hirsutus* DNA. The DNA contents of each species are given in parentheses.

When the sequences 'common' to each species are compared by examining the thermal stabilities of the DNA–DNA hybrids formed *in vitro*, in each case the interspecies hybrid DNAs are less stable than the intraspecies hybrids (table 2). This shows that in fact the 'common' repeated sequences must frequently be different in different species. As the proportion of repeated DNA that can form interspecies duplexes declines in the comparisons the extent of similarity (ΔT_m) between 'common' sequences also declines (table 1). Turnover of repeated sequences must be invoked to explain the extent of divergence.

A prediction of this view of repeated DNA evolution is that the longer that a given repeat family remains in a species then the more structural forms in which it is likely to exist due to reamplification, the more differences it will display from its counterparts in diverging species and the more chromosomal sites it will occupy due to the increased probability of becoming associated with a transposable element. One example fulfilling these predictions is provided by the 120 base pair repeat family of the *Secale* genus shown in table 1. This repeat family is present in all the *Secale* species and also in a wide range of *Aegilops* and *Triticum* species (unpublished results). This suggests the repeat family was established in progenitor species a very long time ago. The sequence has been found to be amplified in combination with several different sequences within the rye genome but these particular forms are not present (or at only low levels) in the *Aegilops* and *Triticum* species tested (see Bedbrook *et al.* 1980*a*) which have their own species-specific variants. The diversity of sequence types related to the 120 b.p. family in rye is illustrated by the low thermal stabilities of the DNA–DNA duplexes formed *in vitro* between such sequences (figure 4). This low thermal stability contrasts with the much higher stabilities of duplexes formed by the 480 or 610 b.p. repeats of rye (table 1) which have not been reamplified in many different forms and are probably much younger than the 120 b.p. repeat family in view of their presence only in the *Secale* genus and their absence from *Secale*

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

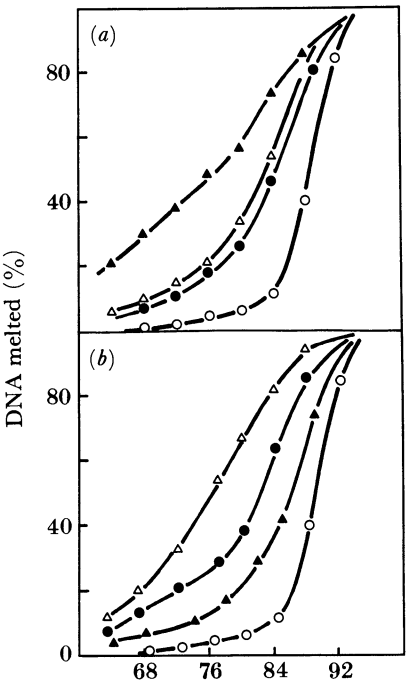


FIGURE 4. Thermal stability of heteroduplexes formed *in vitro* between repeats of different families in the rye genome. Sheared, denatured nick-translated cloned DNAs from the repeat families noted in figure 1 and table 1 were incubated with an over 5000-fold excess of denatured *S. cereale* DNA (fragment length 300–400 nucleotides) at 60 °C in 0.12 M phosphate buffer to a C_0t of 0.3 (C_0t is the concentration of DNA in moles of nucleotide per litre multiplied by incubation time in seconds). Duplex DNA was collected on hydroxyapatite and eluted in 0.12 M phosphate buffer by increasing the temperature in defined steps. (*a*) ●—●, Elution profile of renatured rye repeated DNA (measured by $O.D._{260}$); ○—○, elution profile of native rye DNA; △—△, elution profile of 480 b.p. repeat hybrids; ▲—▲, elution profile of 630 b.p. repeat hybrids. (*b*) ●—● and ○—○ as for (*a*); △—△, elution profile of 120 b.p. repeat hybrids; ▲—▲, elution profile of 610 b.p. repeat hybrids.

silvestre (table 1). The members of the 120 b.p. family are also dispersed over many more sites in the genome than the 610 and 480 b.p. repeats, as revealed by *in situ* hybridization under conditions where minor sites can be seen (Jones & Flavell 1982*a*, unpublished). Thus the time since a sequence was first amplified does seem to correlate with diversity of sequence types and locations. This correlation needs to be examined in many other families.

CONCLUDING REMARKS

The various turnover and fixation processes involving repetitive DNA operate sufficiently rapidly during evolution to create major structural differences between the chromosomes of separate populations and species. Examples involving major tandem arrays are given for the *Secale* genus in table 1. An example of a dispersed repeat is provided by studies on a sequence cloned from the wheat genome. It displays terminal sequence similarities to a transposable element (N. Harris, unpublished) and is present in many hundreds of copies in many *Triticum* and *Aegilops* species but is essentially absent from *Aegilops squarrosa* (Flavell *et al.* 1981; Flavell 1982). Thus these species must differ at hundreds of sites due to this one sequence alone. The biological consequences of such chromosome structural diversity are hard to define on present information. However, elsewhere (Flavell 1982) I have speculated on the relationship between reduced chromosome homology and the observed reduced meiotic chromosome pairing and

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

crossing over within the *Triticum* genus and its relatives (but see Rees *et al.* 1982). Earlier in this paper, I have drawn attention to the possible effects of (i) blocks of heterochromatin on chromosome behaviour; (ii) varying chromosome arm ratios on gene position in the nucleus; and (iii) varying total DNA content on developmental rates and cell size. It also seems inescapable that many of the changes in the secondary DNA of the chromosomes will affect chromosome behaviour, but documenting the evidence is a difficult process (Rees *et al.* 1982).

The turnover processes highlighted in this paper are not restricted to non-coding repeats. There is ample evidence for similar non-reciprocal processes playing important roles in the evolution of multigene families (Ohta 1983*a*), as reviewed elsewhere in this volume for the globin, immunoglobulin and histocompatibility gene families (Smithies & Powers, Hood *et al.*, Bodmer, all this symposium). The ribosomal RNA genes are a major multigene family in all eukaryotes and demonstrate very clearly concerted evolution (as described above) and also between-species differences due to turnover and homogenization processes (Arnheim 1983; Coen *et al.* 1982). The fixation of different repeats in the spacers of rRNA genes during species divergence is a particularly interesting example because these repeats appear to be involved in regulating gene expression by serving as binding sites for polymerase I transcription complexes (Reeder *et al.* 1983). How function might be conserved during the fixation of new variant sequences by turnover processes has been described (Dover & Flavell 1984) and is discussed elsewhere in this volume (Dover & Tautz, this symposium). One outcome of this fixation of different ribosomal RNA spacer repeats during species divergence is that sometimes rDNA loci, when moved into different species by interspecies hybridization, fail to function well because of their lack of compatibility with the polymerase I transcription factors of their new host species (Reeder *et al.* 1983).

The various processes affecting the structure of repeated genes and chromosomes discussed in this paper are clearly very important sources of mutations to be eliminated or spread by selection, drift or molecular drive. The roles of these kinds of mutations in the evolution of new plant phenotypes cannot be generalized: they need to be studied on a case-by-case basis. However, what has clearly been established is that these processes involving repetitive DNA are responsible for major differences in chromosomal phenotypes within and between species.

I am grateful to Nigel Harris for help with compiling the data on transposable elements.

References

Appels, R. & Dvorak, J. 1982 The wheat ribosomal DNA spacer region: its structure and variation in populations and among species. *Theor. Appl. Genet.* **63**, 337–348.

Arnheim, N. 1983 Concerted evolution of multigene families. In *Evolution of genes and proteins* (ed. M. Nei & R. K. Koehn), pp. 38–61. Sunderland, Massachusetts: Sinauer Associates.

Bachmann, K. & Price, H. J. 1977 Repetitive DNA in Cichorieae (Compositae). *Chromosoma (Berl.)* **61**, 267–275.

Bazetoux, S., Jouanin, L. & Huguet, T. 1978 Characterisation of inverted repeated sequences in wheat nuclear DNA. *Nucl. Acid Res.* **5**, 751–769.

Bedbrook, J. R., Jones, J. & Flavell, R. B. 1981 Evidence for the involvement of recombination and amplification events in the evolution of *Secale* chromosomes. *Cold Spring Harb. Symp. quant. Biol.* **45**, 755–760.

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

Bedbrook, J. R., O'Dell, M. & Flavell, R. B. 1980*b* Amplification of rearranged sequences in cereal plants. *Nature, Lond.* **288**, 133–137.

Bennett, M. D. 1972 Nuclear DNA content and minimum generation time in herbaceous plants. *Proc. R. Soc. Lond.* B **181**, 109–135.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Bennett, M. D. 1973 Nuclear characters in plants. *Brookhaven Symp. Biol.* **25**, 344–366.
Bennett, M. D. 1977 Heterochromatin, aberrant endosperm nuclei and grain shrivelling in wheat–rye genotypes. *Heredity* 39, 411–419.
Bennett, M. D. 1982 Nucleotypic basis of the spacial ordering of chromosomes in eukaryotes and the implications of the order for genome evolution and phenotypic variation. In *Genome evolution* (ed. G. A. Dover & R. B. Flavell). London: Academic Press.
Bennett, M. D. & Smith, J. B. 1976 Nuclear DNA amounts in Angiosperms. *Phil. Trans. R. Soc. Lond.* B **274**, 227–274.
Birky, C. W., Jr. 1983 Relaxed cellular controls and organelle heredity. *Science, Wash.* **222**, 468–475.
Bonas, U., Sommer, H. & Saedler, H. 1984 The 17-kb Taml element of *Antirrhinum majus* induces a 3-bp duplication upon integration into the chalcone synthase gene. *EMBO J.* **3**, 1015–1019.
Bostock, C. & Tyler Smith, C. 1982 Changes to genomic DNA in methotrexate-resistant cells. In *Genome evolution* (ed. G. A. Dover & R. B. Flavell), pp. 69–94. New York: Academic Press.
Coen, E. S., Strachen, J. & Dover, G. A. 1982 Dynamics of concerted evolution of ribosomal DNA and histone gene families in the melanogaster species subgroup of *Drosophila*. *J. molec. Biol.* **158**, 17–35.
Dennis, E. S., Gerlach, W. L. & Peacock, W. J. 1980 Identical polypyrimidine-polypurine satellite DNAs in wheat and barley. *Heredity* **44**, 349–366.
Deumling, B. & Greilhuber, J. 1982 Characterisation of heterochromatin in different species of the *Scilla siberica* group (*Liliaceae*) by *in situ* hybridisation of satellite DNAs and fluorochrome banding. *Chromosoma* **84**, 535–555.
Dhar, R., McClements, W. L., Enquist, L. W. & Vande Woude, G. F. 1980 Nucleotide sequence of integrated Moloney sarcoma provirus long terminal repeats and their host viral junctions. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2939–3941.
Doolittle, W. F. & Sapienza, C. 1980 Selfish genes, the phenotype paradigm and genome evolution. *Nature, Lond.* **284**, 601–603.
Doring, H. P., Tillmann, E. & Starlinger, P. 1984 DNA sequence of the maize transposable element *Dissociation*. *Nature, Lond.* **307**, 127–130.
Dover, G. A. 1982 Molecular drive: a cohesive mode of species evolution. *Nature, Lond.* **299**, 111–117.
Dover, G. A. & Flavell, R. B. 1984 Molecular coevolution: DNA divergence and the maintenance of function. *Cell* **38**, 622–623.
Farabaugh, P. J. & Fink, G. R. 1980 Insertion of eukaryotic transposable element Ty1 creates a 5-base pair duplication. *Nature, Lond.* **286**, 352–356.
Federoff, N., Wessler, S. & Shure, M. 1983 Isolation of the transposable maize controlling elements *Ac* and *Ds*. *Cell* **35**, 235–242.
Flavell, R. B. 1980 The molecular characterisation and organisation of plant chromosomal DNA sequences. *A. Rev. Pl. Physiol.* **31**, 569–596.
Flavell, R. B. 1982*a* Amplification, deletion and rearrangement: major sources of variation during species divergence. In *Genome evolution* (ed. G. A. Dover & R. B. Flavell), pp. 301–324. London: Academic Press.
Flavell, R. B. 1982*b* Chromosomal DNA sequences and their organisation. In *Nucleic acids and proteins in plants: II. Encycl. Pl. Physiol.* N.S. **14** B, 46–74.
Flavell, R. B. 1983 Repeated sequences and genome architecture. In *Structure and function of plant genomes* (ed. O. Ciferri & L. Dure), pp. 1–14. New York: Plenum Press.
Flavell, R. B. 1984 DNA transposition – a major contributor to plant chromosome structure. *BioEssays* **1**, 21–22.
Flavell, R. B. 1985 Repeated sequences and genome change. In *Advances in plant gene research*, vol. 2, (ed. L. Dennis & B. Hohn). Vienna: Springer Verlag. (In the press.)
Flavell, R. B., O'Dell, M. & Hutchinson, J. 1981 Nucleotide sequence organisation in plant chromosomes and evidence for sequence translocation during evolution. *Cold Spring Harbor Symp. Quant. Biol.* **45**, 501–508.
Flavell, R. B., Rimpau, J. & Smith, D. B. 1977 Repeated sequence DNA relationships in four cereal genomes. *Chromosoma* **63**, 205–222.
Gerlach, W. L. & Peacock, W. J. 1980 Chromosomal locations of highly repeated DNA sequences in wheat. *Heredity* **44**, 269–276.
Gray, M. W. & Doolittle, W. F. 1982 Has the endosymbiont hypothesis been proven? *Microbiol. Rev.* **46**, 1–42.
Greilhuber, J. & Loidl, J. 1983 On regularities of C banding patterns, and their possible causes. *Third Kew Chromosome Conference*, p. 344. London: George Allen & Unwin.
Heslop-Harrison, J. S. & Bennett, M. D. 1984 Chromosome order – possible implications for development. *J. Embryol. exp. Morph.* **83**, 51–73.
Hickey, D. A. 1982 Selfish DNA: a sexually-transmitted nuclear parasite. *Genetics* **101**, 519–531.
Hinegardner, R. 1976 Evolution of genome size. In *Molecular evolution* (ed. F. J. Ayala), pp. 179–199. Sunderland, Massachusetts: Sinauer Associates.
Hourcade, D., Dressler, D. & Wolfson, J. 1973 The amplification of ribosomal RNA genes involving a rolling circle intermediate. *Proc. natn. Acad. Sci. U.S.A.* **70**, 2926–2930.
John, B. & Miklos, G. L. G. 1979 Functional aspects of satellite DNA and heterochromatin. *Int. Rev. Cytol.* **58**, 1–114.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Jones, J. D. G. & Flavell, R. B. 1982*a* The mapping of highly repeated DNA families and their relationship to C bands in chromosomes of *Secale cereale*. *Chromosoma* (*Berl.*) **86**, 595–612.

Jones, J. D. G. & Flavell, R. B. 1982*b* The structure amount and chromosomal localisation of defined repeated DNA sequences in species of the genus *Secale*. *Chromosoma* (*Berl.*) **86**, 613–641.

Kuff, E. L., Fenestra, A., Lueders, K., Smith, L., Hawley, R., Hozumi, N. & Shulman, M. 1983 The intracisternal A-particle gene belongs to a group of endogenous proretroviral-like elements present in about 1000 copies per haploid genome of mouse. *Proc. natn. Acad. Sci. U.S.A.* **80**, 1992–1996.

Levis, R., Dunsmuir, P. & Rubin, G. M. 1980 Terminal repeats of the *Drosophila* transposable element copia: nucleotide sequence and genomic organisation. *Cell* **21**, 581–588.

Lonsdale, D. M., Hodge, T. P., Howe, C. J. & Stern, D. B. 1983 Maize mitochondrial DNA contains a sequence homologous to the ribulose 1,5 bisphosphate carboxylase large subunit gene of chloroplast DNA. *Cell* **34**, 1007–1014.

McClintock, B. 1951 Chromosome organisation and gene expression. *Cold Spring Harbor Symp. Quant. Biol.* **16**, 13–47.

Narayan, R. K. J. & Rees, H. 1977 Nuclear DNA divergence among *Lathyrus* species. *Chromosoma* (*Berl.*) **63**, 101–107.

Narayan, R. K. J. & Rees, H. 1976 Nuclear DNA variation in *Lathyrus*. *Chromosoma* (*Berl.*) **54**, 141–154.

Ohta, T. 1983*a* On the evolution of multigene families. *Theor. Pop. Biol.* **23**, 216–240.

Ohta, T. 1983*b* Theoretical study on the accumulation of selfish DNA. *Genet. Res.* **41**, 1–15.

Ohta, T. & Dover, G. A. 1983 Population genetics of multigene families that are dispersed in two or more chromosomes. *Proc. natn. Acad. Sci. U.S.A.* **80**, 4079–4083.

Ohta, T. & Dover, G. A. 1984 The cohesive population genetics of molecular drive. *Genetics* **108**, 501–521.

Preisler, R. S. & Thompson, W. F. 1981*a* Evolutionary sequence divergence within repeated DNA families of higher plant genomes. I. Analysis of reassociation kinetics. *J. molec. Evol.* **17**, 78–84.

Preisler, R. S. & Thompson, W. F. 1981*b* Evolutionary sequence divergence within repeated DNA families of higher plant genomes. II. Analysis of thermal denaturation. *J. molec. Evol.* **17**, 85–93.

Reeder, R. H., Roan, J. G. & Dunaway, M. 1983 Spacer regulation of *Xenopus* ribosomal gene transcription: competition in oocytes. *Cell* **35**, 449–456.

Rees, H. & Hazarika, M. H. 1969 Chromosome evolution in *Lathyrus*. *Chromosomes Today* **2**, 157–165.

Rees, H., Jenkins, G., Seal, A. & Hutchinson, J. 1982 Assays of the phenotypic effects of changes in DNA amounts. In *Genome evolution* (ed. G. A. Dover & R. B. Flavell). London: Academic Press.

Rimpau, J., Smith, D. B. & Flavell, R. B. 1978 Sequence organisation analysis of the wheat and rye genomes by interspecies DNA/DNA hybridisation. *J. molec. Biol.* **123**, 327–359.

Rimpau, J., Smith, D. B. & Flavell, R. B. 1980 Sequence organisation in barley and oats chromosomes revealed by interspecies DNA/DNA hybridisation. *Heredity* **44**, 131–149.

Sachs, M. M., Peacock, W. J., Dennis, E. S. & Gerlach, W. L. 1983 Maize *Ac/Ds* controlling elements – a molecular viewpoint. *Maydica* **28**, 289–302.

Saedler, H. & Nevers, P. 1985 Transposition in plants: a molecular model. *EMBO J.* (In the press.)

Scherer, G., Tschudi, C., Perera, J. & Delius, H. 1982 B104, a new dispersed repeated gene family in *Drosophila melanogaster* and its analogies with retroviruses. *J. molec. Biol.* **157**, 435–451.

Schimke, R. 1982 In *Gene amplification* (ed. R. Schimke), pp. 317–333. New York: Cold Spring Harbor Press.

Schwarz-Sommer, Z., Giere, A., Klosgen, R., Wienand, U., Peterson, P. A. & Saedler, H. 1985 The Spm (En) transposable element controls the excision of a 2Kb DNA insert at the wx^{m-8} locus of *Zea mays*. *EMBO J.* **4**, 591–597.

Shepherd, N. S., Swarz-Sommer, Z., vel Spalve, J. B., Gupta, M., Wienand, V. & Saedler, H. 1984 Similarity of the Cin 1 repetitive family of *Zea mays* to eukaryotic transposable elements. *Nature, Lond.* **307**, 185–187.

Smith, G. P. 1976 Evolution of repeated DNA sequences by unequal crossover. *Science, Wash.* **191**, 528–535.

Snape, J. W., Flavell, R. B., O'Dell, M., Hughes, W. G. & Payne, P. I. 1985 Intrachromosomal mapping of the nucleolar organiser region relative to three marker loci on chromosome 1B of wheat (*Triticum aestivum*). *Theoret. Appl. Genet.* **69**, 263–270.

Steele Scott, N. & Timmis, J. N. 1984 Homologies between nuclear and plastid DNA in spinach. *Theoret. Appl. Genet.* **67**, 279–288.

Stein, D. B., Thompson, W. F. & Belford, H. S. 1979 Studies on DNA sequences in the Osmundaceae. *J. molec. Evol.* **13**, 215–232.

Stern, D. B. & Lonsdale, D. M. 1982 Mitochondrial and chloroplast genomes of maize have a 12Kb DNA sequence in common. *Nature, Lond.* **299**, 698–702.

Straus, N. 1972 Reassociation of bean DNA. *Carnegie Inst. Wash. Yb.* **71**, 257–259.

Sutton, W. D., Gerlach, W. L., Schwartz, D. & Peacock, W. J. 1984 Molecular analysis of *Ds* controlling element mutations at the *Adh-1* locus of maize. *Science, Wash.* **223**, 1265–1268.

Tautz, D. & Renz, M. 1984 Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucl. Acid Res.* **12**, 4127–4138.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Thompson, W. F. & Murray, M. G. 1980 Sequence organisation in pea and mung bean DNA and a model for genome evolution. In *Fourth John Innes Symposium* (ed. J. R. Davies & D. A. Hopwood), pp. 31–45. Norwich: John Innes Institute.

Thompson, W. F. & Murray, M. G. 1981 The nuclear genome: structure and function. In *Biochemistry of plants* (ed. P. K. Stumpf & E. E. Conn), pp. 1–81. New York: Academic Press.

Timmis, J. N. & Steele Scott, N. 1983 Sequence homology between spinach nuclear and chloroplast genomes. *Nature, Lond.* **305**, 65–67.

Vodkin, L. O., Rhodes, P. R. & Goldberg, R. B. 1983 A lectin gene insertion has the structural features of a transposable element. *Cell* **34**, 1023–1031.

Ward, B. L., Anderson, R. S. & Bendich, A. J. 1981 The mitochondrial genome is large and variable in a family of plants (*Cucurbitaceae*). *Cell* **25**, 793–803.

Yakura, K., Kato, A. & Tanifuji, S. 1983 Structural organisation of ribosomal DNA in four *Trillium* species and *Paris verticillota*. *Pl. Cell Physiol.* **24**, 1231–1240.

Yakura, K. & Tanifujik, S. 1983 Molecular cloning and restriction analysis of EcoRI fragments of *Vicia faba* rDNA. *Pl. Cell Physiol.* **24**, 1327–1330.

Yamamoto, M. & Miklos, G. L. G. 1978 Genetic studies on heterochromatin in *Drosophila melanogaster* and their implications for the functions of satellite DNA. *Chromosoma* (*Berl.*) **66**, 71–78.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES