Acta Bot. Neerl. 39(4), December 1992, p. 369–384

REVIEW

Nuclear DNA markers in angiosperm taxonomy

K. BACHMANN

Hugo de Vries Laboratory, University of Amsterdam, Kruislaan 318, NL-1098 SM Amsterdam, The Netherlands

CONTENTS

Key-words: angiosperms, molecular phylogeny, RAPD, reticulate evolution, RFLP, VNTR.

CHLOROPLAST VS. NUCLEAR DNA

New methods for the analysis of polymorphisms in DNA sequences have a profound influence on systematic botany. Several recent books deal with molecular methods in evolution and taxonomy (Clegg & O'Brien 1990; Doolittle 1990; Hillis & Moritz 1990; Li & Graur 1991), some of them specifically with plants (Crawford 1990; Soltis *et al.* 1992). A general revision of the evolutionary relationships of plants with molecular methods seems to be well underway (Palmer *et al.* 1988). Chloroplast DNA (cpDNA) has turned out to be the most convenient common molecule for the reconstruction of plant phylogeny. In this review, the use of markers in the nuclear genome will be examined. Obtaining reliable phylogenetic information from nuclear DNA tends to be considerably less straightforward than working wth cpDNA. This pertains both to the collection and to the interpretation of the data. However, the features of cpDNA that make it such a useful molecule for phylogenetic analysis also limit its applicability. A brief and simplified comparison of these features with those of nuclear DNA can help point out where and why nuclear markers become important.

Chloroplasts contain closed circular DNA molecules about 150 000 base pairs (150 kb) in length. The smallest nuclear genomes in angiosperms are about 500-fold bigger, and the variation in haploid nuclear genome sizes ('C values') among the angiosperms is considerable from *Arabidopsis thaliana* (Brassicaceae) with about 10^5 kb in the five chromosomes of the haploid genome to *Fritillaria assyriaca* (Liliaceae) with $1{\cdot}23 \times 10^8$ kb in 12 chromosomes (Bennett *et al.* 1982; Price 1988). There can be considerable differences in the genome sizes of closely related congeneric species and there is an increasing number of reports on intraspecific variation in genome size (Price 1991). A variation in C values from 1·84 pg to 7·14 pg ($1\text{ pg} = 9{\cdot}5 \times 10^5$ kb) among populations of *Collinsia verna*

(Scrophulariaceae) seems to be the widest intraspecific range on record (Greenlee *et al.* 1984). The range of about 2:1 between the largest and the smallest cpDNA in land plants (Palmer 1985) is small in comparison and easy to explain. The availability of the complete sequences of several cpDNAs facilitates considerably the evolutionary comparison (Ohyama *et al.* 1986; Shinozaki *et al.* 1986).

All chloroplast genomes contain more or less the same set of genes. Even the order of the genes on the chloroplast map changes relatively slowly. A segment containing, among others, the genes for the ribosomal RNAs typically is present twice in the cpDNA in the form of an inverted repeat. This cp-rDNA should not be confused with the nuclear rDNA, which I shall discuss below, or with mitochondrial rDNA, which up to now has played only a minor role in plant evolutionary studies (Grabau 1985). The two repeat copies in the cpDNA separate a large single-copy region from a small single-copy region. The problem of multiple copies of identical or similar sequences therefore plays virtually no role in cpDNA evolution while it is ever present in comparisons of nuclear genomes (Doyle 1991, 1992). Differences among cpDNAs are mostly base substitutions at strictly homologous (orthologous) sites. Major changes in genome architecture such as insertions, deletions, inversions and transpositions involving several genes, are relatively rare. When they occur, they can be crucial markers for major monophyletic groups (Downie & Palmer 1992).

Typically, the inheritance of chloroplasts is uniparental. The evolution of cpDNA is essentially that of asexual haploid clones. Each mutation initiates a monophyletic line, and cladistic analysis is strictly applicable at all taxonomic levels. A cladogram based on cpDNA should parallel that of the plants containing the chloroplasts. Usually it does (e.g. Crowhurst *et al.* 1990). However, recombination in the nuclear genomes within species and reticulate evolution (hybridization) between species are essential aspects of organismal evolution, and discrepancies between the phylogenies of chloroplast and nuclear DNAs can reveal crucial evolutionary events. Moreover, chloroplast DNA is essentially a neutral independent marker of phylogeny, while the vast majority of evolutionary change relevant to the living plant is encoded in the nuclear DNA. If we want to go beyond phylogenetic reconstruction and analyse the interactions between organismic and molecular evolution, we have to study the nuclear genome.

One of the advantages of cpDNA for phylogenetic analysis is the relatively slow sequence evolution. The taxonomic level at which cpDNA sequences are optimally informative is the differentiation of genera. With an increasing amount of noise (homoplasy), cpDNA can be compared at higher taxonomic levels (Palmer 1985). At lower taxonomic levels, especially for studies on population genetics and speciation, cpDNA usually provides little informative variation. Here, of course, recombination and reticulate evolution play an essential role, and nuclear DNA markers reveal their full value (Schaal *et al.* 1991). However, there is informative sequence variation for all levels of phylogenetic analysis in the nuclear genome, and we may state that, aside from fossils documenting extinct lineages, an event in phylogenetic history that has not left a recognizable trace in the nuclear DNA is not documented anywhere.

APPROACHES TO THE NUCLEAR GENOME

There can be considerable variation in genome sizes among closely related species, and this shows that the evolution of nuclear genomes is less constrained and faster than that of the organisms harbouring these genomes. The selective amplification of small parts of the

genome can create large fractions of repetitive DNA that may be specific for young evolutionary lineages (McIntyre *et al.* 1988). A significant fraction of the DNA found in one species may be absent in a closely related one (Ganal & Hemleben 1986b; Schmidt *et al.* 1991; Unfried *et al.* 1991). Such DNA sequences cannot play a role in coding for the phenotype, even though their bulk may well have a phenotypic effect via cell-cycle time and cell size (Bennett 1985; Price 1991). The coding sequences shared by all plants are embedded in a vast and variable majority of non-coding DNA, much of it consisting of very many repeats of a few basic sequences. Efforts to derive one overall measure of a genetic similarity among organisms from the common coding sequences in nuclear DNA become technically difficult when genomes with large and variable amounts of DNA are compared. The most general method, DNA–DNA hybridization (Bendich & Bolton 1967), is now only of historical interest in angiosperm taxonomy.

The nearest we can come to an overall measure of genetic relatedness is a random sampling of variation at many strictly homologous sites all across the genome as it is attempted in iso-enzyme studies (Gottlieb 1977). A typical problem with iso-enzyme studies is the limited number of loci that are available for analysis. A more extensive random sampling of genetic variation all across the genome can be achieved with DNA fingerprinting, RFLP and RAPD methods. I shall discuss these methods and their limitations below.

Alternatively, of course, there are many genes common to all plant genomes, and the nucleotide sequence of each one can be used for phylogenetic reconstruction. This yields the phylogeny of that one gene, and (as in the case of cpDNA phylogenies) it will be necessary to check if this reflects the phylogeny of the organisms (Pamilo & Nei 1988; Doyle 1992). The key concept here is that of concordance (Avise & Ball 1990) of gene genealogies (Hudson 1990). A general consensus seems to emerge that the best possible molecular approach to angiosperm phylogeny is the precise sequence comparison of several genetically unlinked and functionally diverse genes, where the chloroplast genome may be considered one linkage group (Martin & Dowd 1991). There is, however, no consensus yet about the nuclear genes that should be included in a general taxonomic survey. The effort involved in characterizing one plant this way excludes this approach for the time being from population sampling. Implicitly this creates a typological approach (i.e., the individual sequence from one specimen represents a species, genus or even family). This seems to be an acceptable practice for the phylogeny of higher categories, where intraspecific sequence variation has to be treated as noise unless it can be recorded statistically as allele frequencies. The continuous improvement in the methods for isolating strictly homologous genes from genomes and for determining their sequences will soon make phylogenetic analysis via gene trees generally available.

Other methods will have to be used at the species level. Here, a representative sampling of individuals and of variation throughout the genome requires simpler methods for determining polymorphisms. The availability of all kinds of direct and indirect methods for the comparison of DNA sequence variation makes it possible to choose the proper combination of techniques for each project and each organism. In the following sections, I shall try to illustrate some typical approaches.

RESTRICTION FRAGMENT LENGTH POLYMORPHISMS (RFLPs)

Many bacterial restriction endonucleases are now commercially available. Each recognizes a short nucleotide sequence, its specific restriction site, at which it cuts double-stranded

DNA. Restriction sites 4 or 6 bp long are typical. Such sites occur statistically in any longer stretch of DNA unless it consists of repetitive simple sequences. A restriction enzyme therefore will digest DNA so that homologous sequences are cut into identical restriction fragments. A nucleotide substitution in a restriction site will prevent recognition and cutting, and will replace two restriction fragments with one of their combined length. It will create a restriction fragment length polymorphism (RFLP). Restriction fragments can be sorted according to length by gel electrophoresis. All fragments of one length then form a band in the gel. Banding patterns on gels based on the restriction fragments from a piece of DNA are routinely used for identification and comparison of DNA sequences. If a piece of DNA is digested with several different restriction enzymes singly and together, the relative positions of the various cutting sites can be determined and their distances in base pairs can be shown on a physical map of the DNA (Fig. 4 of Lanaud *et al.* 1992). Such a physical map is an extract of the actual nucleotide sequence and can be predicted if the sequence is known.

Restriction fragments of identical length need not be homologous. The more bands there are in a pattern, the less easy will it be to determine their homology. This is extreme in the digest of an entire nuclear genome containing thousands of fragments which form a smear all along the gel. In this case, the DNA fragments can be blotted from the gel onto a membrane filter (Southern 1975) and individual bands homologous to a (radioactively) labelled cloned DNA-probe can be detected selectively by DNA/DNA hybridization. The probe may be cloned from the species itself (a 'homologous' probe) or from another species with a sufficiently similar sequence so that the probe finds its homologue by base pairing (a 'heterologous' probe).

For the chloroplast genomes of several species from diverse angiosperm families there exist complete sets of probes covering the entire molecule (e.g. the dicot, lettuce: Jansen & Palmer 1987; and the monocot, *Oncidium excavatum*: Chase & Palmer 1989). The conservative evolution of cpDNA facilitates the cross-hybridization with heterologous probes, and phylogenetic comparisons of cpDNA can be based on complete physical maps. Much of the variation among the maps from related species is due to the gain or loss of a restriction site due to a point mutation. Such mutations can be arranged as character-state changes in a strict cladistic analysis. In this way, RFLPs are a statistically characterized random sample (based on the random occurrence of restriction sites) of sequence variation all across the cpDNA. RFLPs can also arise by the insertion, deletion, or inversion of DNA. Such events can be recognized and differentiated from restriction site mutations in cpDNA analyses. RFLP analysis of nuclear DNA is never as complete and usually not as precise. It involves either the detailed comparison of a gene (essentially a simplified sequence comparison) or the search for RFLPs at a random set of loci across the genome. The first is exemplified by RFLP analysis of the ribosomal RNA genes. The latter depends on the availability of a set of appropriate cloned probes for various regions of the genome.

RIBOSOMAL RNA GENES AS AN EXAMPLE

The nuclear genes coding for ribosomal RNAs (rDNA) occur hundreds of times as tandem (one after the other) repeats at the nucleolar organizer regions on one or more chromosomes of a haploid set. This and the abundance of their RNA products in the cell has facilitated their isolation and they have been used for a variety of molecular studies including the earliest attempts at molecular phylogeny of angiosperms (Vodkin &

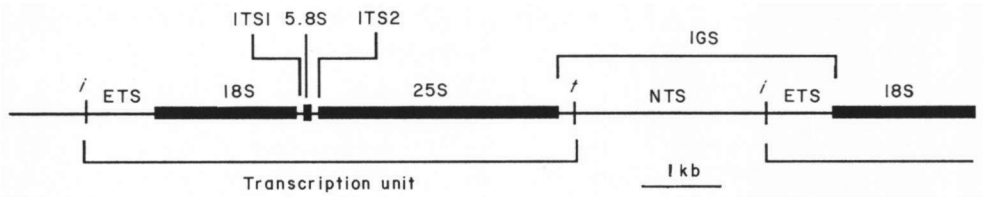

Fig. 1. Structure and terminology of the nuclear ribosomal genes (rDNA). Each of the repeat units (*i* to *i*) consists of a transcription unit between a transcription initiation site (TIS, *i*) and a transcription termination site (TTS; *t*) and a non-transcribed spacer (NTS). Black bars mark the regions coding for ribosomal RNAs, 18S, 5.8S, and 25S. These are preceded by an external transcribed spacer (ETS) and separated by the two internal transcribed spacers, ITS1 and ITS2. The entire region between the end of one 25S coding sequence and the start of an 18S coding sequence is the intergenic spacer (IGS).

Katterman 1971). Each rDNA gene is transcribed into one continuous primary transcript (precursor rRNA), from which the ribosomal 18S, 5·8S, and 25S RNAs are cut, whereby sequences corresponding to an 'external transcribed spacer' (ETS) and two short 'internal transcribed spacers' (ITS1 and ITS2) are discarded (Fig. 1). The ribosomal 5S RNAs are coded by a separate, unlinked tandem array (Gerlach & Dyer 1980; Hemleben & Werts 1988; Dvorak *et al.* 1989; Lapitan *et al.* 1991). The rDNA transcription units are separated by 'non-transcribed spacers' (NTSs, intergenic spacers: IGSs). In general, all the numerous copies of the 18S and 25S rRNA genes in a genome have the same sequence, and the sequences of these two RNAs evolve slowly (Zimmer *et al.* 1989; Martin & Dowd 1991; Hamby & Zimmer 1992). This conservation of the coding sequences contrasts with the nearly unrestrained evolution of the non-transcribed spacer region (Jorgensen & Cluster 1988). DNA probes from the IGS of *Cucurbita pepo* (courgette) do not even hybridize with DNA from other genera of the Cucurbitaceae such as *Cucumis* (cucumber and melon; Torres *et al.* 1989; Torres & Hemleben, 1991). Each intergenic spacer typically contains tandem repeats of simple sequences, 80–325 bp in length in different species. These simple-sequence sub-repeats are usually conserved within a genus or among closely related genera (Ganal & Hemleben 1986a). Length heterogeneity among the rDNA repeat units within a species can arise when the number of simple-sequence repeats in the IGS region varies (Yakura *et al.* 1984; Rogers & Bendich 1987; Hemleben *et al.* 1988; Reddy *et al.* 1990). Enzymes that do not cut within the simple-sequence repeats of the IGS transform spacer length heterogeneity into RFLPs (e.g. Bellarosa *et al.* 1990; Reddy *et al.* 1990). This, rather than mutations in restriction sites, is the typical origin of RFLPs in the rDNA. As (heterologous) probes for the rDNA coding sequences have been available for some time, there are quite a few taxonomic and evolutionary studies using RFLP analysis of the rDNA.

The value of any polymorphism for phylogenetic analysis depends on our knowledge of the way in which this polymorphism evolves. Even where we have a general understanding for the evolution of a type of polymorphism (such as for base pair replacements), we have to check the specific parameters empirically in each case. There are several possible mechanisms leading to changes in the repeat number of tandemly repetitive DNA, but there are also mechanisms that eliminate variation among the repeats in one linked cluster (Appels & Dvorak 1982; Dover 1986; Sytsma & Schaal 1990). In general, recombination tends to promote sequence homogeneity within populations, and the magnitude of fragment length differences among species need not be a measure of the phylogenetic distance.

RFLPs based on simple-sequence repeats in spacers are being used for the determination of (phylo)genetic relatedness, mostly among populations of one species and among congeneric species. There are considerable differences among the levels of variability in different cases. Twenty length variants were detected among the rDNA repeats of a single plant of *Vicia faba* (Yakura *et al.* 1984) and up to five per individual in *Clematis fremontii* (Learn & Schaal 1987). No intraspecific spacer length variation, but some variation in restriction sites was observed in an extensive survey of *Rudbeckia missouriensis* (King & Schaal 1989). In many species restriction site variation occurs only among races (Doyle *et al.* 1984) or at the interspecific level (Sytsma & Schaal, 1985; Doyle & Beachy 1985; Rieseberg *et al.* 1988). Qualitative and quantitative differentiation among populations and subspecies was found in wild barley (Saghai-Maroof *et al.* 1984), *Phlox divaricata* (Schaal *et al.* 1987) and *Clematis fremontii* (Learn & Schaal 1987). Sytsma & Schaal (1985) have analysed the *Lisianthius skinneri* complex in Panama. *Lisianthius* is one of the apparently rare examples of length variation in the ITS rather than the intergenic spacer (Sytsma & Schaal 1990). Bellarosa *et al.* (1990) have used RFLPs in the rDNA to determine the relationship among mediterranean oaks (*Quercus* spp.), D'Ovidio *et al.* (1990a) that of poplars (*Populus* spp.), Reddy *et al.* (1990) have compared 23 populations from 12 species of *Secale*, and Zentgraf *et al.* (1992) have used diagnostic RFLPs in the rDNA cluster to distinguish the Asian and African species groups within *Cucumis*.

There is no uniform method for treating the data in the various studies using RFLPs in the nuclear rDNA. While Reddy *et al.* (1990) employ an essentially phenetic approach based on frequencies of spacer-length variants, the classification by Bellarosa *et al.* (1990) is based on the shared presence of length variants and is therefore essentially cladistic. The recent analysis of nuclear rRNA variation in *Krigia* (Asteraceae) by Kim & Mabry (1991) is an instructive example for the evaluation of rDNA RFLP data.

RFLP ANALYSIS WITH RANDOM GENOMIC AND cDNA PROBES

RFLPs all across the genome can be used to determine the relatedness and the phylogeny of genomes. The crucial factor is the availability of cloned DNA-probes for the detection of homologous sequences by DNA/DNA hybridization on blots of restriction digests. RFLPs in the nuclear DNA can be used as co-dominant markers in genetic crosses, and their segregation and linkage relationships can be determined by Mendelian analysis. They can be used to produce an RFLP map of the genome. In contrast to the physical map of a segment of a genome in which restriction sites are mapped in base-pair distances, distances on an RFLP map of the genome are in recombination units. The genetic loci detected by RFLPs can be incorporated in the general genetic map of the species and greatly facilitate Mendelian analysis as they show a simple monogenic co-dominant inheritance pattern (Helentjaris *et al.* 1985; Landry *et al.* 1987; Paterson *et al.* 1988). Nuclear RFLPs form a direct link between the genetic analysis of the genome and phylogenetic analysis. Ideally their map positions are known, and their contribution to a truly representative sampling of the genome can be determined (Kesseli *et al.* 1991). The sequence responsible for an RFLP and the relevant chromosomal region are experimentally accessible via the cloned probe, and a direct link can be laid between molecular evolution and the genetic basis of morphological evolution.

Unfortunately, the expenses in time and money for the full use of nuclear RFLP analysis can become limiting. This primarily concerns the production of a set of informative

probes, the preparation and the characterization of a clone 'library'. These probes are either cDNA, i.e. DNA that has been synthesized *in vitro* by reverse transcription from mRNA (Helentjaris *et al.* 1985; Bernatzky & Tanksley 1986), or pieces of DNA cut directly from the nuclear genome (Kochert *et al.* 1991; Neuhausen 1992). cDNA clones contain only the coding sequences of expressed genes, while random genomic clones also contain regulatory and non-coding repetitive sequences, which require a more intensive screening for useful probes (Xu & Sleper 1991) but provide access to a wider range of variation.

Due to the considerable technical investment, nuclear RFLP analysis has been applied mainly to crop species. These studies, however, frequently involve interesting phylogenetic problems, mostly on the level of species within a genus. The levels of intraspecific or interspecific variation for nuclear RFLPs are rather unpredictable in plants from various genera or families (Shattuck-Eidens *et al.* 1990). Neuhausen (1992) found that RFLPs within species of melons (*Cucumis*) mostly were due to mutations in restriction sites, while RFLPs between species frequently involved chromosome restructuring. Song *et al.* (1988a,b, 1990) have constructed cladograms of *Brassica* species and cultivars from nuclear RFLPs, which agree well with other data and provide new insights in the evolution of this group. Their investigation raises the difficult question of the incorporation of polyploid hybrid taxa among diploids in a cladistic analysis. Brummer *et al.* (1991) constructed separate cladograms for diploid and tetraploid cultivars of alfalfa (*Medicago*), but also faced the problem of reticulate evolution via hybridization at one ploidy level ('homoploid reticulate evolution'). Kochert *et al.* (1991) compared 8 cultivars and 14 species of peanuts (*Arachis*) and found that they could recognize specific restriction fragments of the diploid parental species in the banding pattern of the tetraploids. Two cottonwood species (*Populus fremontii* and *P. angustifolia*) along the Weber River drainage in Utah could be identified by interspecific RFLPs and the parentage of individual plants of a hybrid swarm occurring in their overlap zone could be determined. The hybrids were either F1 plants or backcrosses to the *P. angustifolia* parent (Keim *et al.* 1989). The topology of the relationship among four species of *Actinidia* based on nuclear RFLPs with cDNA and random genomic clones agrees with that based on RFLPs in the cpDNA (Crowhurst *et al.* 1990).

MINI-SATELLITE VARIATION: VNTR FINGERPRINTS

Various relatively short DNA sequences seem to cause frequent irregularities in DNA replication and/or recombination wherever they occur (Levinson & Gutman 1987). Such sequences are found as moderately repetitive tandem repeats ('mini-satellites') of more or less identical units that are up to about 60 bp in length (Jeffreys *et al.* 1985). Many such mini-satellites seem to occur at various positions throughout all eukaryote genomes. Martienssen & Baulcombe (1989) illustrate a mini-satellite in the wheat genome that consists of 21 repeats of 15–26 bp in length each containing a 14 bp 'core sequence' (averaged over all 21 repeats, there are 12·1 of 14 bp identical with the core consensus).

The number of repeats in a particular mini-satellite can vary among the individuals of a population (variable number tandem repeats: VNTRs). This variation is detected as restriction fragment length variation when DNA is digested with restriction enzymes that do not cut within the repeat. As VNTRs based on one basic sequence regularly occur at various places in the genome, a probe detecting the repeat sequence usually reveals several or many restriction fragments per individual. Some of these may be alleles from loci heterozygous for repeat numbers. Due to the high rate of length variation, the

corresponding pattern of bands often is specific for an individual and is called a 'DNA fingerprint'.

Several probes containing core sequences of VNTRs have been used for fingerprinting in various organisms. The first demonstrations of the existence of mini-satellites in plants made use of a human mini-satellite probe (Dallas 1988; Rogstad *et al.* 1988a) or of a probe derived from bacteriophage M13 (Rogstad *et al.* 1988b, 1991a,b; Ryskov *et al.* 1988; Nybom *et al.* 1989; Zimmerman *et al.* 1989). Relatively simple and reliable techniques for detecting fingerprints in plants make use of synthetic oligonucleotides that can be four-base repeats such as $(GATA)_4$ or $(GACA)_5$ (Weising *et al.* 1989, 1990, 1991). Such probes are easily available and likely to reveal RFLPs at many loci throughout the genome (Tautz & Renz 1984; Tautz 1989; Condit & Hubble 1991). The great advantage of this method is set against two limitations to its use in phylogenetic analysis.

The most severe limitation is the high rate of change of VNTR lengths. Van Houten *et al.* (1991), using $(GATA)_4$ as a probe, found that very few plants in a population of *Microseris pygmaea* (Asteraceae) had identical fingerprint patterns and that the patterns could change unpredictably between a plant and its offspring from selfing. In the closely related species, *M. elegans*, fingerprint patterns were considerably more stable and could be used to determine the genetic relationship of genotypes isolated from various populations (Van Heusden & Bachmann, 1992a). In a comparison of three species of *Plantago*, Wolff *et al.* (in press) found no intrapopulation variation with the M13 probe in the selfing species *P. major* and high levels of mini-satellite variation in the outcrossing species *P. lanceolata*. Population differentiation for mini-satellites was found in *Acer negundo* by Nybom & Rogstad (1990) and in *Asimina triloba* (Annonaceae; Rogstad *et al.* 1991a). The interpretation of fingerprint patterns confronts us with the second limitation of the method. It is impossible to determine the homology (or the allelism) of the bands without performing breeding experiments. The comparison is usually limited to the presence or absence of a band at a certain position in the gel, and it is assumed that this comes from identical alleles of one locus in all the plants. The typical data matrix will list, for each plant, the absence or presence of all bands found in the entire sample. These data can either be treated as cladistic characters or, using an index of band sharing, in a distance matrix (Lynch 1988, 1990).

The most useful application of mini-satellite profiles in plant evolution seems to be for the determination of clonal identity and evolution in asexually reproducing plant species such as apomictic blackberries (*Rubus*, Nybom & Schaal 1990; Nybom & Hall 1991) or clones of quaking aspen (*Populus tremuloides*, Rogstad *et al.* 1991b). Fingerprinting of apomictic dandelions (*Taraxacum*) has shown that plants of the 'microspecies' *T. hollandicum* from The Netherlands, Czechoslovakia and France are indeed members of one clone, but has thrown doubt on the clonal nature of some other *Taraxacum* microspecies (Van Heusden *et al.* 1991).

The term 'DNA fingerprinting' is not exclusively applied to banding patterns based on mini-satellite length variation, but may signify any individual- or strain-specific banding pattern on a gel based on length polymorphisms of DNA fragments (Gebhardt *et al.* 1989). This may include banding patterns based not on restriction fragments but on amplification products (D'Ovidio *et al.* 1990b; Welsh & McClelland 1990).

POLYMORPHISMS IN PCR AMPLIFICATION PRODUCTS, RAPDs

The introduction of the polymerase chain reaction (PCR) for the *in-vitro* amplification of specific DNA segments out of a mixture of many others (Saiki *et al.* 1985) has greatly

simplified many molecular protocols. PCR has made comparative sequencing of homologous stretches of DNA accessible to systematists, because the purification of homologous sequences from new material, which used to be forbiddingly complex, is now a simple routine. The method depends on sufficient knowledge of the desired sequence to synthesize stretches of about 30 bp length flanking the piece to be amplified with opposite complementarity. These hybridize to the homologous sequences on denatured, single-strand (template) DNA and serve as starting points (primers) for the synthesis of the complementary strand including the site of the opposite primer. A heat-stable DNA polymerase is used that survives the temperatures used for denaturing DNA. Repeated denaturing, primer attachment and complementary strand synthesis eventually produces an enormous amount of double-stranded DNA corresponding precisely to the stretch between the opposing primers. Sequencing the amplification product will then reveal polymorphisms at this locus among the various specimens from which template DNA has been isolated (Barbier *et al.* 1991).

PCR can also be used for a random sampling of sequence variation all over the genome. For this, in place of two long (*c.* 30 bp) primers flanking a specific sequence, one shorter (10 bp) primer with an arbitrary sequence is used. Amplification will depend on the existence somewhere in the genome of two sites that are complementary to the primer and lie on opposite strands at a distance so that the intervening sequence can be filled in by *in vitro* DNA synthesis. The statistical chance for such a constellation is minute, but genomes are large and there is a virtually unlimited variety of possible 10 bp primers. In fact, quite regularly several amplification products of different lengths are obtained from a plant genome with an arbitrary primer and detected as bands on a gel after electrophoresis (Williams *et al.* 1990; Quiros *et al.* 1991; Van Heusden & Bachmann 1992a). These banding patterns are usually polymorphic in populations (random amplified polymorphic DNAs: RAPDs; Williams *et al.* 1990), and the various patterns obtained with different primers form a detailed DNA 'fingerprint' for genome recognition (Welsh & McClelland 1990; Klein-Lankhorst *et al.* 1991). Mendelian analysis shows that the polymorphisms usually consist of the absence or presence of an amplification product from a site on the genome, possibly reflecting base substitutions in the primer attachment sites. The dominant inheritance of RAPDs makes them less convenient genetic markers than the co-dominantly inherited RFLPs, but RAPD amplification products can be mapped (Williams *et al.* 1990; Quiros *et al.* 1991). They can also be used as presence/absence characters for phylogenetic analysis.

Populations of the Californian autogamous annual, *Microseris elegans* (Asteraceae), are geographically isolated from each other and consist essentially of one local 'biotype' each. The morphological and genetic differences among these biotypes have been analysed in inbred strains derived from various populations. The virtual absence of heterozygosity in these strains simplifies their comparison with dominant RAPD markers (Van Heusden & Bachmann 1992a). These revealed monophyletic groups of strains, of which one represented populations that were not geographically nearest neighbours. Apparently, populations can arise after long-distance fruit dispersal to unoccupied sites within the range of the species and then resemble their source populations more than their neighbours. The close genetic relationship among these populations could not be concluded from morphological characters, but once it was known, it could be detected in the first two factors of a principal components analysis of the morphological variation (Bachmann & van Heusden 1992). Similar analyses of the closely related species, *M. bigelovii* (Van Heusden & Bachmann 1992b) and *M. pygmaea* (Van Heusden & Bachmann 1992c) have

revealed subtle differences among the species both in the degree of recombination (gene flow) among populations and in the overall variation within the species. The method seems to be very useful for the analysis of intraspecific variation, at least in inbreeding species. However, Van Heusden & Bachmann (1992c) have shown that there are RAPDs for various taxonomic levels of analysis. In a comparison of the two most divergent strains of each of three species of *Microseris* with 22 primers, 143 bands were shared by at least two strains. Of these, 6, 12, and 14 amplification products were shared by the two strains of each of the three species and are therefore candidates for species identification with DNA. Forty-nine amplification products were present in all six strains and are possibly informative at higher taxonomic levels.

RAPD polymorphisms are typical molecular markers. They are individually of little value, but can become immensely valuable when they give information about many unique loci randomly distributed across the genome. The method is relatively new and not yet fully tested. In fact, theoretical predictions based on the assumption that the primers recognize and reveal all complementary sites in the genome are frequently not met. For instance, combinations of two primers usually do not produce the bands of the two single-primer patterns plus new ones (Klein-Lankhorst *et al.* 1991; Williams *et al.* in press), and a combination of two template DNAs does not necessarily produce the combined bands amplified from the single templates (Williams *et al.* in press). Moreover, if amplification products are isolated from a gel, labelled and hybridized to a blot of this gel, more bands are labelled than are seen in the original ethidium bromide stained gel (Williams *et al.* in press; K. Buchmann, unpublished data). This suggests that there is some sort of selection among various potential sites for amplification during the amplification process. Moreover, it pays to remember that the bands revealed by RAPD primers (and unspecific fingerprint probes) may come from any genome including those of insects and molds infecting the plants under study. Within limits and with relevant controls, though, the method is ideal for a quick and easy genetic survey. As the amplification products can be isolated and cloned, the method can be converted into a rigorous RFLP test.

CONCLUSIONS AND OUTLOOK

The genetic variation in nuclear-DNA markers covers the entire range from polymorphisms among repeated sequences within a single plant to evolutionary stability over long geological periods. The discussion of the ribosomal RNA genes has shown that all levels of variability can occur within different parts of a single gene. Various genomic fractions revealed by the different methods have their typical levels of variability. The examples cited above have shown much variation within the expected range of each fraction, and the level at which a particular marker is informative in a specific plant group needs to be empirically determined.

Most of the recent studies in plant evolution employing nuclear-DNA markers use them at the level of species and populations. There they complement cpDNA data by revealing variation at all levels of genetic relationship, by their biparental inheritance and by their genetic linkage with the genes involved in morphological and physiological evolution. With nuclear markers, the parental genomes in diploid and polyploid hybrids can be recognized (Arnold *et al.* 1990; Kochert *et al.* 1991; Zhang & Dvorak 1991). When polymorphisms in the nuclear DNA of a diploid parental species are found back in the allotetraploid derivative, they suggest a multiple origin of the tetraploid (Soltis & Soltis 1991).

Comparisons of cpDNA and nuclear DNA (or phenotypic) relationships reveal an increasing number of discrepancies. In these cases, a plant contains chloroplasts from one lineage and nuclear markers from another one. This has been shown for wild maize (*Zea*; Doebley 1989), poplars (*Populus*: Smith & Sytsma 1990), Australian species of cotton (*Gossypium*; Wendel *et al.* 1991), *Heuchera* (Soltis *et al.* 1991) and many other plants (Rieseberg & Brunsfeld 1992). Various mechanisms can lead to 'chloroplast capture' (Wendel *et al.* 1991), but these allow no prediction about the frequency of its occurrence.

The ease and reliability of phylogeny reconstruction with a combination of DNA markers, including nuclear ones, begin to exceed those based on morphological characters. Where discrepancies between phylogenies based on DNA and on morphology are found (Sytsma 1990), the morphological evidence more often than not turns out to have been ambiguous. In a few cases, this has already led to a formal taxonomic revision (Jansen *et al.* 1991). The availability of molecular markers uncouples, as far as that is possible, the reconstruction of the historical phylogeny from its interpretation and causal analysis. Temporarily this might draw some taxonomists' attention from phenotypic differences and phenotypic evolution to the new fashionable method. However, molecular phylogeny is likely to become a routine approach very soon. Its real value will become apparent, when it provides the independent phylogenetic background information for the interpretation of phenotypic evolution. With nuclear markers, the genetic basis of phenotypic evolution can be investigated in detail, beginning with a genetic map of the genes involved in phenotypic differences relative to molecular polymorphisms. The analysis of the evolution of maize by Doebley *et al.* (1990) shows the way. The first steps for a similar analysis of the evolution in wild plants have been taken (Vlot *et al.* 1992).

REFERENCES

Appels, R. & Dvorak, K. (1982): Relative rates of divergence of species and gene sequences within the rDNA region of species in the Triticeae: implications for the maintenance of homogeneity of a repeated gene family. *Theor. Appl. Genet.* **63:** 361–365.

Arnold, M.L., Bennett, B.D. & Zimmer, E.A. (1990): Natural hybridization betweeen *Iris fulva* and *Iris hexagona*: pattern of ribosomal variation. *Evolution* **44:** 1512–1521.

Avise, J.C. & Ball, R.M. Jr. (1990): Principles of genealogical concordance in species concepts and biological taxonomy. *Oxford Surveys in Evolutionary Biology* **7:** 45–67.

Bachmann, K. & Van Heusden, A.W. (1992): Phenotypic similarity and genetic relationship in *Microseris elegans* (Asteraceae: Lactuceae). *Biol. Zentralblatt* **111:** 38–53.

Barbier, P., Morishima, H. & Ishihama A. (1991): Phylogenetic relationships of annual and perennial wild rice: probing by direct sequencing. *Theor. Appl. Genet.* **81:** 693–702.

Bellarosa, R., Delre, V., Schirone, B. & Maggini, F. (1990): Ribosomal RNA genes in *Quercus* spp. (Fagaceae). *Plant Syst. Evol.* **172:** 127–139.

Bendich, A.J. & Bolton, E.T. (1967): Relatedness among plants as measured by the DNA-agar technique. *Plant Physiol.* **42:** 959–967.

Bennett, M.D. (1985): Intraspecific variation in DNA amount and the nucleotypic dimension in plant genetics. In: Freeling, M. (ed.): *Plant Genetics.* 283–302. Alan R. Liss, New York.

Bennett, M.D., Smith, J.B. & Heslop-Harrison, J. (1982): Nuclear DNA amounts in angiosperms. *Proc. Roy Soc. Lond.* **B216:** 179–199.

Bernatzky, R. & Tanksley, S.D. (1986): Majority of random cDNA clones correspond to single loci in the tomato genome. *Mol. Gen. Genet.* **203:** 8–14.

Brummer, E.C., Kochert, G. & Bouton, J.H. (1991): RFLP variation in diploid and tetraploid alfalfa. *Theor. Appl. Genet.* **83:** 89–96.

Chase, M.W. & Palmer, J.D. (1989): Chloroplast DNA systematics of lilioid monocots; resources, feasability, and an example from the Orchidaceae. *Am. J. Bot.* **76:** 1720–1730.

Clegg, M.T. & O'Brien, S.J. (eds.). 1990: *Molecular Evolution.* Wiley-Liss, New York.

Condit, R. & Hubble, S.P. (1991): Abundance and DNA sequence of two-base repeat regions in tropical tree genomes. *Genome* **34:** 66–71.

Crawford, D.J. (1990): *Plant Molecular Systematics. Macromolecular Approaches.* Wiley Interscience, New York.

Crowhurst, R.N., Lints, R., Atkinson, R.G. & Gardner, R.C. (1990): Restriction fragment length polymorphisms in the genus *Actinidia* (Actinidiaceae). *Plant Syst. Evol.* **172:** 193–203.

Dallas, J.F. (1988): Detection of DNA 'fingerprints' of cultivated rice by hybridization with a human minisatellite probe. *Proc. Natl. Acad. Sci. USA* **85:** 6831–6835.

Doebley, J. (1989): Molecular evidence for a missing wild relative of maize and the introgression of its chloroplast genome into *Zea perennis*. *Evolution* **43:** 1555–1559.

Doebley, J., Stec, J.A., Wendel, J. & Edwards, M. (1990): Genetic and morphological analysis of a maize-teosinte F2 population: implications for the origin of maize *Proc. Natl. Acad. Sci. USA* **87:** 9888–9892.

Doolittle, R.F. (ed.). (1990): *Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences.* Academic Press, San Diego.

Dover, G.A. (1986): Molecular drive in multigene families: how biological novelties arise, spread, and are assimilated. *Trends Genet.* **2:** 159–165.

D'Ovidio, R., Scarascia Mugnozza, G. & Tanzarella, O.A. (1990a): Ribosomal RNA gene structure in some *Populus* spp. (Salicaceae) and their hybrids. *Plant Syst. Evol.* **173:** 187–196.

D'Ovidio, R., Tanzarella, O.A. & Porceddu, E. (1990b): Rapid and efficient detection of genetic polymorphism in wheat through amplification by polymerase chain reaction. *Plant Mol. Biol.* **15:** 169–171.

Downie, S.R. & Palmer, J.D. (1992): use of chloroplast rearrangements in reconstructing plant phylogeny. In: Soltis, P.S., Soltis, D.E. and Doyle, J.J. (eds): *Molecular Systematics of Plants.* 14–35. Routledge, Chapman & Hall, New York.

Doyle, J.J. (1991): Evolution of higher plant glutamine synthetase genes: tissue specificity as a criterion for predicting orthology. *Mol. Biol. Evol.* **8:** 366–377.

Doyle, J.J. (1992): Gene trees and species trees: molecular systematics as one-character taxonomy. *Syst. Bot.* **17:** 144–163.

Doyle, J.J. & Beachy, R.N. (1985): Ribosomal gene variation in soybean (*Glycine*) and its relatives *Theor. Appl. Genet.* **70:** 369–376.

Doyle, J.J., Beachy R.N. & Lewis. W.H. (1984): Evolution of rDNA in *Claytonia* polyploid complexes. In: Grant W.F. (ed.): Plant Biosystematics. 321–341. Academic Press, Toronto.

Dvorak, J., Zhang, H.-B., Kota, R.S. & Lassner, M. (1989): Organization and evolution of the 5S ribosomal RNA gene family in wheat and related species. *Genome* **32:** 1003–1016.

Ganal, M. & Hemleben, V. (1986a): Comparison of the ribosomal RNA genes in four closely related Cucurbitaceae. *Plant Syst. Evol.* **154:** 63–77.

Ganal, M. & Hemleben, V. (1986b): Different AT rich satellite DNAs in *Cucurbita pepo* and *Cucurbita maxima*. *Theor. Appl. Genet.* **73:** 129–135.

Gebhardt, C., Blomendahl, C., Schachtschabel, U., Debener, T., Salamini, F. & Ritter, E. (1989): Identification of 2*n* breeding lines and 4*n* varieties of potato (*Solanum tuberosum*, ssp. *tuberosum*) with RFLP fingerprints. *Theor. Appl. Genet.* **78:** 16–22.

Gerlach, W.L. & Dyer, T.A. (1980): Sequence organization of the repeating units in the nucleus of wheat which contain 5S rRNA genes. *Nucleic Acids Res.* **8:** 4851–4865.

Gottlieb, L.D. (1977): Electrophoretic evidence and plant systematics. *Ann. Missouri Bot. Garden* **64:** 161–180.

Grabau, E. A. (1985): Nucleotide sequence of the soybean mitochondrial 18S rRNA gene: evidence for a slow rate of divergence in the plant mitochondrial genome. *Plant Mol. Biol.* **5:** 119–124.

Greenlee, J.K., Rai, K.S. & Floyd, A.D. (1984): Intraspecific variation in nuclear DNA content in *Collinsia verna* Nutt. (Scrophulariaceae). *Heredity* **52:** 235–242.

Hamby, R.K. & Zimmer, E.A. (1992): Ribosomal RNA as a phylogenetic tool in plant systematics. In: Soltis, P.S., Soltis, D.E. and Doyle, J.J. (eds): *Molecular Systematics of Plants.* 50–91. Routledge, Chapman & Hall, New York.

Helentjaris, T., King, G., Slocum, M., Siedenstrang, C. & Wegman, S. (1985): Restriction fragment polymorphisms as probes for plant diversity and their development as tools for applied plant breeding. *Plant. Mol. Biol.* **5:** 109–118.

Hemleben, V., Ganal, M., Gerstner, J., Schiebel, K. & Torres, R.A. (1988): Organization and length heterogeneity of plant ribosomal RNA genes. In: Kahl, G. (ed.): *Architecture of Eucaryotic Genes* VCH Verlagsgesellschaft, Weinheim.

Hemleben, V. & Werts, D. (1988): Sequence organization and putative regulatory elements in the 5S rRNA genes of two higher plants (*Vigna radiata* and *Matthiola incana*). *Gene* **62:** 165–169.

Hillis, D.M. & Moritz, C. (1990): *Molecular Systematics.* Sinauer Associates, Sunderland, Mass.

Hudson, R.R. (1990): Gene genealogies and the coalescent process. *Oxford Surveys in Evolutionary Biology* **7:** 1–44.

Jansen, R.K. & Palmer, J.D. (1987): Chloroplast DNA from lettuce and *Barnadesia* (Asteraceae): structure, gene localization, and characterization of a large inversion. *Curr. Genet.* **11:** 553–564.

Jansen, R.K., Wallace, R.S., Kim, K.-J. & Chambers, K.L. (1991): Systematic implications of chloroplast DNA variation in the subtribe Microseridinae (Asteraceae: Lactuceae). *Am. J. Bot.* **78:** 1015–1027.

Jeffreys, A.J., Wilson, V. & Thein, S.L. (1985): Hypervariable 'minisatellite' regions in human DNA. *Nature* **314:** 67–73.

Jorgensen, R.A. & Cluster, P.D. (1988): Modes and tempos in the evolution of nuclear ribosomal DNA: new characters for evolutionary studies and new markers for genetic and population studies. *Ann. Missouri Bot. Garden* **75:** 1238–1247.

Keim, P., Paige, K.N., Whitham, T.G. & Lark, K.G. (1989): Genetic analysis of an interspecific hybrid swarm of *Populus*: occurrence of unidirectional introgression. *Genetics* **123:** 557–565.

Kesseli, R., Ochoa, O. & Michelmore, R. (1991): Variation in RFLP loci in *Lactuca* spp. and origin of cultivated lettuce (*L. sativa*). *Genome* **34:** 2430–2436.

Kim, K.-J. & Mabry, T.J. (1991): Phylogenetic and evolutionary implications of nuclear ribosomal DNA variation in dwarf dandelions (*Krigia*, Lactuceae, Asteraceae). *Plant Syst. Evol.* **177:** 53–69.

King, L.M. & Schaal, B. (1989): Ribosomal DNA variation and distribution in *Rudbeckia missouriensis*. *Evolution* **53:** 1117–1119.

Klein-Lankhorst, R.M., Vermunt, A., Weide, R., Liharska, T. & Zabel, P. (1991): Isolation of molecular markers for tomato (*L. esculentum*) using random amplified polymorphic DNA (RAPD). *Theor. Appl. Genet.* **83:** 108–114.

Kochert, G., Halward, T., Branch, W.D. & Simpson, C.E. (1991): RFLP variability in peanut (*Arachis hypogaea* L.) cultivars and wild species. *Theor. Appl. Genet.* **81:** 565–570.

Lanaud, C., Tezenas du Montcel, H., Jolivot, M.P., Glaszmann, J.C. & Gonzalez de Leon, D. (1992): Variation of ribosomal gene spacer length among wild and cultivated banana. *Hered.* **68:** 147–156.

Landry, B.S., Kesseli, R.V., Farrara, B. & Michelmore, R.W. (1987): A genetic map of lettuce (*Lactuca sativa* L.) with restriction fragment length polymorphism, isozyme, disease resistance and morphological markers. *Genetics* **116:** 331–337.

Lapitan, N.L.V., Ganal, M.W. & Tanksley, S.D. (1991): Organization of the 5S ribosomal RNA genes in the genome of tomato. *Genome* **34:** 509–514.

Learn, G.H., Jr. & Schaal, B.A. (1987): Population subdivision for ribosomal DNA repeat variants in *Clematis fremontii*. *Evolution* **41:** 433–438.

Levinson, G. & Gutman, G.A. (1987): Slipped-strand mispairing: major mechanism for DNA sequence evolution. *Mol. Biol. Evol.* **4:** 203–221.

Li, W.H. & Graur, D. (1991): *Fundamentals of Molecular Evolution*. Sinauer Associates. Sunderland, Mass.

Lynch, M. (1988): Estimation of relatedness by DNA fingerprinting. *Mol. Biol. Evol.* **5:** 584–599.

Lynch, M. (1990): The similarity index and DNA fingerprinting. *Mol. Biol. Evol.* **7:** 478–484.

Martienssen, R.A. & Baulcombe, D.C. (1989): An unusual wheat insertion sequence (WIS*1*) lies upstream of an α-amylase gene in hexaploid wheat, and carries a 'minisatellite' array. *Mol. Gen. Genet.* **217:** 401–410.

Martin, P.G. & Dowd, J.M. (1991): A comparison of 18S ribosomal RNA and rubisco large subunit sequences for studying angiosperm phylogeny. *J. Mol. Evol.* **33:** 274–282.

McIntyre, C.L., Clarke, C. & Appels, R. (1988): Amplification and dispersion of repeated DNA sequences in the Triticeae. *Plant Syst. Evol.* **160:** 39–59.

Neuhausen, S.L. (1992): Evaluation of restriction fragment length polymorphism in *Cucumis melo*. *Theor. Appl. Genet.* **83:** 379–384.

Nybom, H. & Hall, H.K. (1991): Minisatellite DNA 'fingerprints' can distinguish *Rubus* cultivars and estimate their degree of relatedness. *Euphytica* **53:** 107–114.

Nybom, H. & Rogstad, S.H. (1990): DNA 'fingerprints' detect genetic variation in *Acer negundo* (Aceraceae). *Plant Syst. Evol.* **173:** 49–56.

Nybom, H. & Schaal, B.A. (1990): DNA 'fingerprints' reveal genotypic distributions in natural populations of blackberries and raspberries (*Rubus*, Rosaceae). *Am. J. Bot.* **77:** 883–888.

Nybom, H., Schaal, B.A. & Rogstad, S.H. (1989): DNA 'fingerprints' can distinguish cultivars of blackberries and raspberries. *Acta Horticulturae* **262:** 305–310.

Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S., Inokuchi, H. & Ozeki, H. (1986): Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. *Nature* **332:** 572–574.

Palmer, J.D. (1985): Evolution of chloroplast and mitochondrial DNA in plants and algae. In: MacIntyre, R.J. (ed.): *Monographs in Evolutionary Biology, Molecular Evolutionary Genetics*. 131–240. Plenum, New York.

Palmer, J.D., Jansen, R.K., Michaels, H.J., Chase, M.W. & Manhart, J.R. (1988): Chloroplast DNA variation and plant phylogeny. *Ann. Missouri Bot. Garden* **75:** 1180–1206.

Pamilo, P. & Nei, M. (1988): Relationship between gene trees and species trees. *Mol. Biol. Evol.* **5:** 568–583.

Paterson, A.H., Lander, E.S., Hewitt, J.D., Peterson, S., Lincoln, S.E. & Tanksley, S.D. (1988): Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction length polymorphisms. *Nature* **335:** 721–726.

Price, H.J. (1988): DNA content variation among higher plants. *Ann. Missouri Bot. Garden* **75:** 1248–1257.

Price, H.J. (1991): Genomic stress, genome size, and plant adaptation. In: Taylor, E. Jr., Pitelka, L.F. and Clegg, M.T. (eds.): *Ecological Analysis of Air Pollution*. 277–287. Springer, New York.

Quiros, C.F., Hu, J., This, P., Chevre, A.M. & Delseny, M. (1991): Development and chromosomal localization of genome specific markers by polymerase chain reaction in *Brassica*. *Theor. Appl. Genet.* **82:** 627–632.

Reddy, P., Appels, R. & Baum, B.R. (1990): Ribosomal DNA spacer length variation in *Secale* spp. (Poaceae). *Plant Syst. Evol.* **171:** 205–220.

Rieseberg, L.H. & Brunsfeld, S.J. (1992): Molecular evidence and plant introgression. In: Soltis, P.S., Soltis, D.E. and Doyle, J.J. (eds): *Molecular Systematics of Plants*. 151–176. Chapman & Hall, New York.

Rieseberg, L.H. Soltis, D.E. & Palmer, J.D. (1988): A molecular reexamination of introgression between *Helianthus annuus* and *Helianthus bolanderi* (Compositae). *Evolution* **42:** 227–238.

Rogers, S.O. & Bendich, A.J. (1987): Ribosomal RNA genes in plants; varability in copy number and in the intragenic spacer. *Plant Mol. Biol.* **9:** 509–520.

Rogstad, S.H., Nybom, H. & Schaal, B. (199b): The tetrapod 'DNA fingerprinting' M13 repeat probe reveals genetic diversity and clonal growth in quaking aspen (*Populus tremuloides, Salicaceae*). *Plant Syst. Evol.* **175:** 115–123.

Rogstad, S.H., Patton, J.C. & Schaal, B.A. (1988a): A human minisatellite probe reveals RFLPs among individuals of two angiosperms. *Nucleic Acids Res.* **16:** 11378.

Rogstad, S.H., Patton, J.C. & Schaal, B.A. (1988b): M13 probe detects DNA minisatellite-like sequences in gymnosperms and angiosperms. *Proc. Natl. Acad. Sci. USA* **85:** 9176–9178.

Rogstad, S.H., Wolff, K. & Schaal, B. (1991a): Geographical variation in *Asimina triloba* Dunal (Annonaceae) revealed by the M13 'DNA fingerprint' probe. *Am. J. Bot.* **78:** 1391–1396.

Ryskov, A.P., Jincharadze, A.G., Prosnyak, M.I., Ivanov, P.L. & Limborskaya, S.A. (1988): M13 phage DNA as a universal marker for DNA fingerprinting of animals, plants, and microorganisms. *FEBS Letters* **233:** 388–392.

Saghai-Maroof, M.A., Soliman, K.M., Jorgensen, R.A. & Allard, R.W. (1984): Ribosomal DNA spacerlength polymorphism in barley: Mendelian inheritance, chromosomal location, and population dynamics. *Proc. natl. Acad. Sci USA* **81:** 8014–8018.

Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A. & Arnheim, N. (1985): Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **230:** 1350–1354.

Schaal, B.A., Leverich, W.J. & Nieto-Sotelo, J. (1987): Ribosomal DNA variation in the native plant *Phlox divaricata*. *Mol. Biol. Evol.* **4:** 611–621.

Schaal, B.A., O'Kane, Jr., S.L. & Rogstad, S.H. (1991): DNA variation in plant populations. *Trends Ecol. Evol.* **6:** 329–333.

Schmidt, T., Jung, C. & Metzlaff, M. (1991): Distribution and evolution of two satellite DNAs in the genus *Beta*. *Theor. Appl. Genet.* **82:** 793–799.

Shattuck-Eidens, D.M., Bell, R.N., Neuhausen, S.L. & Helentjaris, T. (1990): DNA sequence variation within maize and melon: observations from polymerase chain reaction amplification and direct sequencing. *Genetics* **126:** 207–217.

Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matsubayashi, T., Zaita, N., Chungwongse, J., Obokata, J., Yamaguchi-Shinozaki, K., Ohto, C., Tovazawa, K., Meng, B.Y., Sugita, M., Deno, H., Kamogashira, T., Yamata, K., Kusuda, J., Takaiwa, F., Kato, A., Tohdoh, N., Shimada, H. & Sugiuva, M. (1986): The complete nucleotide sequence of tobacco chloroplast genome: its gene organization and expression. *EMBO J.* **5:** 2043–2049.

Smith, R.L. & Sytsma, K.J. (1990): Evolution of *Populus nigra* (Sect. *Aigeros*): introgressive hybridization and the chloroplast contribution of *Populus alba* (Sect. *Populus*). *Am. J. Bot.* **77:** 1176–1179.

Soltis, D.E., Soltis, P.S., Collier, T.G. & Edgerton, M.L. (1991): Chloroplast DNA variation within and among genera of the *Heuchera* group (Saxifragaceae): evidence for chloroplast transfer and paraphyly. *Am. J. Bot.* **78:** 1091–1112.

Soltis, P.S. & Soltis, D.E. (1991): Multiple origins of the allotetraploid *Tragopogon mirus* (Compositae): rDNA evidence. *Syst. Bot.* **16:** 407–413.

Soltis, P.S., Soltis, D.E. & Doyle, J.J. (eds.). (1992): *Molecular Systematics of Plants*. Routledge, Chapman & Hall, New York.

Song, K.M., Osborn, T.C. & Williams, P.H. (1988a): *Brassica* taxonomy based on nuclear restriction fragment length polymorphisms (RFLPs). 1. Genome evolution of diploid and amphidiploid species. *Theor. Appl. Genet.* **75:** 784–794.

Song, K.M., Osborn, T.C. & Williams, P.H. (1988b): *Brassica* taxonomy based on nuclear restriction fragment length polymorphisms (RFLPs). 2. Preliminary analysis of subspecies within *B. rapa* (syn. *campestris*) and *B. oleracea*. *Theor. Appl. Genet.* **76:** 593–600.

Song, K.M., Osborn, T.C. & Williams, P.H. (1990): *Brassica* taxonomy based on nuclear restriction fragment length polymorphisms (RFLPs). 3. Genome relationships in *Brassica* and related genera and the origin of *B. oleracea* and *B. rapa* (syn. *campestris*). *Theor. Appl. Genet.* **79:** 497–506.

Southern, E.M. (1975): Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98:** 503–517.

Sytsma, K.J. (1990): DNA and morphology: inference of plant phylogeny. *Trends Ecol. Evol.* **5:** 104–110.

Sytsma, K.J. & Schaal, B.A. (1985): Phylogenetics of the *Lisianthius skinneri* (Gentianaceae) complex in Panama using DNA restriction fragment analysis. *Evolution* **39:** 594–608.

Sytsma, K.J. & Schaal, B.A. (1990): Ribosomal DNA variation within and among individuals of *Lisianthius* (Gentianaceae) populations. *Plant Syst. Evol.* **170:** 97–106.

Tautz, D. (1989): Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Res.* **17:** 6463–6471.

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

Torres, R.A. & Hemleben, V. (1991): Use of ribosomal DNA spacer probes to distinguish cultivars of *Cucurbita pepo* L. and other Cucurbitaceae. *Euphytica* **53:** 11–17.

Torres, R.A., Zentgraf, U. & Hemleben, V. (1989): Species and genus specificity of the intergenic spacer (IGS) in the ribosomal RNA genes of Cucurbitaceae. *Z. Naturforschg.* **44c:** 1029–1034.

Unfried, K., Schiebel, K. & Hemleben, V. (1991): Subrepeats of rDNA intergenic spacer present as prominent independent satellite DNA in *Vigna radiata* but not in *Vigna angularis*. *Gene* **99:** 63–68.

Van Heusden, A.W. & Bachmann, K. (1992a): Genotype relationships in *Microseris elegans* (Asteraceae, Lactuceae) revealed by DNA amplification from arbitrary primers (RAPDs). *Plant Syst. Evol.* **179:** 221–233.

Van Heusden, A.W. & Bachmann, K. (1992b): Nuclear DNA polymorphisms among strains of *Microseris bigelovii* (Asteraceae, Lactuceae) amplified from arbitrary primars. *Bot. Acta* **105:** 331–336.

Van Heusden, A.W. & Bachmann, K. (1992c): Genetic differentiation of *Microseris pygmaea* (Asteraceae, Lactuceae) studied with DNA amplification from arbitrary primers (RAPDs) *Acta Bot. Neerl.* **41:** 385–396.

Van Heusden, A.W., Rouppe van der Voort, J. & Bachmann, K. (1991): Oligo-(GATA) fingerprints identify clones in asexual dandelions (*Taraxacum*, Asteraceae). *Fingerprint News* **3**(2): 13–15.

Van Houten, W.H.J., Van Heusden, A.W., Rouppe van der Voort, J., Raijmann, L. & Bachmann, K. (199): Hypervariable DNA fingerprint loci in *Microseris pygmaea* (Asteraceae, Lactuceae). *Bot. Acta* **104:** 252–256.

Vlot, E.C., Van Houten, W.H.J., Mauthe, S. & Bachmann, K. (1992): Genetic and non-genetic factors influencing deviations from five pappus parts in a hybrid between *Microseris douglasii* and *M. bigelovii* (Asteraceae, Lactuceae). *Int. J. Plant Sci.* **153:** 89–97.

Vodkin, M. & Katterman, F.R.H. (1971): Divergence of ribosomal RNA sequences within angiospermae. *Genetics* **69:** 435–451.

Weising, K., Fiala, B., Ramloch, K., Kahl, G. & Epplen, J.T. (1990): Oligonucleotide fingerprinting in angiosperms. *Fingerprint News* **2**(2): 5–8.

Weising, K., Ramser, J., Kaemmer, D., Kahl, G. & Epplen, J.T. (1991): Oligonucleotide fingerprinting in plants and fungi. In: Burke, T., Dolf, G., Jeffreys, A.J. and Wolff, R. (eds): *DNA Fingerprinting, Approaches and Applications*. Birkhäuser, Basel.

Weising, K., Weigand, F., Driesel, A.J., Kahl, G., Zischler, H. & Epplen, J. (1989): Polymorphic simple GATA/GACA repeats in plant genomes. *Nucleic Acids Res.* **17:** 10128.

Welsh, J. & McClelland, M. (1990): Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Res.* **18:** 7213–7218.

Wendel, J.F., Stewart, J. McD. & Rettig, J.H. (1991): Molecular evidence for homoploid reticulate evolution among Australian species of *Gossypium*. *Evolution* **45:** 694–711.

Williams, J.G.K., Kubelik, A.R., Litvak, K.J., Rafalsky, J.A. & Tingey, S.V. (1990): DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res.* **18:** 6531–6535.

Williams, J.G.K., Rafalski, J.A. & Tingey, S.V. (in press): Genetic analysis using RAPD markers. *Methods in Enzymology*

Wolff, K., Schaal, B.A. & Rogstad, S.H. (in press): Population structure of three *Plantago* species revealed by minisatellite DNA analysis.

Xu, W.W. & Sleper, D.A. (1991): A survey of restriction fragment length polymorphisms in tall fescue and its relatives. *Genome* **34:** 686–692.

Yakura, K., Kato, A. & Tanifuji, S. (1984): Length heterogeneity of the large spacer of the *Vicia faba* rDNA is due to the differing number of a 325 bp repetitive sequence element. *Mol. Gen. Genet.* **193:** 400–405.

Zentgraf, U., King, K. & Hemleben, V. (1992): Repetitive sequences are valuable as molecular markers in studies of phylogenetic relationships within the genus *Cucumis*. *Acta Bot. Neerl.* **41:** 397–406.

Zhang, H.-B. & Dvorak, J. (1991): The genome of tetraploid species of *Leymus* (Poaceae: Triticeae) inferred from variation in repeated nucleotide sequences. *Amer. J. Bot.* **78:** 871–884.

Zimmer, E.A., Hamby, R.K., Arnold, M.L., LeBlanc, D.A. & Theriot, E.C. (1989): Ribosomal RNA phylogenies and flowering plant evolution. In: Fernholm, M., Bremer, K. and Jornvall, H. (eds): *Nobel Symposium 17: The Hierarchy of Life.* 205–214. Elsevier, Amsterdam.

Zimmerman, P.A., Lang-Unnasch, N. & Cullis, C. (1989): Polymorphic regions in plant genomes detected by an M13 probe. *Genome* **32:** 824–828.