DNA replication, recombination, and repair in plastids

Anil Day and Panagiotis Madesis

Abstract

Plastid DNA is conserved, highly polyploid and uniform within a plant reflecting efficient plastid DNA replication/recombination/repair (DNA-RRR) pathways. We will review the current understanding of the DNA sequences, proteins, and mechanisms involved in plastid genome maintenance. This includes analysis of the topological forms of plastid DNA, models of plastid DNA replication, homologous recombination, replication slippage, DNA repair, and plastid DNA-RRR-proteins. We will focus on flowering plants but include information from algae when relevant. Plastid DNA is comprised of a multimeric series of circular, linear, and branched forms. Variant plastid DNA molecules include small linear palindromes with hairpin ends. Plastid transformation has demonstrated an efficient homologous recombination pathway, acting on short ~200 bp sequences, that is active throughout shoot development. These functional studies involving plastid transformation to manipulate DNA sequences, combined with genomics and reverse genetics to isolate mutants in plastid DNA-RRR proteins, will be particularly important for making progress in this field.

1 The importance of DNA replication, recombination, and repair pathways in plastids

All life on earth relies on DNA-replication, recombination, and repair (DNA-RRR) pathways for stable maintenance and propagation of DNA. Plants are dependent on light for growth and are exposed to the damaging effects of radiation. Chloroplasts, the light harvesting plastids of plants and algae, are the sites at which radiation might be expected to cause the greatest damage. Radiation itself and toxic reactive oxygen species, produced as by products of photosynthesis, are examples of destructive agents that can damage plastid DNA (Fig. 1). Despite these damaging agents plastid DNA is relatively well conserved with respect to sequence and gene content (Palmer 1990) and is widely used for phylogenetic studies (Soltis et al. 1999). Plastid genomes are present in multiple copies per cell, which are identical resulting in a uniform population of DNA molecules within an

Topics in Current Genetics, Vol. 19
R. Bock (Ed.): Cell and Molecular Biology of Plastids
DOI 10.1007/4735_2007_0231 / Published online: 13 June 2007

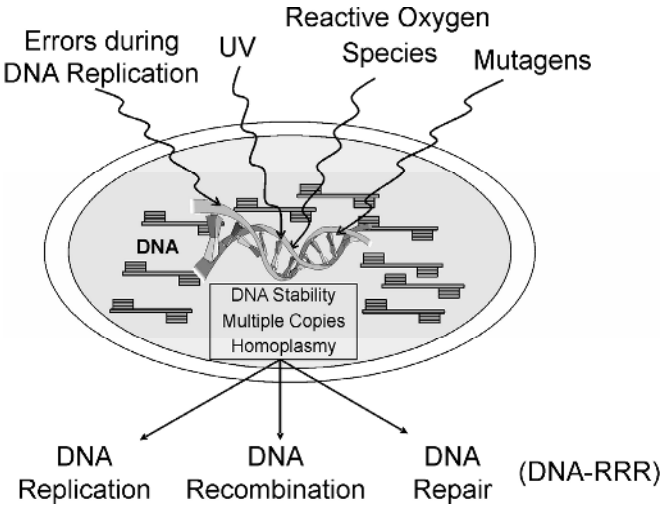

Fig. 1. DNA-RRR pathways are responsible for the high copy number, uniformity, and stable maintenance of plastid genomes.

individual cell or multicellular plant. Highly effective DNA maintenance pathways in plastids must underpin the evolutionary stability and uniformity of plastid DNA.

DNA-RRR pathways must overcome two potential problems associated with the mode of inheritance and ploidy of plastid DNA. First, in sexual crosses plastid DNA often exhibits uniparental inheritance (Corriveau and Coleman 1988; Reboud and Zeyl 1994) reducing the possibility of DNA recombination between parental plastid genomes (Chapter 3). Moreover, in flowering plants, when two plastid types are present in the same cell they rarely recombine (Medgyesy et al. 1985) and segregate away from each other to form cells with pure populations of each plastid type. Segregation of plastids during vegetative growth is known as cytoplasmic sorting or vegetative segregation (Birky 1994). Lack of DNA recombination between different plastid types means that plastids propagate asexually and do not have the benefits of sex and DNA recombination between parental alleles (applicable to nuclear genes) to eliminate deleterious mutations. Muller's ratchet (Muller 1964) would operate leading to an accumulation of mutations in plastid DNA. Second, because plastid DNA is present in multiple copies any new mutations in plastid genes would be masked by the wild type (WT) alleles present in the cell. Asexual propagation and a high degree of polyploidy are two features of plastid DNA that would be expected to promote the accumulation of mutations. Without effective plastid DNA-RRR pathways, plastid mutations would accumulate with time resulting in loss of fitness and death.

1.1 Proteins and DNA targets of plastid DNA-RRR pathways

Plastid DNA maintenance is governed by cis-acting plastid DNA sequences which are the targets for trans-acting proteins that replicate, recombine, and repair plastid genomes. The plastid genomes of green algae and plants that have been characterised do not encode any known DNA-RRR proteins. A number of non-green algae, including diatoms, red and cryptomonad algae contain a plastid *dnaB*-like gene (Kowallik et al. 1995; Reith and Munholland 1995; Douglas and Penny 1999). The *dnaB* gene encodes a DNA helicase involved in replication (Nakayama et al. 1984). In angiosperms the absence of plastid-encoded DNA-RRR proteins is demonstrated by the observation that plastid DNA is replicated in albino cereal (Hess et al. 1994; Zubko and Day 2002) and *Brassica* plants (Zubko and Day 1998) lacking plastid-encoded proteins.

DNA replication, recombination, and repair were once considered to be distinct pathways but more recent work in bacteria has shown they are interrelated processes (Kreuzer 2005). Pathways for recombination-dependent DNA replication and DNA replication-dependent recombination have been described (Kowalczykowski 2000; Kreuzer 2000, 2005) and are applicable to plastid DNA (see Section 7 below). This review will summarise our current knowledge on the mechanisms, DNA sequences and proteins involved in the maintenance of plastid DNA. We will focus on plastid DNA in flowering plants but will include relevant work from algal plastids where appropriate.

2 Plastid DNA polyploidy, packaging, and segregation

2.1 Plastid DNA copy number

One thousand to 1,700 copies of plastid DNA are present per cell in *Arabidopsis thaliana* leaves (Zoschke et al. 2007) whilst five thousand to over ten thousand copies of plastid DNA per cell are present in leaves of *Pisum sativum* (Lamppa and Bendich 1979), *Triticum aestivum* (Day and Ellis 1984), *Spinacia oleracea* (Lawrence and Possingham 1986), and *Hordeum vulgare* (Baumgartner et al. 1989). Fewer plastid genomes per cell are found in other organs containing non-green plastids, such as the roots of *P. sativum* (~500 copies per cell, Lamppa and Bendich 1979) and *T. aestivum* (~300 copies per cell, Day and Ellis 1984). An increase in plastid genome copies is associated with the development of chloroplasts from precursor plastids. Copy number estimates based on quantifying the DNA present in purified plastids from leaf cells of different ages indicate the number of genomes per chloroplast reaches a maximum value in young leaves and then decreases in older cells well before senescence. For example, in the developing primary leaf of four-day-old *H. vulgare* seedlings, plastids in the basal meristem were estimated to contain ~130 genomes, this increased to ~210 genomes in chloroplasts in older cells located one to three cm above the meristem, and decreased to ~50 genomes per chloroplast in the oldest cells in the leaf tip

(Baumgartner et al. 1989). More recent publications also report decreases in genomes per chloroplasts in mature leaves compared to young leaves. Decreases observed include 225 to 106 genomes per chloroplast in *Zea mays* (Oldenburg and Bendich 2004a; Shaver et al. 2006), 135 to 53 genomes per chloroplast in *P. sativum*, 122 to 47 genomes per chloroplast in *Medicago truncatula*, and 190 to 70 genomes per chloroplast in *Nicotiana tabacum* (Shaver et al. 2006). These results are consistent with the idea that replication of plastid DNA takes place predominantly in meristematic cells and leaf primordia (Kuroiwa 1991; Fujie et al. 1994; see Section 13.1) and as plastids divide during leaf development the number of genomes per plastid falls. We know very little about the replication mechanisms regulating the copy number of plastid DNA. Some progress has been made in this field with the recent finding that copy number is influenced by specific plastid DNA sequences. Deletion of the OriA plastid DNA sequence implicated in DNA replication (Section 4.1 below) reduces the copy number of plastid DNA in developing leaves of *N. tabacum* (Scharff and Koop 2007).

Recent publications detailing two to threefold reductions in DNA levels per plastid during leaf maturation also suggest the apparent absence of DNA in some mature chloroplasts: DNA was not observed in approximately 11% of *M. truncatula*, 9% of *P. sativum*, 80-90% of *Z. mays* (Shaver et al. 2006) and 29% of *A. thaliana* chloroplasts (Rowan et al. 2004). Loss of DNA from chloroplasts during leaf ageing was not observed in *N. tabacum* (Shaver et al. 2006). Based on the results obtained in *Z. mays* and *A. thaliana* a mechanism that actively degrades DNA in maturing chloroplasts was proposed by the authors (Oldenburg and Bendich 2004a; Rowan et al. 2004). These results appear to suggest that chloroplasts lacking DNA retain photosynthetic activity for long periods (Oldenburg and Bendich 2004a; Rowan et al. 2004), which conflicts with our current understanding of the importance of plastid gene expression for maintaining chloroplast functions. An alternative explanation for the apparent absence of DNA in isolated plastids is that it is an artefact resulting from the experimental approaches used (Li et al. 2006). In particular, degradation of plastid DNA during the purification of plastids or during treatment of plastids with DNase I (to remove contaminating extraplastidic DNA outside plastids) will give rise to low copy number estimates. This might be more problematic for old leaf cells of some species where the release of DNA nucleases during homogenisation and changes in plastid porosity might allow nucleases to enter plastids. DNA fluoresces when stained with the DNA-binding dye 4',6-diamidino-2-phenylindole (DAPI). DAPI stained chloroplasts in leaf sections of *A. thaliana* and *N. tabacum* fixed immediately after sectioning appeared to show the same pattern of DNA reduction or loss obtained with isolated chloroplasts supporting the data with isolated chloroplasts (Shaver et al. 2006). Quantitation of plastid DNA levels in total DNA preparations from liquid nitrogen frozen leaves using Southern blot analysis is less sensitive to plastid DNA degradation during sample preparation and provides a robust method for estimating plastid DNA levels. Such an approach using rapidly extracted total DNA can be used to confirm or dismiss the findings based on isolated chloroplasts. Copy number estimates can also be obtained using quantitative real-time PCR on total DNA with controls to rule out contamination or amplification of 'promiscuous' plastid

DNA sequences present in mitochondria and nuclei (Zoschke et al. 2007). Using these methods it appears that once chloroplast development is completed the levels of plastid DNA in total DNA appear to remain relatively constant during further leaf development in *A. thaliana* (Zoschke et al. 2007; Li et al. 2006) and *N. tabacum* (Li et al. 2006). These results do not support the idea of a dramatic reduction in plastid DNA levels during leaf development in *A. thaliana*.

Replication of plastid and nuclear DNA do not appear to be tightly co-ordinated in *N. tabacum* (Heinhorst et al. 1985) or the green alga *Chlamydomonas reinhardtii* (Chiang and Sueoka 1967). In contrast to the stringent controls restricting nuclear DNA synthesis to one round of replication during the S phase of each cell cycle, plastid DNA replication appears to be less stringent and is not limited to the S phase (Heinhorst and Cannon 1993). Moreover, plastid genomes appear to be chosen randomly for replication (Birky 1994). The prereplication factor CDT1 appears to affect both nuclear DNA replication and plastid division in *A. thaliana* and provides a possible link between the cell cycle and plastid division (Raynaud et al. 2005). In synchronous cultures of *C. reinhardtii*, duplication of plastid DNA could be localised to a particular time period (Chiang and Sueoka 1967) whereas plastid DNA synthesis, monitored by ^{32}P incorporation, was observed throughout the cell cycle (Grant et al. 1978). The ^{32}P-incorporation was suggested to be due to DNA repair activities which were required to maintain plastid genomes throughout the cell cycle (Grant et al. 1978).

2.2 Packaging of plastid DNA

Within plastids, the DNA is not dispersed but localised into aggregates of DNA and protein called nucleoids (Kuroiwa 1991; Sakai et al. 2004). The uniformity of plastid DNA is governed by DNA-RRR pathways that are likely to be carried out in nucleoids. The organisation of multiple plastid genome copies into a smaller number of units will govern the segregation of plastid DNA during plastid and cell divisions (VanWinkle-Swift 1980) and will facilitate cytoplasmic sorting. The number, sizes, morphologies, and distribution of nucleoids, visualised by DAPI staining, vary during development of chloroplasts from proplastids (Miyamura et al. 1986). *T. aestivum* proplastids contain one to ten nucleoids and 30 to 40 plastid genomes whereas chloroplasts contain ten to thirty nucleoids and 70 to 100 plastid genomes (Miyamura et al. 1990). In *Nicotiana*, mature chloroplasts contain eight to forty nucleoids, each with about ten plastid genomes (Kuroiwa 1991). Nucleoids appear to be located in the stroma or attached to the envelope or thylakoids depending on the plastid type (Sato et al. 2003; Sakai et al. 2004). The functional significance of changes in the intra-plastidic location of nucleoids is not known. However, it is interesting to note that constitutive expression of the *B. napus* homologue of the *P. sativum* plastid envelope DNA (PEND) binding protein in *N. tabacum* nuclear transformants leads to an albino phenotype possibly due to a lack of release of DNA from the envelope (Wycliffe et al. 2005). The PEND protein is targeted to plastids and might be involved in anchoring plastid DNA to the inner envelope during early chloroplast development (Sakai et al. 2004),

2.3 Segregation of plastid genomes

The replication and segregation mechanisms in plastids prevent the persistence of two different plastid genomes in cells. Heteroplasmy can only be maintained for long periods of time by selection for both plastid genomes (Drescher et al. 2000; Shikanai et al. 2001; Kode et al. 2005). The fine details of heteroplasmy are not known and the two plastid DNA types might be mixed within single nucleoids, or localised to separate nucleoids within a plastid or be separated into two populations of plastids within a cell. Heteroplasmy within a plastid is required when a lethal mutation is plastid autonomous (Kode et al. 2005) and cannot be rescued by import of cytoplasmic metabolites.

3 Topological forms of plastid DNA

The mechanisms of plastid DNA replication and maintenance will be reflected in the topologies of DNA molecules found in plastids. Plastids are most likely to be descendents of ancient cyanobacteria (Martin et al. 2002), which contain circular double-stranded DNA genomes. Circular DNA overcomes the problems of replicating gaps at the ends of linear DNA molecules following RNA primer removal at the 5' ends of newly synthesized DNA (Cavalier-Smith 1974). The sequence maps of all the plastid genomes that have been characterized are circular (Chapter 3). In the majority of species the genomes can be represented as a single circular double-stranded DNA molecule containing all genes. Dinoflagellates are an exception and contain genes dispersed over a number of DNA mini-circles each with one to three genes (Koumandou et al. 2004). A circular sequence or restriction map does not necessarily imply the physical structure of a DNA species is a circle (Streisinger et al. 1964). Tandemly repeated DNA sequences (Fig. 2a), such as nuclear ribosomal RNA genes, on a linear chromosome or circularly permuted sequences arranged on separate linear DNA molecules of defined (Fig. 2b) or varying lengths (Fig. 2c) will also give rise to circular maps (Fig. 2d). To study the structure of plastid DNA requires the analysis of intact DNA isolated from chloroplasts. Because double-stranded DNA is prone to breakage by shearing, the analysis of plastid DNA topology requires distinguishing breakage products of the extraction process from intact plastid DNA molecules. Most studies have involved chloroplasts, which are easily identified, abundant in leaves and relatively easy to purify. The structure of chloroplast DNA has been studied by microscopic and gel electrophoretic methods.

An early electron microscopic study showed that monomer circles corresponding in size to a single set of plastid genes represented 37% of the DNA extracted from *P. sativum* chloroplasts (Kolodner and Tewari 1972). The remaining DNA

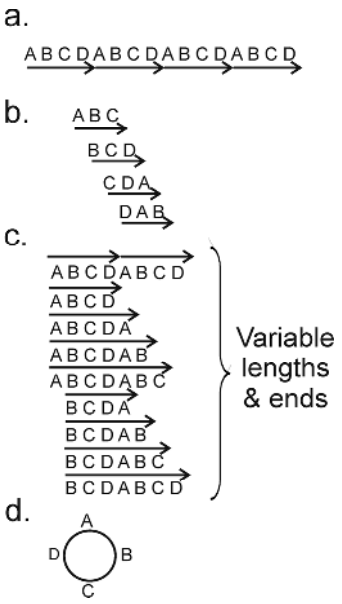

Fig. 2. a) Tandemly repeated linear DNA sequences, b) Circularly permuted linear DNA sequences of fixed length, and c) Multimeric linear DNA molecules of variable sizes and a number of dispersed ends will give rise to d) Circular sequence maps. Arrows indicate orientation of sequences. Molecules starting with the letters c or d are not shown in c).

was mainly comprised of sub-genomic linear DNA forms, which could represent breakage products of circular plastid DNA molecules or even contaminating nuclear DNA (Fig. 3a). Breakage during extraction would give rise to variable ratios of circles to linear products with each preparation. In a later study, three to four percent of circular species were found to be dimers (Kolodner and Tewari 1979). These were arranged head-to-tail in *P. sativum* chloroplasts, and both head-to-head and head-to-tail in *S. oleracea* and *Lactuca sativa* chloroplasts. The majority of *L. sativa* and *S. oleracea* chloroplast DNA dimers (about 80%) were arranged head-to-head (Kolodner and Tewari 1979; see also Section 8 and Fig. 8a). In an independent study on *S. olereacea* and other dicots 80% of chloroplast DNA molecules were found as monomer circles and 10 to 15% as dimers. About 15% of circles were supercoiled (Herrmann et al. 1975).

An elegant more recent study utilised fluorescence *in situ* hybridization (FISH) involving extended DNA fibres and plastid DNA probes (Lilly et al. 2001). At the time of writing this single report remains the only published source for FISH-based analysis of plastid DNA in flowering plants. Purified chloroplasts were lysed and DNA fixed directly on a slide before hybridization providing less opportunity for DNA breakage. Using this method, chloroplast DNA from *A. thaliana* and *N. tabacum* was found to be comprised of a multimeric series of circular and linear DNA molecules (Fig. 3b). Circles comprised about 40-50% and linear DNA species about 20-25% of chloroplast DNA molecules. The remaining molecules

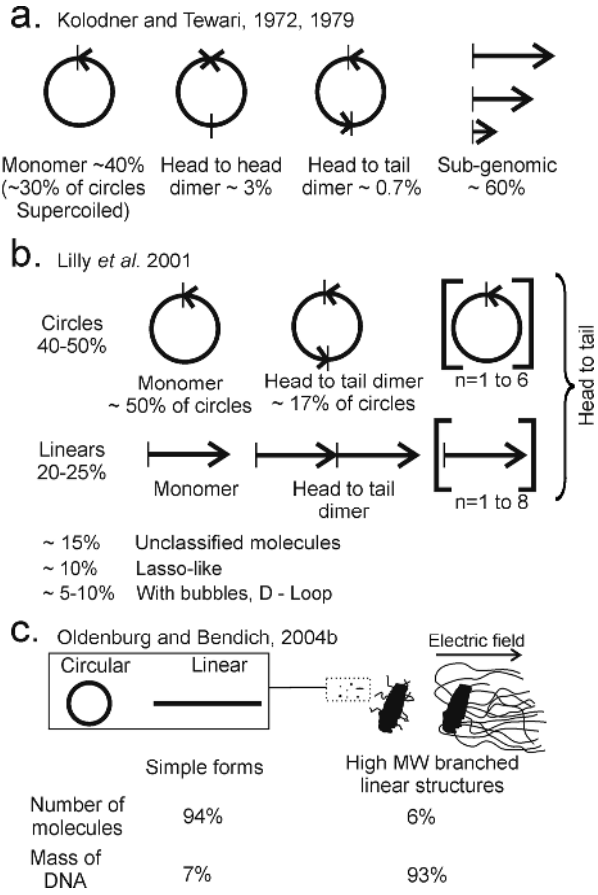

Fig. 3. a) Topological forms of plastid DNA revealed by electron microscopy (Kolodner and Tewari 1972). Only head-to-tail dimers were found in *P. sativum* while head-to-head dimers were predominant in *L. sativa* and *S. oleracea* (Kolodner and Tewari 1979). Breakage of circles during extraction will give rise to a variable percentage of circles and sub-genomic linear forms. b) Topological forms of plastid DNA in *A. thaliana* and *N. tabacum* revealed by DNA fibre-based FISH with plastid DNA probes (Lilly et al. 2001). Only the monomer and dimer are shown in the circular and linear multimeric series. Arrows indicate sequence orientation. c) Structures of plastid DNA molecules in fluorescent images of ethidium-stained DNA from purified *Z. mays* chloroplasts lysed in agarose plugs. DNA fibres in compact high MW structures were extended by an electric field or flowing liquid (Oldenburg and Bendich 2004b).

were unclassified (15%), lasso-like (~10%) or contained bubbles or D-loops (5-10%). In *N. tabacum* chloroplasts (Lilly et al. 2001), monomeric circles are the most abundant (55%) followed by dimers (17%), trimers (10%), tetramers (7%), pentamers (5%), and hexameric circles (1%). Rare higher-order multimers of unit genome sized (genome size = 156 kbp) linear DNA molecules were found up to

the octomer. In *A. thaliana* and *N. tabacum* multimers were comprised of monomers arranged head-to-tail (Lilly et al. 2001) in contrast to the earlier results of predominantly head-to-head dimers in *L. sativa* and *S. oleracea* (Kolodner and Tewari 1979). The ends of the linear DNA molecules were not mapped and they could represent real linear plastid DNA species or the breakage products of large circles (Lilly et al. 2001). A number of lines of evidence indicate that linear DNA molecules found in plastids are not simply breakage products of large circles (see Section 3.1, 3.2, and 7 below). Using a similar DNA fibre-based FISH method, *C. reinhardtii* plastid DNA was found to be mainly comprised of monomeric and dimeric linear and circular forms (Maul et al. 2002). The multimeric series of linear and circular DNA molecules found in plastids (Fig. 3b) must result from the action of plastid DNA-RRR pathways.

Embedding cells in agarose plugs prior to cell and chloroplast lysis reduces DNA breakage and allows the isolation of large DNA molecules. DNA in agarose plugs can then be analysed by pulsed-field gel electrophoresis or microscopy after staining with ethidium bromide. Circular DNA does not enter pulsed field gels and remains within the agarose plugs at the origin at the relatively short pulse times used to fractionate DNA in the 100-1000 kbp range (Bendich and Smith 1990; Backert et al. 1995). Linear chloroplast DNA molecules enter the agarose gel and can be identified by blot hybridization with chloroplast DNA probes. Some of these linear DNA molecules might result from breakage of circular DNA molecules (Backert et al. 1995; Bendich 2004). A multimeric series comprised of monomer (most abundant) and higher molecular weight (MW) linear plastid DNA forms can be visualised on pulsed-field gels. The largest multimers found were tetramers for *S. oleracea* (Deng et al. 1989) and *N. tabacum* (Lilly et al. 2001), dimers (Lilly et al. 2001) or trimers for *P. sativum*, and up to the octomer for *Citrullus vulgaris* plastid DNA (watermelon, Bendich and Smith 1990). The banding pattern can be disrupted by altering the activities of plastid DNA-RRR proteins. Inhibition of plastid-targeted gyrase (see Section 13.4 below), which is required to decatenate newly replicated DNA, reduces the levels of discrete bands corresponding to the monomer and dimer, and gives rise to a heterogeneous mixture of plastid DNA molecules, some of which are greater than 1000 kbp in size on pulsed field gels (Cho et al. 2004).

Whilst the bands seen on pulsed-field gels were useful for visualising multimers of plastid DNA they represent a minor proportion of plastid DNA and give a distorted view of the topological forms of plastid DNA molecules (Bendich 2004). The bulk of plastid DNA molecules including circles, high MW linear branched forms (Bendich 2004; Oldenburg and Bendich 2004b), tangled DNA fibres and any DNA in unlysed plastids remains immobile in the agarose plugs at the origin and does not enter pulsed field gels. *N. tabacum* leaf chloroplast DNA remained at the origin (migration into the gel was not detected) whereas about 35% of *Chenopidium album* plastid DNA from a non-green suspension-culture entered the gel revealing monomer and dimer bands (pulse times of 30-60 seconds, Backert et al. 1995). The presence of electrophoretically-mobile linear plastid DNA molecules in *C. album* non-pigmented plastids but not in *N. tabacum* chloroplasts might reflect changes in plastid DNA topologies in different plastid

types, also indicated from other studies on *Z. mays* (Oldenburg and Bendich 2004a), or result from breakage during extraction. Mild DNase I treatment of high MW *N. tabacum* chloroplast DNA and blot hybridization with a plastid probe revealed a smear of DNA (representing molecules of different lengths) within which discrete monomer to tetramer bands were clearly visualised (Backert et al. 1995). These discrete linear bands are likely to be derived from circular DNA because a single double strand break mediated by DNase I will convert a circle to its linear form.

The structures of plastid DNA molecules in agarose plugs prepared from 10-14 day old *Z. mays* seedlings have been studied by fluorescence microscopy following ethidium bromide staining (Oldenburg and Bendich 2004b). The DNA was present as simple DNA molecules and high MW DNA complexes with a central core and attached DNA fibres (Fig. 3c). In the presence of an electric field or liquid flow the simple molecules migrate whereas linear fibres extend from the immobile cores of the high MW complexes. Simple DNA molecules are comprised of circles and linear molecules and represent 94% of the DNA molecules but only 7% of the mass of DNA in plastids due to their small sizes relative to the high MW DNA complexes. The high MW complexes contained on average a minimum of eight plastid genomes (not including bright fluorescent cores) and were suggested to be largely comprised of linear and complex-branched molecules (Oldenburg and Bendich 2004b). A reduction in high MW complexes and an increase in simple forms were correlated with chloroplast maturation during leaf development in *Z. mays* (Oldenburg and Bendich 2004a). The multigenome complexes were reported to represent 93% of plastid DNA by mass. Following removal of linear DNA molecules from multigenome complexes by pulsed-field gel electrophoresis the immobile high MW core was suggested to be comprised of complex-branched DNA structures representing 50% of the mass of DNA in plastids (Oldenburg and Bendich 2004b). The complex high MW branched forms have been suggested to represent replication intermediates and their analysis is particularly important (Oldenburg and Bendich 2004b; Scharff and Koop 2006). The ~15% of tangled DNA fibres that were unclassified (Fig. 3b) by Lilly et al. (2001) might correspond to these high MW DNA complexes identified by Oldenburg and Bendich (2004b). DNA fibre-based FISH using plastid DNA probes would confirm the presence of plastid DNA in these high MW complexes and might be a useful tool to study their sequence organisation.

3.1 Linear hairpin DNA molecules in plastids

Genuine linear plastid DNA molecules can be distinguished from linear products of broken DNA circles by studying their ends. Breakage products would be expected to possess ends that map to randomly selected regions of the plastid genome, and these ends would be expected to be indistinguishable from double-strand breaks with flush or short single-stranded 5' or 3' DNA extensions. Analysis of plastid DNA deletion mutants in albino cereal plants regenerated from pollen provided the first evidence for the presence of linear plastid DNA molecules

with special ends (Day and Ellis 1985; Ellis and Day 1986). Small linear subgenomic molecules have also been found in albino somatic cells from cereals (Kawata et al. 1997; Zubko and Day 2002) and can represent the predominant plastid DNA species in albino cereal plants. Their abundance facilitates the analysis of their ends, which have been examined in detail.

Small linear plastid DNA molecules are inverted repeat palindromes with hairpin ends, which map to a number of sites in the large single copy region of plastid DNA near the *trn*E(UUC) gene (Ellis and Day 1986; Harada et al. 1992; Kawata et al. 1997). The centres of a subset of these linear palindromes are located between *trn*G(GCC) and *trnf*M(CAU) of the 135 to 140 kbp cereal plastid genome (Ogihara et al. 2000) and retain only 5.2 kbp of plastid DNA (Zubko and Day 2002). Small linear DNA molecules all contain the plastid *trn*E(UUC) gene, which is probably essential for heme synthesis (Howe and Smith 1991; Zubko and Day 2002). Linear hairpin DNA molecules are also found in eubacteria including the spirochete genus *Borrelia* (Casjens 1999) and prophage N15 of *Escherichia coli* (Rybchin and Svarchevsky 1999). Models for the origin of these linear DNA molecules include: strand switching during DNA replication (Ellis and Day 1986), possibly promoted by short inverted repeats (Fig. 4a); repair of double strand DNA breaks by intra-strand annealing at inverted repeats (Fig. 4b; Qin and Cohen 2000); and an *E. coli* linear prophage N15-like mechanism involving two cleavages, sealing DNA ends to form hairpins and resolution of replicated DNA into a linear palindrome (Fig. 4c; Rybchin and Svarchevsky 1999). Hairpin ends provide a mechanism to overcome the end-replication-problem and stabilise the ends of linear DNA molecules by protecting them from nucleases (Cavalier-Smith 1974).

Studies on albino cereal plants demonstrate that plastids contain the enzymes required to maintain and replicate linear DNA molecules. If hairpin molecules play a role in maintenance of intact plastid DNA they should also be found in the green chloroplasts of WT plants. Revealingly, hairpin molecules are found in WT *H. vulgare* chloroplasts. The hairpin ends do not appear to be localised but map to various sites within the plastid genome (Collin and Ellis 1991), which is consistent with their derivation from a population of linear DNA molecules with heterogeneous ends. Dispersed ends that are not defined in location are also found in the heterogeneous populations of linear DNA molecules in plant (Backert and Börner 2000; Oldenburg and Bendich 2001) and *Saccharomyces cerevisiae* (bakers' yeast) mitochondria. Mitochondrial DNA in *S. cerevisiae* is comprised of a polydisperse population of linear DNA molecules ranging in size between the 75 kb monomer and 150 kbp dimer (Williamson 2002).

3.2 Linear plastid DNA molecules with discrete ends in WT plastids

Restriction enzymes that cleave plastid DNA rarely (once or twice) have been used to map the ends of linear molecules in high MW DNA prepared in agarose plugs. When *Z. mays* plastid DNA was cleaved with an enzyme that cuts once, the predicted linear 140 kbp genome band was observed. In addition, discrete smaller

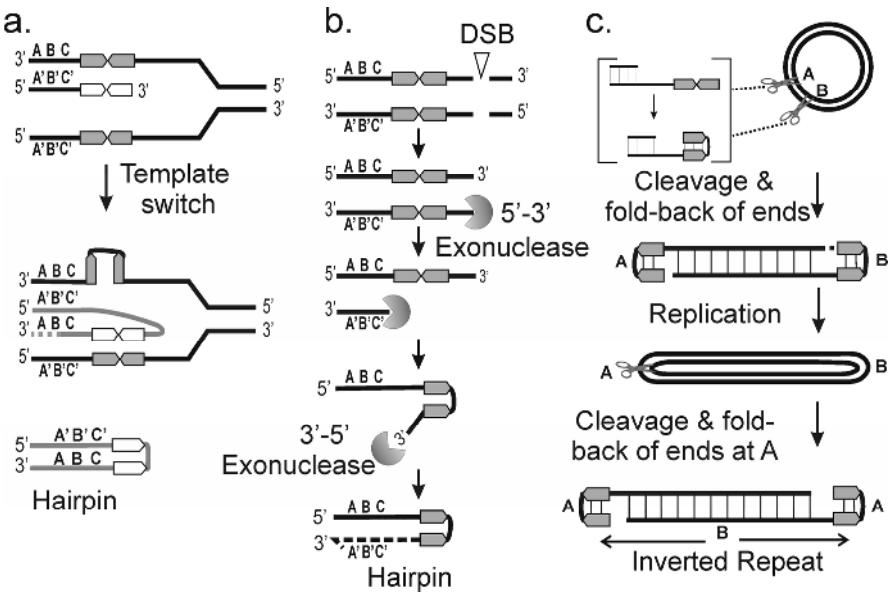

Fig. 4. Models for generation of linear hairpin plastid DNA molecules by a) Template strand switching at the replication fork (Ellis and Day 1986), b) Intra-strand annealing at inverted repeat sequences (Qin and Cohen 2000). A double-strand break (DSB) initiates the pathway, which involves exonuclease and fold-back of single-stranded DNA at inverted repeats. c) A bacteriophage N15-like mechanism involving DNA cleavage at two sites, fold-back and repair of ends to form hairpins and replication followed by cleavage at one site to form a linear palindrome (Rybchin and Svarchevsky 1999). Converging box arrows indicate inverted repeats.

sub-genomic bands were found (Oldenburg and Bendich 2004b). These sub-genomic bands can be explained if they arise from long linear DNA molecules containing one natural end found *in vivo* and one site created by restriction enzyme cleavage. These natural ends in sub-genomic fragments map to the large inverted repeats of *Z. mays* plastid DNA. A similar but more detailed analysis on high MW *N. tabacum* plastid DNA identified eleven natural ends (Scharff and Koop 2006). The majority of breaks mapped to the large inverted repeats but ends were also found in the large and small single copy regions (Scharff and Koop 2006). Some of the *Z. mays* and *N. tabacum* ends map close to plastid DNA sequences promoting DNA synthesis or exhibiting features resembling D-loops or replication bubbles (see Section 4.1 below). Only one end corresponded to the proposed site for initiation of rolling circle replication (Kolodner and Tewari 1975) located at 180˚ from the two D-loops (Section 4; Fig. 5a) in *N. tabacum* (Scharff and Koop 2006). These mapped ends define the termini of linear DNA present in high MW plastid DNA complexes (Oldenburg and Bendich 2004b; Scharff and Koop 2006). The structures of the ends of these linear sections of *Z. mays* and *N. tabacum* plastid DNA are not known but their elucidation (e.g. protected or exposed, hairpin or secondary DNA structure, or simply double-strand

DNA breaks with flush or 5' or 3' protruding ends) is likely to provide information on the mechanisms underlying their formation.

4 A replicon model for plastid genome maintenance

Research on plastid DNA replication has been heavily influenced by the 'replicon model' put forward by Jacob, Cuzin, and Brenner (Jacob et al. 1963). The model proposes a specific DNA element that is recognised by an initiator protein. If plastid DNA replication conforms to the model it would predict initiation of replication at specific sites in plastid DNA. Replication of bacterial genomes and plasmids in bacteria and *S. cerevisiae* conform to the 'replicon model' and involve origins of replication recognised by specific origin recognition proteins (Gilbert 2004). A variety of methods have been used to try and localise origins of replication in plastid genomes. Early electron microscopy (EM) studies on DNA isolated from *P. sativum* and *Z. mays* chloroplasts identified structures resembling D-loops and rolling circles (Kolodner and Tewari 1975). These are well known DNA replication intermediates and provided early models for plastid DNA replication. Unidirectional replication from an origin of replication creates a displacement loop (D-loop), comprised of double stranded DNA and a displaced single stranded DNA loop. Two D-loops spaced 7 kbp apart were found in monomer circles of *P. sativum* plastid DNA and gave rise to the dual D-loop model (Fig. 5a) for initiation of chloroplast DNA replication (Kolodner and Tewari 1975). Convergent replication forks from the two D-loops pass each other and a bidirectional replication bubble is formed once the forks pass the starting points of replication. The complete genome is replicated by the replication forks continuing round the circle in opposite directions, with discontinuous replication on the lagging strands, until they meet at 180^{o} from the origin of D-loop synthesis (Fig. 5a). For rolling circle replication, a replication fork displaces the lagging strand at a nick and continues round the circle (Fig. 5b). Rolling circle replication enables multiple tandem head-to-tail copies of plastid DNA to be made from a single round of replication initiation. The rolling circles appeared to be initiated at the terminus of bidirectional replication (Kolodner and Tewari 1975). This early EM study stimulated research to locate D-loops on sequence maps of plastid DNA.

4.1 Replication origins mapped to the large inverted repeat

Electron-microscopy combined with restriction enzyme digestion enables D-loops and replication bubbles to be mapped onto restriction fragments of plastid DNA. The *P. sativum* D-loops (OriA and OriB) flank the 23S ribosomal RNA gene (Fig. 6a; Meeker et al. 1988). Unlike most angiosperm plastid genomes *P. sativum* lacks a large inverted repeat (Chapter 3). Restriction fragments of proplastid DNA with a high frequency of D-loops from *N. tabacum* BY2 suspension culture cells

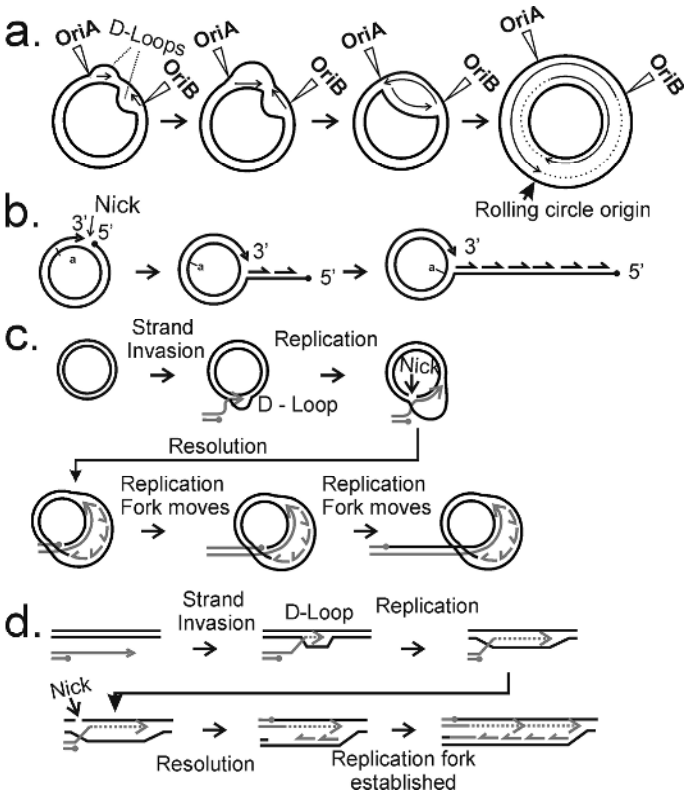

Fig. 5. Plastid DNA replication models. a) Displacement-loop (D-loop) model of plastid DNA replication. Two D-loops converge to give rise to bidirectional replication (Kolodner and Tewari 1975). b) Rolling circle replication arising from strand displacement at a nick. Movement of the replication fork is shown by anti-clockwise rotation of the circle marked by "a" c) Recombination-dependent DNA replication (Kowalczykowski 2000) on a circular template with D-loop gives rise to a bubble-containing circle with tail. d) Recombination-dependent DNA replication on a linear DNA template gives rise to a branched molecule.

mapped close to the end of the 23S rRNA gene in the large inverted repeat (Fig. 6b Nt (pro)) and a less active D-loop mapped to a 2.3 kbp Stu I fragment containing part of the *psaA* and *psaB* genes in the large single copy region (Takeda et al. 1992). Later work, using two dimensional agarose gel electrophoresis to map bubbles in cloned plastid DNA templates replicated in chloroplast fractions, *in vitro* DNA replication assays and primer extension on nascent DNA strands, suggested different positions for two D-loops (named OriA and OriB) in plastid DNA from *N. tabacum* leaves. *N. tabacum* OriA mapped to the intron of the *trnI* (GAU) gene located between the 16S and 23S rRNA genes (Lu et al. 1996; Kunnimalaiyaan and Nielsen 1997a). OriB mapped to the large inverted repeat close to the border of the small single copy region in orf350 or *ycf1* (Kunnimalaiyaan and Nielsen 1997a; Kunnimalaiyaan et al. 1997b).

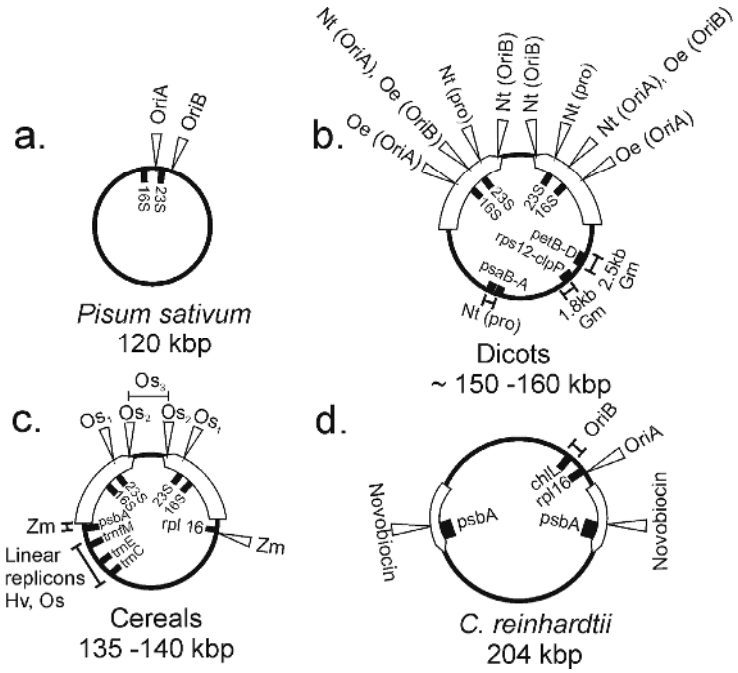

Fig. 6. Schematic diagram showing the locations of potential replication origins. Origin locations are shown outside the circular maps as triangles or bar-ended lines. a) *P. sativum* (Meeker et al. 1988) lacks a large inverted repeat, b) Dicots containing large inverted repeats. *N. tabacum* Nt (Ori A) and Nt (Ori B) (Kunnimalaiyaan and Nielsen 1997; Kunnimalaiyaan et al. 1997), *N. tabacum* D-loops Nt (pro) in proplastids (Takeda et al. 1992), *Oenothera hookeri* Oe (Ori A) and Oe (Ori B) (Chiu and Sears 1992; Sears et al. 1996), *Glycine max* bubbles Gm (Hedrick et al. 1993). c) *Oryza sativa* (Os) replication origins in suspension culture cells (Os_1), leaf blades (Os_2), and coleoptiles (Os_3) mapped by Wang et al. 2003, *Z. mays* (Zm, Gold et al. 1987), linear DNA replicons (Ellis and Day 1986; Harada et al. 1992; Zubko and Day 2002), *Hordeum vulgare* (Hv), *Oryza sativa* (Os). d) *C. reinhardtii* Ori A and Ori B (Waddell et al. 1984; Chang and Wu 2000) and initiation of novobiocin-resistant replication (Woelfle et al. 1993). The large inverted repeats are shown as converging box arrows on maps b-d; arrow orientation according to rrn operon transcription direction.

Two D-loops separated by 4 kbp (Chiu and Sears 1992; Sears et al. 1996) were found in the large inverted repeat of *Oenothera hookeri* plastid DNA, where they flank the 16S ribosomal genes (Fig. 6b). The locations of two origins in the large inverted repeat suggest four potential replication origins in *N. tabacum* and *O. hookeri* plastid DNA; two in each inverted repeat. The complexity of mapping plastid origins is illustrated by the observation that the locations of origins appear to vary in different cells, tissues and organs from the same species. Differences in the mechanism of DNA replication and locations of plastid origins have been observed in suspension culture cells, coleoptiles, and leaves of *Oryza sativa* (rice,

Wang et al. 2003). *O. sativa* plastid replication origins were mapped to the small single copy region and to two positions in the large inverted repeat (Fig. 6c; Wang et al. 2003). Comparisons of mapped origins show that the location of only one is conserved in *N. tabacum* (OriA), *O. hookeri* (OriB), and *P. sativum* (OriA). This is located in the intergenic region between the 16S and 23S rRNA genes (Fig. 6a, 6b; Lu et al. 1996). Conservation in location might suggest this region is important for plastid genome maintenance. However, deletion of OriA in *N. tabacum* using plastid transformation has revealed that it is not essential for plastid DNA replication and maintenance (Mühlbauer et al. 2002). Of the two copies of OriB present in the large inverted repeat the one located in orf350 (OriB2) could be deleted. The other copy of OriB (OriB1) cannot be removed without mutating the essential *ycf1* gene and hence OriB1 dispensability cannot be addressed by a deletion that removes *ycf1* function (Mühlbauer et al. 2002). In a recent study, a stem-loop in OriB1 was mutated that left *ycf*1 intact (Scharff and Koop 2007). This allowed the isolation of plants in which OriB1 was mutated and OriB2 was deleted indicating that neither OriB sequence was essential. The copy number of plastid DNA appeared to be the same in shoot tips but lower in young and older leaves of deleted OriA lines (down ~1.5-fold) and lines lacking both OriA and OriB2 (down ~2-fold) compared to WT. The plastid DNA copy number in young and older leaves of OriB mutated lines (OriB2 deleted, OriB1 mutated) was higher (up ~1.7-fold) than WT plants (Scharff and Koop 2007).

4.2 Replication origins located in the single copy regions

Potential replication origins have also been located well away from the ribosomal genes. In *C. reinhardtii* two D-loops (*Ori*A and *Ori*B) spaced 7 kb apart (Waddell et al. 1984) map to the single copy region of plastid DNA (Fig. 6d). *C. reinhardtii* OriA was localised to a 224 bp region containing the *rpl16* gene (Chang and Wu 2000) whereas "OriB is located in or adjacent to *chl* L" (Wu personal communication). Replication at OriA is influenced by transcription across *rpl16* (Chang and Wu 2000). Novobiocin inhibition of D-loop replication led to novobiocin-resistant replication starting close to a 'hot spot' of recombination near the 3' end of the *psbA* gene in the large inverted repeat of *C. reinhardtii* plastid DNA (Woelfle et al. 1993). In *Z. mays*, preferential DNA synthesis from a 1368 bp plastid DNA sequence amongst cloned templates representing 94% of the *Z. mays* plastid genome by a partially purified *P. sativum* plastid DNA polymerase suggested this region promoted replication (Gold et al. 1987). The *Z. mays* region promoting DNA synthesis was found to promote bidirectional replication using a partially purified ~90 kDa *Z. mays* DNA polymerase. This was localised to a 455 bp sequence containing the 3' end of the *rpl16* gene in the large single copy region of plastid DNA (Fig. 6c; Carrillo and Bogorad 1988). A one kbp region including the 3' end of the *psbA* gene also promoted DNA synthesis using the *Z. mays* DNA polymerase fraction (Fig. 6c; Carrillo and Bogorad 1988). Whilst the sequence organizations of *Z. mays* and *C. reinhardtii* plastid DNA are not conserved it is interesting that a po-

tential replication origin overlaps with the *rpl16* genes in both species (Gold et al. 1987; Chang and Wu 2000).

Two replication bubbles were mapped to 1.8 kbp *Sac* I-*Bam* HI and 2.5 kbp *Bam* HI fragments in the large single copy region of *Glycine max* (soybean) chloroplast DNA by two-dimensional gel electrophoresis (Fig. 6b; Hedrick et al. 1993). The recently published *G. max* plastid genome (Saski et al. 2005) locates these bubble-containing fragments to regions containing the *rps12-clp* P1 and *pet* B-*petD* genes in the large single copy region. *Petunia hybrida* and *N. tabacum* plastid DNA sequences located in the small copy region near the location of *N. tabacum Ori*B promote autonomous replication in yeast nuclei (Ohtani et al. 1984; Dehaas et al. 1986). Although interesting, the finding that some A + T rich plastid sequences resembling yeast autonomously replicating elements (ARS) promote replication in *S. cerevisiae* nuclei would appear to provide weak evidence for locating plastid origins of replication. The locations of these ARS regions are not shown in Figure 6. The lack of agreement between different experimental approaches in locating origins of replication (Fig. 6) has hindered progress in this difficult area of research. It might indicate a lack of conservation in the location of plastid replication origins between different species, the presence of multiple origins, differences in the accuracy of locating origins or limitations of the plastid dual D-loop replication model.

5 Maintenance of small DNA molecules in plastids

Promotion of autonomous replication of plasmids within plastids themselves would provide a function-based assay in a homologous system for locating plastid origins of replication. Free monomeric plasmids are known to persist for a short period after transformation in *C. reinhardtii* (Boynton et al. 1988). This persistence might reflect replication and/or excision of complete plasmids integrated by single recombination events into the plastid genome. Homologous recombination between 16 bp direct repeats gave rise to an 868-bp DNA minicircle in *N. tabacum* plastids that was unstable but persisted as a multimeric series for several months during growth and development of transplastomic plants (Staub and Maliga 1994). The 868 bp excised sequence extends from the *trnI*(GAU) intron to the *trnA*(UGC) intron in the large inverted repeat and is located about 300 bases from the 82 bp region containing Ori A (Kunnimalaiyaan and Nielsen 1997a).

In *C. reinhardtii*, rescue of an *atpB* photosynthetic mutant with a partial function *atpB* allele led to amplification of the transforming plasmid in the form of large tandem arrays that appeared to be episomal (Suzuki et al. 1997). The maintenance of these episomal tandem arrays in plastids required a region of homology with the resident plastid genome. Revealingly, plastid transformation was not observed if there was no homology between resident plastid genome and transforming plasmid. Moreover, the inclusion of *C. reinhardtii* Ori A promoted integration rather than persistent autonomous replication (Suzuki et al. 1997). These results demonstrate that small plasmids containing mapped plastid replication origins do

not contain sufficient sequence information for replication and partition to exist as stable autonomous replicons in plastids. Recombination appears to play a role in plasmid maintenance in plastids and this may also be true of WT plastid genomes (see Section 7 below).

6 Deletion mapping delimits DNA sequences capable of self-replication in plastids

In cereals, most plastid genes are dispensable allowing the isolation of deletion mutants lacking most of the plastid genome (see Section 3.1 above). This natural deletion mapping identifies a region of plastid DNA located around the *trnE*(UUC) gene (Fig. 6c) that is self-replicating as linear DNA molecules (Day and Ellis 1985; Ellis and Day 1986; Harada et al. 1992; Kawata et al. 1997; Zubko and Day 2002; Cahoon et al. 2003). The retained region does not contain any known plastid replication origins mapped by electron microscopy, gel electrophoresis or biochemical methods (Fig. 6b, 6c). Maintenance of these linear DNA molecules might involve a replication origin that is not highly active in WT shoots and leaves. The relationship between the replication mechanisms maintaining small linear molecules and WT plastid DNA is not known but hairpin ends have been found at low frequency in WT *H. vulgare* plastids (Collin and Ellis 1990; Section 3.1).

Recombination events between sequences in the large single copy region and those located in either the large inverted repeat or small single copy region give rise to deleted plastid DNA molecules with circular maps (Day and Ellis 1984) and circular structures (Day and Ellis 1985) that have been found in albino *T. aestivum* plants from anther culture. The region present in the smallest circles (39 kbp, Day and Ellis 1985), containing only 30% of the plastid genome, stretches from the *trnE* gene region to the end of the adjacent large inverted repeat (Fig. 6c). These molecules contain only one large inverted repeat sequence and lack the replication origin mapped near *rpl16* in *Z. mays* (Fig. 6c; Gold et al. 1987). Deleted circular plastid DNA molecules containing only one large inverted repeat have also been found in WT *N. tabacum* chloroplasts by DNA fibre-based FISH (Lilly et al. 2001). These results show that sub-genomic plastid DNA molecules lacking the small single copy region, most of the large single copy region and one large inverted repeat can be maintained as circular DNA molecules in plastids.

7 A recombination-dependent DNA replication model of plastid DNA

The replicon model (Jacob et al. 1963) has been very successful and been substantiated in bacteria, animal viruses and budding yeasts such as *S. cerevisiae*. Problems in localising *bona fide* replication origins in plastids (Section 4.1, 4.2) and

the nuclear genomes of multicellular organisms (Gilbert 2004) have hindered universal application of the replicon model. The apparent failure of plastid genomes to conform to the replicon model with one or two well-defined replication origins might suggest the standard model (Fig. 5a, 5b) for replication of plastid genomes (Kolodner and Tewari 1975) requires revision (Bendich 2004). The lack of progress in understanding plastid genome replication has been matched by illuminating advances in bacterial genetics, particularly by Kogoma and colleagues (Asai et al. 1994), that have identified new replication mechanisms initiated by recombination for genome maintenance. Recombination-dependent DNA replication allows stalled replication forks at double-strand DNA breaks to re-establish and enables initiation of replication in the absence of a defined origin of replication (Asai et al. 1994; Kowalczykowski 2000). The mechanism requires a linear DNA end and strand invasion to prime DNA synthesis on a circular (Fig. 5c) or linear DNA (Fig. 5d) template followed by resolution of the recombination-junction to re-establish the replication fork. Replication forks resulting from D-loops primed by strand invasion were first described in bacteriophage T4 DNA replication (Kreuzer 2000, 2005). The role of recombination in maintaining replication forks has been suggested to be the main function of recombination systems (Cox et al. 2000; Goodman 2000). Plastids are known to contain a highly active homologous recombination pathway (see Section 8 below), which is a requirement for recombination-dependent DNA replication.

The only requirement for recombination-dependent DNA replication is a free end that could be located at any position on plastid DNA. As mentioned above (Section 3.1), the linear genomes in *S. cerevisiae* and plant mitochondria appear to have heterogeneous ends rather than a limited number of defined ends. Recombination-dependent DNA replication has been put forward as a mechanism for replication of fungal and plant mitochondrial genomes (Oldenburg and Bendich 1998; Williamson 2002), and the minicircles present in the plastids of dinoflagellates (Nelson and Green 2005). Linear ends with 3' overhangs would allow strand invasion to prime DNA replication on other plastid DNA molecules acting as templates. Induction of double-strand breaks at specific sites in *E. coli* gives rise to recombination-dependent DNA replication origins that can be mapped (Asai and Kogoma 1994). Linear DNA molecules with defined ends have been found in *Z. mays* (Oldenburg and Bendich 2004b) and *N. tabacum* (Scharff and Koop 2006, 2007) plastids (Section 3.2 above). One common end in both species maps close to OriA in the large inverted repeat of *N. tabacum* between the 16S and 23S rRNA genes (Fig. 6b). An origin of replication has not yet been located in this position in cereal plastid genomes (Fig. 6c). These natural ends have been suggested to invade template DNA and prime DNA synthesis by recombination-dependent DNA replication (Bendich 2004; Oldenburg and Bendich 2004b). As a result they could define sites at which replication of the plastid genome is initiated. Interestingly, the locations of these ends appeared to change when OriA and OriB2 were deleted (Scharff and Koop 2007). Recombination-based-replication will lead to circular DNA molecules with tails (Fig. 5c) and linear branched structures (Fig. 5d). Linear DNA molecules are extended when their ends invade and replicate template genomes. The process leads to multimerization of linear DNA molecules and has

been documented in detail during replication of bacteriophage T4 in *E. coli* (Kreuzer 2000). Highly branched complexes will be formed if several independent DNA molecules are connected by recombination-dependent DNA replication events. The complexity of branching increases if resolution (see resolution step shown in Fig. 5c, 5d) is not completed in some of the branches (Kreuzer 2000). Complex branched networks arising from recombination-dependent DNA replication might explain the 90-95% (by mass) of *Z. mays* plastid DNA found in high MW complexes (Oldenburg and Bendich 2004b; Section 3; Fig. 3c).

8 DNA recombination in plastids

Plastid fusion and DNA recombination between different plastid types are rare in flowering plants. Rapid segregation is observed when two plastid types with different genomes are forced into the same cell by protoplast fusion (Morgan and Maliga 1987). In *C. reinhardtii*, recombination between parental plastid genomes in exceptional zygotes is well established (Gillham 1974). The development of plastid transformation has demonstrated an active homologous DNA recombination pathway in *C. reinhardtii* (Boynton et al. 1988) and flowering plant plastids (Svab et al. 1990). The rarity of plastid fusion in angiosperms probably explains the lack of DNA recombination between ”parent” plastid genomes in protoplast fusion experiments. In one successful protoplast fusion experiment a single plant with a recombinant plastid genome resulting from at least six crossover events between parental genomes was isolated (Medgyesy et al. 1985).

Most characterised plastid genomes contain a large inverted repeat sequence. Recombination between the large inverted repeat sequences (flip-flop recombination) is responsible for the two isomers of plastid DNA, which differ with respect to the orientation of the single copy regions (Palmer 1983). Flip-flop recombination giving rise to the two isomers can take place between circular (Fig. 7a) or linear DNA substrates (Fig. 7b). The head-to-head circular dimers (Fig. 3a) in *L. sativa* and *S. oleracea* plastids observed by Kolodner and Tewari (1979) were explained by intermolecular recombination between opposite large inverted repeats in circular DNA substrates (Fig. 8a). These head-to-head dimers are comprised of an inverted sequence representing ~90% of the unit genome size separated by small spacer loops comprised of the small single copy sequences. *P. sativum* plastid DNA lacks a large inverted repeat providing an explanation for the lack of head-to-head dimers in plastids from this species (Kolodner and Tewari 1979). Head-to-head inverted sequences representing ~90% of the unit genome length will also be produced by recombination events between large inverted repeat sequences involving linear DNA substrate (Fig. 8b). Homologous recombination that is not limited to specific sequences appears to be responsible for generating these isomers. Intermolecular recombination between inverted repeats in long chain multimers of plastid DNA would be expected to place at any point and would give rise to a large number of isomers. Intramolecular recombination events between tandemly repeated copies of the unit genome in linear multimers will give

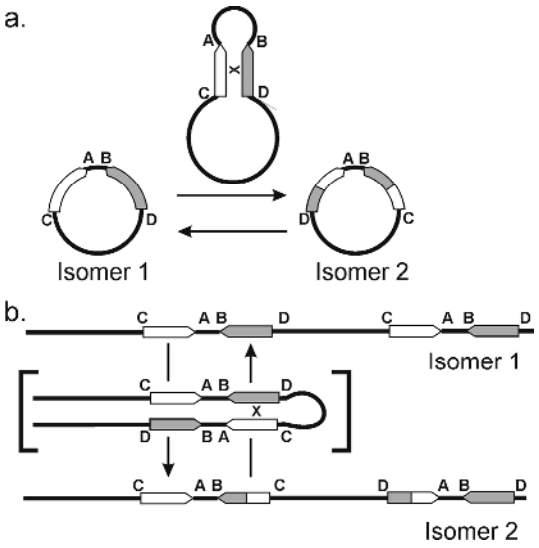

Fig. 7. Intramolecular flip-flop recombination between large inverted repeat sequences in a) Circular, and b) Linear DNA molecules. The large inverted repeat sequences are shown as converging grey and white box arrows. Note that the linear product in b contains one of the head-to-head inverted sequences high-lighted in Fig. 8b.

rise to circular DNA molecules. Oldenburg and Bendich (2004b) have pointed out that recombination-dependent DNA replication primed by an end within the large inverted repeat will also result in head-to-head inverted sequences and flipping of single copy regions. This can be visualised by looking at Figure 8b where the products of reciprocal recombination can also be obtained by strand-invasion by the top molecule on the bottom template at the crossover site followed by D-loop replication, resolution, and replication fork movement (see Fig. 5d) to the end of the template molecule.

Figure 9 shows a recombination event between large inverted repeats following replication of one copy of the repeat (Futcher 1986). This switches the direction of the replication fork allowing many identical head-to-tail copies of the unit genome to be made in a multimeric circle from a single template without re-initiation of DNA replication. The absence of a large inverted repeat in some plastid genomes (Palmer and Thompson 1982) would suggest that this double rolling circle mechanism is not essential for amplification of plastid DNA. A linear multimeric chain replicated from a circular template will be formed if the lagging strand is not replicated following recombination between the duplicated and unreplicated copies of the large inverted repeats (Ellis and Day 1985).

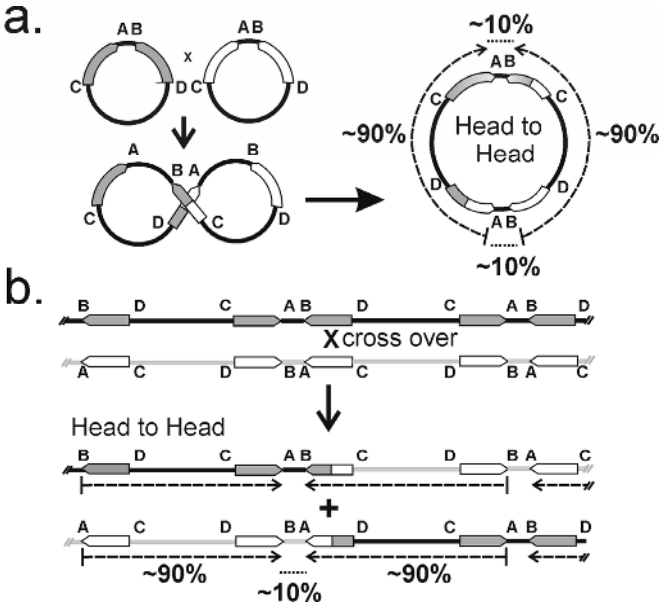

Fig. 8. Intermolecular recombination between opposite large inverted repeat sequences gives rise to plastid genomes orientated head-to-head. a) Circular head-to-head and b) Linear head-to-head DNA molecules. Head-to-head inverted sequences are shown as dotted arrowed lines. Length is expressed as a percentage of the unit genome size.

8.1 Integration of foreign genes by homologous recombination

The era of plastid transformation enables the detailed study of recombination events in plastids. In algae and plants foreign DNA integrates by homologous recombination between common DNA sequences in the transforming vector and resident plastid genome. Large regions of donor plastid DNA integrate into the resident plastid genome well beyond the markers used to select transformants resulting in incorporation of all or almost all of a 6.2 kb recombinant plastid sequence in *N. tabacum* plastid transformants (Staub and Maliga 1992). Reciprocal recombination or gene conversion events between transforming plasmid and resident plastid genome will result in transgene integration (Fig. 10a). A variety of plastid DNA sequences have been used to target integration of foreign genes to different sites in the plastid genome (Chapter 14). This indicates the plastid recombination machinery is not limited to specific substrates but can act on a wide selection of DNA sequences. When a vector containing a gene-of-interest flanked by targeting DNA is introduced into angiosperm plastids, double recombination events in both arms will insert the transgene into the plastid genome. Alternatively, the entire plasmid can integrate as a result of homologous recombination in

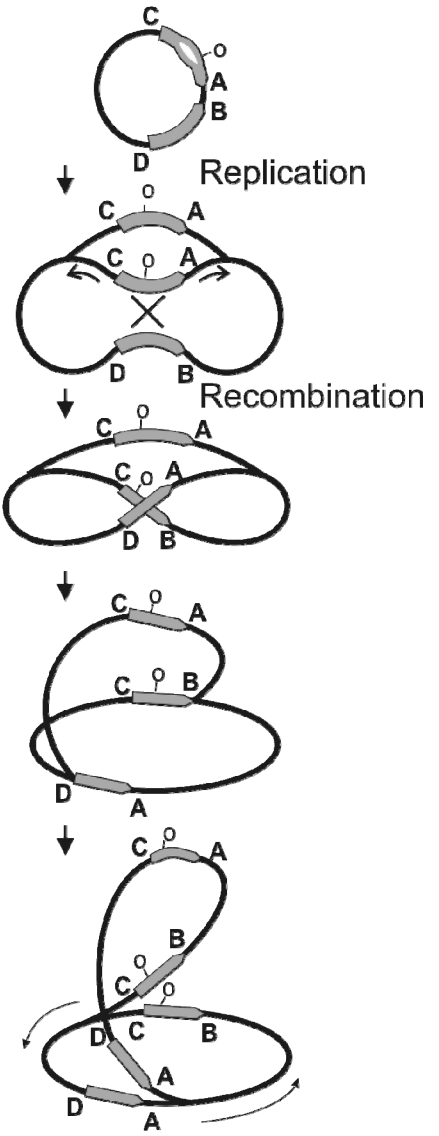

Fig. 9. Double rolling circle model of DNA replication gives rise to large circular multimers from a single round of replication initiation (Futcher 1986). Not to scale: the small single copy region is enlarged to illustrate the model. O= replication origin.

one flanking region of plastid DNA. The resulting co-integrate contains duplications of left and right targeting regions. The co-integrate can be selected by placing the marker gene in the vector sequences (Klaus et al. 2004). When selection is removed further homologous recombination events between these duplicated sequences excise vector and marker sequences to leave either the gene-of-interest or

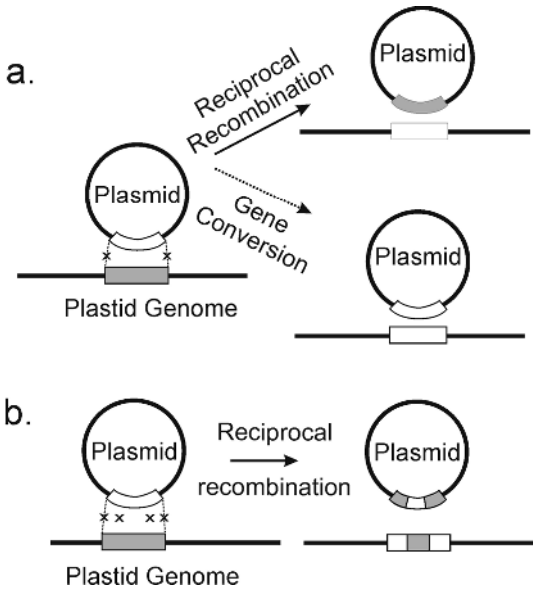

Fig. 10. Integration of foreign DNA into the plastid genome by homologous recombination. a) Integration requires homology and involves reciprocal recombination and possibly gene conversion events. b) Integration of homeologous DNA reveals multiple crossover events between target and donor DNA (Kavanagh et al. 1999).

WT plastid sequences (see Chapter 14; Klaus et al. 2004). These co-integrate experiments show that the crossover events in both arms required for integration are not tightly linked and this integration pathway must reflect properties of the DNA-RRR pathways in angiosperm plastids.

In *E. coli,* homologous recombination is stimulated by 8 base 5'GCTGGTGG chi sequences which are recognised by the *recBCD* complex (Kowalczykowski 2000). The eight base chi motif is absent in *N. tabacum* plastids (Shinozaki et al. 1986) but is present in the 16S ribosomal RNA genes in grass plastid genomes (Hiratsuka et al. 1989). Any role for chi sequences in plastids would appear to be ruled out by the finding that cyanobacterial and angiosperm genomes do not appear to contain homologues of genes encoding the *recBCD* complex. In *C. reinhardtii,* sequences that appear to stimulate recombination have been identified within the large inverted repeat and have been localised to a 400 bp region of plastid DNA containing the 3' end of the *psbA* gene (Newman et al. 1992).

8.2 Homologous recombination between short DNA repeats

The introduction of short repeated sequences into plastid genomes by transformation has demonstrated that they are effective substrates for homologous recombination (Table 1). Recombination between direct repeats excises the intervening

Table 1. Recombination events between engineered DNA repeats in transgenic plastids

Species	Repeat length (bp)	Repeated sequence	Reference
Nicotiana tabacum	174[1]	16S rrn promoter & *rblC* RBS	Iamtham and Day 2000
	418[1]	*psbA* 3' UTR	Iamtham and Day 2000
	210	*rbcL* 3' UTR	Dufourmantel et al. 2007
	232 (IR)	*psbA* 3' UTR	Rogalski et al. 2006
	649	atpB 5'UTR & promoter	Kode et al. 2006
Glycine max	403	HPPD gene	Dufourmantel et al. 2007
Lactuca sativa	~200	16S rrn promoter & RBS	Lelivelt et al. 2005
Chlamydomonas reinhardtii	216	*chlL* coding region	Cerutti et al. 1995
	483	pACYC184 NruI-BspHI[2]	Fischer et al. 1996

Repeats were in direct orientation apart from Rogalski et al. 2006. Most recombinant genomes contained two engineered repeats apart from [1]three 418 bp and two 174 bp repeats (Iamtham and Day 2000). [2]Restriction fragment from pACYC184 plasmid. IR = inverted repeat. RBS = ribosome binding site. HPPD = 4-hydroxyphenylpyruvate dioxygenase.

DNA (Fig. 11a) while recombination between inverted repeats reverses the orientation of the intervening DNA. Both length and number of repeated sequences influence recombination frequency. Whereas two 418 bp direct repeats were ineffective in deleting intervening DNA, three 418 bp repeats promoted high frequencies of excision to leave a single 418 bp direct repeat (Iamtham and Day 2000). Intermediate forms containing two copies of the 418 bp repeat were not detected indicating that once activated the homologous recombination pathway goes to completion (Fig. 11b). A variety of direct repeats promote excision (Table 1) but a systematic study on the relationship between repeat length, DNA sequence and recombination frequency has not been carried out. Whereas recombination was barely detected between two 3'UTR *psb* A direct repeats of 418 bp (Iamtham and Day 2000), excision was promoted by repeated ~200 bp 16S *rrn* promoter elements in *L. sativa* (Lelivelt et al. 2005), and *N. tabacum* (Zou et al. 2003), and 210 bp *rbc* L 3' UTR repeats in *N. tabacum* (Dufourmantel et al. 2007). Recombination between 232 bp inverted *psbA* 3' UTR repeats has been shown in *N. tabacum* plastids (Rogalski et al. 2006). Excision can take place at any time during the transformation process and non-deleterious transgenes flanked by direct repeats, might be stabilised by the high copy number of plastid DNA once homoplasmy has been reached. Promoter regions of ~120 bp are duplicated in a number of plastid transformation vectors with no apparent reports of instability (Zoubenko et al. 1994) hinting at a lower size limit for efficient recombination.

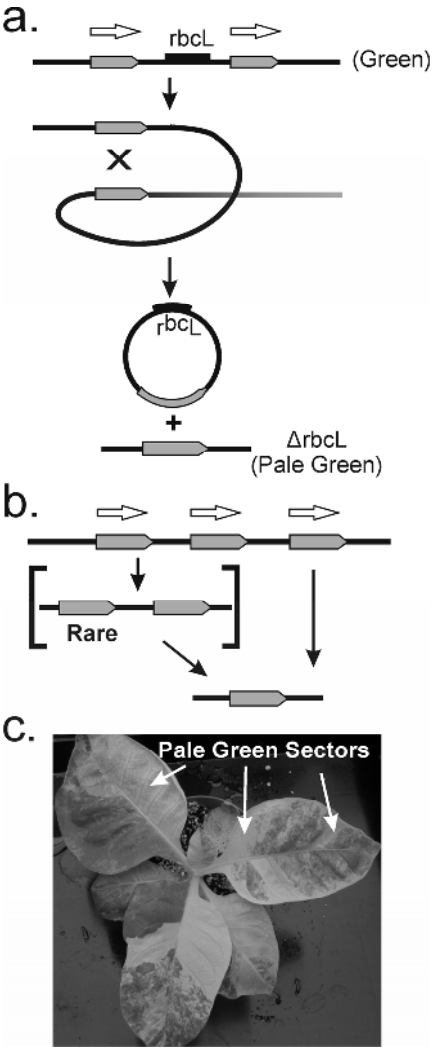

Fig. 11. Homologous recombination between engineered direct repeats (arrows) in plastid DNA. a) Recombination between two 649 base *atp*B promoter repeats (Kode et al. 2006) excises the *rbcL* gene and a marker gene (not shown). b) Recombination between three 418 bp 3' *psb*A repeats leaves a single 418 bp 3' *psbA* sequence. Once recombination is activated the pathway appears to go to completion because intermediates containing two repeats are rare and were not found (Iamtham and Day 2000). c) The recombination event shown in a) gives rise to pale green sectors (light areas) which appear randomly throughout leaf development.

The types of intervening genes excised by flanking direct repeats might also influence the accumulation of recombination products. Selection on plastid fitness might be expected to promote the division of plastids from which foreign genes

with a negative impact on plastid functions and division have been deleted. The finding that excision mediated by recombination between duplicated 649 bp *atpB* 5' regulatory regions allows the isolation of mutant defective plastid genomes lacking the *rbcL* gene (Fig. 11a; Kode et al. 2006) indicates that products of deleterious recombination events can be isolated under suitable conditions; in this case sucrose was provided in the media to allow non-photosynthetic growth. Recombination-mediated excision of *rbcL* and segregation of plastid genomes gives rise to pale-green sectors allowing the process to be monitored. Pale-green sectors of varying sizes are visualised in leaves (Fig. 11c, lighter areas represent pale-green sectors) indicating the recombination and segregation pathways are active throughout leaf development. The variable sizes and random appearance of pale-green sectors reflect random spontaneous excision events combined with stochastic replication and segregation of plastid genomes.

In *E. coli*, a minimum sequence length of 23-27 bp is considered to be required for efficient homologous recombination via the recBC-dependent pathway (Shen and Huang 1986). A minimum identical stretch of 150 to 200 bases appears to be required for homologous recombination in *C. reinhardtii* (Newman et al. 1992). Recombination between engineered direct repeats of 216 bp (Cerutti et al. 1995) has been observed in *C. reinhardtii* (Cerutti et al. 1995). In another study recombination was not detected between 100 bp or 230 bp direct repeats but frequent recombination was detected between 483 bp repeats (Fischer et al. 1996). This is probably because the recombination assay for the 216 bp repeats relied on restoration of gene function giving rise to green sectors and was more sensitive than the loss of antibiotic resistance assay involving the 230 bp direct repeats flanking *aad*A (Fischer et al. 1996). However, sequence-dependent differences in recombination rates between direct repeats cannot be ruled out (Fischer et al. 1996).

Site-specific recombination is an alternative to homologous recombination for manipulating plastid genomes. The Cre site-specific recombinase from the P1 bacteriophage of *E. coli* mediates strand-exchange between 34 bp LoxP sites (Sternberg and Hamilton 1981). When Cre is introduced into plastids it recombines LoxP sites as intended (Corneille et al. 2001; Hajdukiewicz et al. 2001). Unexpectedly, Cre recombinase also appears to stimulate recombination between ~120 bp direct repeats comprised of the 16S rrn promoter region (Corneille et al. 2001; Hajdukiewicz et al. 2001). The WT *N. tabacum* plastid genome lacking LoxP sites is stable in the presence of Cre recombinase (Corneille et al. 2003). This suggests that creation of double strand breaks at LoxP sites by Cre recombinase stimulates native recombination events in plastids. Cre-stimulated illegitimate recombination events between a LoxP site and a recombination hotspot in the promoter region of the rps7/3'rps12 operon were also reported (Hajdukiewicz et al. 2001; Corneille et al. 2003). The recombination hotspot contained multiple copies of a TATTA sequence (Hajdukiewicz et al. 2001). Short repeats are often associated with recombination hotspots in plastid DNA. The role of short multiple 18 to 37 bp repeats near a recombination hot spot in *C. reinhardtii* was addressed by deleting them. Their deletion did not reduce recombination frequency (Newman et al. 1992) indicating the repeats were not responsible for increased recombination. The observation that Cre-cleaved DNA ends are recombinogenic

might suggest that the natural ends of the linear DNA molecules found in plastids (Oldenburg and Bendich 2004b; Scharff and Koop 2006) are protected by proteins or secondary structures, for example, DNA loops or hairpins.

Spontaneous excision of an 868 bp sequence following apparent recombination between 16 bp imperfect direct repeats (5' GTACTGc/tGCTCTCCAA) was reported to accompany plastid transformation in *N. tabacum* (see Section 5 above; Staub and Maliga 1994). This might indicate that some plastid sequences of less that 20 bp are effective substrates for recombination. Evolutionary comparisons of plastid genomes have found DNA sequence inversions in the large single copy region that distinguish related species of flowering plants (Doyle et al. 1992). Analyses of the end points of an inversion in rice plastid DNA relative to *N. tabacum* suggest recombination events between sequences as short as 16 bp in length (Hiratsuka et al. 1989). In another study, short inverted repeats of 7-11 bp were found to be associated with inversions of the intervening 4 bp region (Kelchner and Wendel 1996). These studies indicate recombination events between short repeated stretches of nucleotides. Whether the same recombination pathway acts on the very short (10-20 bp) and longer (~200 bp and above) substrates remains to be determined. These questions can only be addressed once mutants in specific plastid recombination pathways have been isolated.

9 Recombination and plastid genome stability

An active homologous recombination pathway in plastids must underlie maintenance of the plastid genome and shape its evolution. Plastid transformation experiments have demonstrated recombination events between short repeats of ~200 bp in length in flowering plant plastids resulting in the deletion of genes (Table 1). Such excision events would be deleterious to WT plastid genomes. Grossly deleted *N. tabacum* transgenic plastid genomes resulting from recombination between distant *psbA* 3'UTR sequences do not persist (Svab and Maliga 1993) presumably because of strong selection against dysfunctional plastids with defective plastid genomes. Deleterious recombination events in WT plastid genomes would be avoided if they lacked DNA repeats or contained small repeats that were below the size needed for homologous recombination. One set of inverted repeats would be allowed because they would only flip the relative orientation of the single copy regions (Fig. 7). Any repeats that lie on either side of the large inverted repeat would be converted into direct repeats by large inverted repeat mediated flipping of single copy regions and destabilise plastid genomes (Day and Ellis 1984). Loss of the large inverted repeat in *P. sativum* is associated with rearrangements in plastid gene order (Palmer and Thompson 1982); presumably because inversions are not restricted by the presence of a large inverted repeat (Day and Ellis 1984).

Whilst most angiosperm plastid genomes contain a large inverted repeat of 20-76 kbp they are deficient in repeated DNA sequences over 100 bp in size. The plastid genome of *Pelargonium* x *hortorum* contains a large number of repeated sequences, including nine pairs of dispersed repeats of 31-101 bp in size

(Chumley et al. 2006). Over 20% of the *C. reinhardtii* plastid genome is comprised of repetitive DNA of less than 50 bp in length (Maul et al. 2002). Presumably these repeated sequences of 100 bp or less are too small to act as efficient substrates for homologous recombination. Plastid DNA is uniform within a plant and recombination events that would destabilize and fragment plastid genomes are not normally observed. However, low frequency recombination events between short repeats might give rise to the plastid DNA rearrangements observed during evolution (Hiratsuka et al. 1989; Chumley et al. 2006). Deleted plastid genomes with circular maps (Day and Ellis 1984; Cuzzoni et al. 1995) and circular topologies (Day and Ellis 1985) have been described in albino cereal plants from anther culture demonstrating the instability of the plastid genome when cells are rescued by heterotrophic growth *in vitro*. A recent report suggests albino *Bambusa edulis* (bamboo) plants contain deleted plastid genomes (Liu et al. 2007). Aberrant subgenomic circles have also been described in WT *N. tabacum* chloroplasts by DNA fibre-based FISH analysis (Lilly et al. 2000). Two factors appear to be required to maintain plastid genomes, first, DNA-RRR surveillance mechanisms that either reduce the frequency of deleterious recombination events or repair deleted plastid genomes, and second, selection for functional plastids with an intact plastid genome.

10 Homeologous recombination in plastids

In *E. coli*, sequence divergence dramatically reduces the rate of homologous recombination. A 10% reduction in identity between DNA sequences reduces homologous recombination frequency by 40-fold (Shen and Huang 1986). The effect of mismatches on integration of transgenes can be studied by using heterologous plastid DNA to target integration of foreign genes into plastids. *N. tabacum* plastid DNA has been used to target integration of foreign genes in *Lycopersicon esculentum* (tomato; Ruf et al. 2001), *Petunia hybrida* (Zubko et al. 2004) and *Solanum tuberosum* (potato; Sidorov et al. 1999). Because the transformation frequency of homeologous *N. tabacum* plastid DNA was not compared with homologous plastid DNA the influence of mismatch on plastid transformation frequency in these species is not known. In a more detailed study, a 7.8 kbp region of *Solanum nigrum* plastid DNA was introduced into *N. tabacum* plastids (Kavanagh et al. 1999). Recombinant plastid genomes exhibited a mosaic structure comprised of several patches of *S. nigrum* DNA interspersed with *N. tabacum* plastid DNA sequences (Fig. 10b). This is consistent with multiple recombination events during integration of 7.8 kbp *S. nigrum* plastid DNA and random resolution of Holliday junctions. Although *S. nigrum* and *N. tabacum* plastid DNA showed 2.4% sequence divergence, plastid transformation frequencies were not reduced relative to using homologous sequences for *N. tabacum* plastid transformation. Stringent mismatch repair processes which reduce recombination between diverged DNA sequences (Evans and Alani 2000) might be suppressed in flowering plant plastids (Kavanagh et al. 1999). In contrast, homeologous plastid DNA exhibiting around

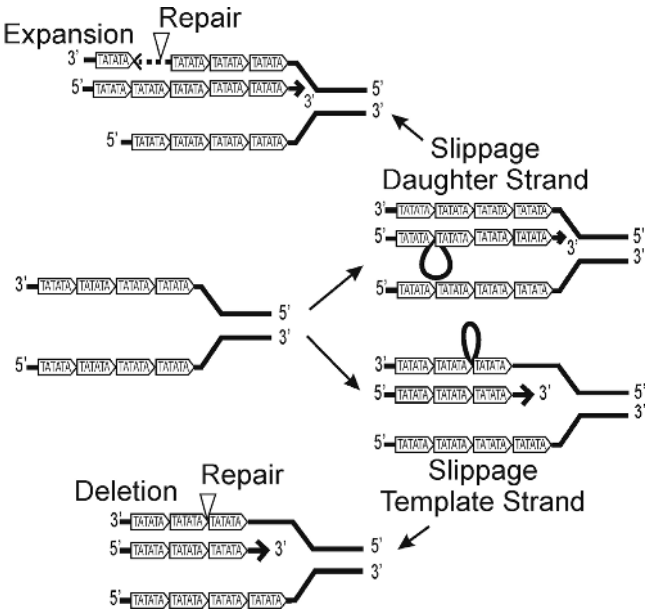

Fig. 12. Expansion and deletion of short direct repeats by replication slippage (Lovett 2004). Slippage in the daughter strand increases the number of repeat units while slippage in the template strand reduces the number of repeat units.

2-4% mismatch decreased transformation frequency by two to fivefold (Newman et al. 1990) in *C. reinhardtii*. This might indicate a more efficient plastid mismatch repair system in this green alga.

11 Replication slippage in plastids

Very short tandem repeats, based on mononucleotide, dinucleotide, trinucleotide, and consecutive nucleotide repeats up to the ~30-mer, are found in plastid genomes. Very short tandem repeats are considered to result from slippage of the replication fork. Replication slippage in the newly replicated daughter strand inserts a repeat whilst slippage of the template strand deletes a repeat (Fig. 12). A number of hot spots of variation in plastid genomes are associated with short tandem repeats (Newman et al. 1992; Sears et al. 1996; Stoike and Sears 1998; Ogihara et al. 2002). Plastome-mutator is a nuclear mutation in *O. hookeri* associated with a 200 to 1000-fold increase in pigment-deficient sectors and changes in plastid DNA (Epp 1973; Stoike and Sears 1998). The product of the plastome-mutator gene is not known but it has been suggested to be involved in plastid DNA-RRR pathways (Stoike and Sears 1998). Examination of alterations induced by plastome-mutator in the intergenic region between the 16S and 23S ribosomal RNA

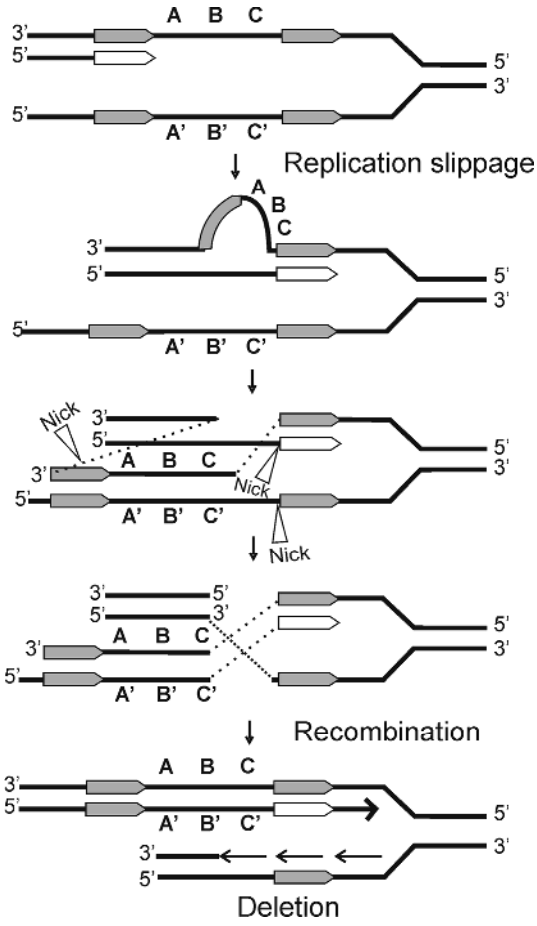

Fig. 13. Direct repeat mediated deletion by replication slippage and recombination (Bi and Liu 1996).

genes suggested they were the result of replication slippage rather than recombination events (Stoike and Sears 1998).

Replication slippage between direct repeats and recombination (Fig. 13; Bi and Liu 1996) provides an alternative mechanism to homologous recombination (Fig. 11a) for excision of genes from plastid DNA. Replication slippage induced recombination is reduced with increasing distance between repeats and does not increase in frequency when repeat length is increased above ~100 bp in *E. coli* (Bi and Liu 1996). The observation that recombination between engineered repeats in plastids increases with longer repeats favours a mechanism involving homologous recombination rather than slippage (see Section 8.2 above). Replication slippage can give rise to insertions and deletions (Fig. 12) and appears to be a major force in plastid genome evolution. The 70 kb plastid genome of the non-photosynthetic parasite *Epifagus virginiana* is less than half the size of the 156 kb *N. tabacum*

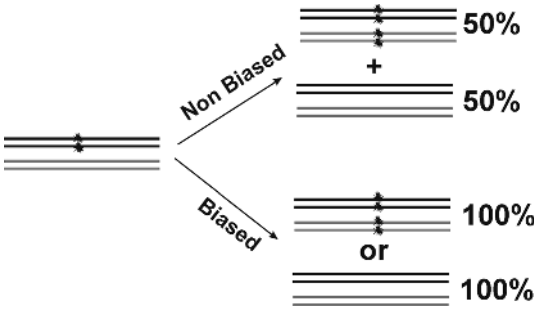

Fig. 14. Recombination events between WT and mutant (*) DNA duplexes can be non-biased or biased. Biased gene conversion events will favour one of the alleles and either fix or remove mutations.

plastid genome (Wolfe et al. 1992). Gene order in *E. virginiana* and *N. tabacum* is conserved despite the loss of all photosynthesis-related genes. A large number of deletion events on a local scale mediated by replication slippage rather than recombination events between distant parts of the plastid genome would appear to be the major mechanism underlying selective gene loss and genome reduction in *E. virginiana* (Wolfe et al. 1992).

12 DNA repair in plastids

DNA replication, recombination, and repair are interrelated processes and the homologous recombination pathway in plastids is likely to play a role in DNA repair (Cerutti et al. 1995). The estimated mutation rate of plastid genes is approximately two-fold lower than that of nuclear genes (Wolfe et al. 1987). Within plastids the synonymous substitution rate of genes located in the large inverted repeat is about two-fold lower than that for genes located in the single copy regions (Perry and Wolfe 2002). This has been interpreted to be the result of the two-fold higher dosage of inverted repeat sequences and biased gene conversion in favour of WT plastid DNA sequences (Birky and Walsh 1992; Perry and Wolfe 2002). Non-biased repair will either correct the mutation to WT or fix the mutation (convert WT to mutant) and give rise to both outcomes in equal proportions. Biased repair favours one of these outcomes to give 100% of only one product (Fig. 14). Direct experimental confirmation for biased gene conversion has been obtained by monitoring correction of mutations tightly linked (31 bp distance) to an *aad*A insertion in transgenic *N. tabacum* plastids (Khakhlova and Bock 2006). Whilst the *aad*A gene was retained by spectinomycin selection the mutations were repaired to WT with a bias for reversing AT to GC changes more efficiently than GC to AT mutations. It has been suggested that this bias towards AT might underlie the high overall AT content (>70% AT) of plastid genomes (Khakhlova and Bock 2006). Multiple copies of plastid DNA and biased gene conversion in favour of WT would reduce the rate at which mutations are fixed.

Alternatives to RecA-based recombination repair include photoreactivation, base excision repair, nucleotide excision repair, and mismatch repair (Kimura and Sakaguchi 2006). Little is known on these alternative repair pathways in plastids. A putative plastid-localised uracil-DNA glycosylase activity probably involved in base excision repair was partially purified from *Z. mays* chloroplasts (Bensen and Warner 1987). UV-induced lesions in the plastid *psbA* gene of *G. max* suspension culture cells were repaired in the light (but not in the dark) with kinetics that were considerably slower than expected for photoreactivation by photolyases (Cannon et al. 1995). Experiments on purified *S. oleracea* chloroplasts (Hada et al. 2000) and the lack of identification of plastid transit peptides in the products of plant genes encoding photolyases (Draper and Hays 2000) led to the possibility that plastids might be deficient in photolyase-mediated photoreactivation. However, tolerance of plastid DNA replication to UV-B lesions in *A. thaliana* plants grown in blue (photorepair-compatible) light might suggest the presence of as yet unidentified photolyases in plastids (Draper and Hays 2000). Under gold light where light-dependent photorepair does not take place, a UV-B fluence rate of 5 kJ m-2 inhibits replication of plastid DNA but not nuclear and mitochondrial DNA indicating a deficiency in light-independent (dark) repair pathways in chloroplasts (Draper and Hays 2000).

Endonuclease activities that could act on apurinic sites following base removal were purified from *H. vulgare* chloroplasts (Velemínskíý et al. 1980). A single-strand specific nuclease activity from *T. aestivum* chloroplasts cleaves single stranded DNA or RNA regions including 5' flaps, 5' overhangs, and 3' pseudoflaps and has been suggested to be involved in DNA repair (Przykorska et al. 2004). The multi-subunit replication protein A (RPA) binds to single-stranded DNA and is involved in pathways including nucleotide excision repair. RPA subunits appear to be targeted to plastids (Kimura and Sakaguchi 2006). Plastid-localised homologues of RecQ have been implicated in DNA repair (see Section13.5 below). Nitroso-methy-urea and nitroso-guanidine are particularly effective for inducing mutations in flowering plant plastid genomes (Hagemann 1976). Methyl transferases reverse the damage to bases caused by these alkylating agents. The presence of methyl transferases in nuclei but their absence in plastids might explain the utility of these mutagens for inducing plastid mutations (Sears 1998).

13 Identification of proteins involved in plastid DNA RRR-pathways

Studies on bacteria have identified a suite of DNA metabolism enzymes including DNA polymerases, DNA primase, RecA, topoisomerases, and helicases (CameriniOtero and Hsieh 1995). The cyanobacterium *Anabaena* contains 93 genes encoding proteins with significant similarity to known DNA-RRR proteins (Kaneko et al. 2001). Early studies to identify plastid DNA-RRR proteins involved purifying enzymes with DNA-RRR activities, such as DNA synthesis, from chloroplasts (McKown and Tewari 1984). More recently, whole genome da-

tabases have been used to identify candidate plastid-targeted proteins with significant matches to well-known bacterial DNA-RRR proteins. Further experimental support is then required to confirm *in silico* predictions of plastid location. The list of genes encoding homologues of DNA-RRR proteins for which there is experimental support for a plastid location is short and is reviewed below. Proteomics of purified chloroplasts (Chapter 12) provides an alternative approach to identify DNA-RRR proteins. However, the limited abundance of plastid DNA-RRR proteins hinders their identification in whole chloroplast preparations. Further purification of sub-chloroplast fractions containing DNA-protein complexes is required to identify plastid DNA-RRR proteins (Sakai et al. 1999; Phinney and Thelen 2005).

13.1 Plastid DNA polymerases

Eukaryotic cellular template-dependent DNA polymerases can be classified into α, β, γ, δ, and ε DNA polymerases based largely on work in vertebrates and yeast (Wang 1991). The α, β, δ, and ε DNA polymerases are located in nuclei whereas the γ–DNA polymerase is located in mitochondria. This original list of five eukaryotic DNA polymerases has expanded to fourteen (Mori et al. 2005) as new DNA polymerases are identified including those involved in translesion DNA synthesis (Hubscher et al. 2000). Nine classes of DNA polymerase have been identified by analyses of plant genomes (Mori et al. 2005). The activity of γ–DNA polymerases is resistant to aphidicolin, sensitive to N-ethylmaleimide, and relatively resistant to low concentrations of dideoxynucleoside 5' triphosphates (Wang 1991). This γ–DNA polymerase activity can be distinguished from α, δ, and ε DNA polymerases which are sensitive to aphidicolin but resistant to dideoxynucleoside 5' triphosphates, and β–DNA polymerase, which is resistant to both aphidocolin and N-ethylmaleimide but sensitive to low concentrations of dideoxynucleoside 5' triphosphates (Wang 1991). Plastid DNA polymerases that have been characterised from *S. oleracea* (Spencer and Whitfeld 1969; Sala et al. 1980), *P. sativum* (McKown and Tewari 1984), *G. max* (Heinhorst et al. 1990; Bailey et al. 1995), *N. tabacum* (Sakai et al. 1999), and *C. reinhardtii* (Wang et al. 1991) resemble γ-DNA polymerases based on their resistance to aphidocolin and sensitivity to N-ethylmaleimide. Furthermore, the plant but not the *C. reinhardtii* plastid DNA polymerases are resistant to low concentrations of dideoxynucleoside 5' triphosphates (Heinhorst and Cannon 1993). The activities of purified plastid DNA polymerases appear to be stimulated by KCl (Spencer and Whitfeld 1969; Sala et al. 1980; McKown and Tewari 1984; Heinhorst et al. 1990; Sakai et al. 1999) and inhibited by ethidium bromide (McKown and Tewari 1984; Wang et al. 1991; Sakai et al. 1999).

Estimated sizes of plastid DNA polymerases were 87 kDa (McKown and Tewari 1984) and 70 kDa for *P. sativum* (Gaikwad et al. 2002), 85-90 kDa for *G. max* (Heinhorst et al. 1990), 116 kDa for *N. tabacum* (Sakai et al. 1999), and

Table 2. Plastid DNA-RRR proteins encoded by *Arabidopsis thaliana* nuclear genes.

Protein	Length (aa)	Acc. no.	Gene no.	Reference
DNA polymerase	1049[1]	AAL58915	At1g50840	Christensen et al. 2005
	1034[1]	BAE98907	At3g20540	
RecA	439	Q39199	At1g79050	Cao et al. 1997
Gyrase A	950[1]	AAG51377	At3g10690	Wall et al. 2004
Gyrase B	732[1]	Q94BZ7	At5g04130	Wall et al. 2004;
	657[1]	Q9SS38	At3g10270	Christensen et al. 2005
RecQ	858[2]	Q9FT69	At5g27680	Saotome et al. 2006;
	606[2,3]	Q9FT74	At3g05740;	Hartung and Puchta 2006

Dual targeting or alternative translation start sites target product to both plastids and mitochondria. Length will vary depending on translation initiation at alternative start sites or at non-AUG start codons in the 5' UTR (Christensen et al. 2005). 2. Homologues of *Oryza sativa* (rice) plastid-localised proteins (Saotome et al. 2006). 3. Targeted to plastids and nuclei in *O. sativa.*

possibly two catalytic subunits of 80 and 116 kDa for *C. reinhardtii* (Wang et al. 1991). The size discrepancies might be explained by proteolytic cleavage of proteins during purification. In the absence of primary sequence information the relationships between these proteins are not known and it is unclear whether they are different DNA polymerases or homologues of the same protein. Analyses of DNA polymerase genes present in sequenced genomes (discussed below; Mori et al. 2005) are likely to help resolve some of the discrepancies encountered in the earlier biochemical work. A number of proteins have been found to be associated with purified plastid DNA polymerases. These include a 43 kDa protein (related to ribonuclease T2, GenBank Acc. P93845) that stimulates the activity and processivity of an 87 kDa (Chen et al. 1996) and a 70 kDa *P. sativum* DNA polymerase (Gaikwad et al. 2002), and a 3' to 5' 20 kDa exonuclease subunit of a 105 kDa *S. oleracea* DNA polymerase complex (Keim and Mosbaugh 1991).

There are similarities in the sizes and properties of DNA polymerases purified from plant mitochondria and chloroplasts (Heinhorst et al. 1990; Sakai et al. 1999). The idea of related DNA polymerases in mitochondria and plastids is supported by analyses of genes in sequenced genomes. The *A. thaliana* genome encodes at least two genes encoding organelle DNA polymerases (Table 2) sharing 70% amino acid identity that are expressed in the shoot apical meristem (Mori et al. 2005). The 116 kDa 1034-long *A. thaliana* DNA polymerase-like protein (gene no. At3g20540) contains a presequence that is predicted (Emanuelsson et al. 2000) to target mitochondria (TargetP score 0.74) and chloroplasts (TargetP score 0.59). A related 117 kDa, 1049 amino acid *A. thaliana* protein (gene no. At1g50840) contains a putative plastid targeting presequence (TargetP score 0.933). Both N-termini deliver GFP to chloroplasts but the 116 kDa presequence appears to also target GFP to the mitochondria (Mori et al. 2005). Interestingly, the 117 kDa protein is targeted to mitochondria when translation initiates upstream of the predicted AUG start codon most probably at an in frame CUG lying seven codons

upstream of AUG (Christensen et al. 2005). Use of this non-AUG start codon suggests both *A. thaliana* organelle DNA polymerases are dual-targeted to plastids and mitochondria adding complexity to the regulation of organelle DNA polymerases.

Homologues of the *A. thaliana* organelle targeted DNA polymerases are present in *O. sativa*. Two *O. sativa* organellar DNA polymerases are predicted to be dual-targeted to plastids and mitochondria (Christensen et al. 2005). Polyclonal antibodies raised against one of these polymerases cross-reacted with a protein in isolated chloroplasts (Kimura et al. 2002). Expression of this DNA polymerase was studied by *in situ* hybridization. RNA was detected in tissues with dividing cells including leaf primordia, and the apical meristem of shoots and roots but not in mature leaves (Kimura et al. 2002). These genome based studies suggest plastids contain at least two ~110 kDa DNA polymerases that appear to be dual-targeted to mitochondria and plastids (Christensen et al. 2005; Mori et al. 2005). Future experiments should delineate the roles of these polymerases in organelle DNA replication versus repair. Recombinant forms of these enzymes will facilitate the characterization of their properties (Mori et al. 2005). These *A. thaliana* and *O. sativa* organelle DNA polymerases contain 3'-5' exonuclease and DNA polA-like domains. The distantly related apicoplast in the malaria parasite *Plasmodium falciparum* appears to contain a different DNA polymerase. A 235 kDa multidomain protein with DNA primase, DNA helicase, DNA polymerase and 3' to 5' exonuclease regions is implicated in replication of apicoplast DNA (Seow et al. 2005).

13.2 DNA primase activities in plastids

A primase activity purified from *P. sativum* chloroplasts was linked to a 115-120 kDa protein that was distinct from the 90 kDa DNA polymerase. The primase was resistant to tagetitoxin, an inhibitor of RNA polymerase, and able to initiate replication on poly dT, poly dC and cloned ssDNA templates containing *P. sativum* OriA (Nielsen et al. 1991). No primary sequence information is available on this primase.

13.3 Plastid localised RecA

The RecA protein promotes strand transfer and heteroduplex formation between DNA duplexes in prokaryotes (CameriniOtero and Hsieh 1995). RecA is required for DNA replication, recombination and repair pathways (Kowalczykowski 2000). The central role of RecA protein in homologous recombination is illustrated by its position in the double-strand break recombination model (Szostak et al. 1983) in Figure 15. Absence of *recA* function reduces the frequency of homologous recombination by 10,000–fold in *E. coli* (CameriniOtero and Hsieh 1995). Plastids contain a homologue of RecA, which was first identified as a 39 kDa protein in *P. sativum* chloroplasts using polyclonal antibodies against *E. coli* RecA protein

(Cerutti et al. 1992). This provided the first evidence for a RecA-mediated homologous recombination pathway in plastids. Accumulation of the *P. sativum* RecA-like protein appears to increase following exposure of protoplasts to DNA damaging agents (Cerutti et al. 1993). The presence of a RecA homologue in *P. sativum* chloroplasts has been recently confirmed by proteomics of protein-nucleic acid particles from purified chloroplasts (Phinney and Thelen 2005). The *A. thaliana* nuclear genome contains four genes encoding RecA-like proteins related to the bacterial RecA proteins (Khazi et al. 2003). Two of these genes encode organelle-targeted proteins (Khazi et al. 2003). Gene At1g79050 encodes a 439 amino acid protein (Table 2) that is predicted to be targeted to chloroplasts (TargetP score 0.841; Emanuelsson et al. 2000) and is imported into isolated *P. sativum* chloroplasts (Cao et al. 1997), whereas gene At3g10140 encodes a 389 amino acid protein that is targeted to mitochondria (Khazi et al. 2003). Both RecA-like proteins share 36% identity excluding N- and C-terminal extensions (Khazi et al. 2003). The At3g10140 mitochondrial RecA-like protein partially complements an *E. coli recA* mutant and provides tolerance to the methyl methane sulfonate (MMS) and mitomycin C (Khazi et al. 2003), which are DNA damaging agents.

C. reinhardtii plastids contain a RecA homologue (Nakazato et al. 2003). The influence of altered RecA activity on plastid recombination was addressed by expressing *E. coli* WT and dominant negative RecA proteins in *C. reinhardtii* chloroplasts (Cerutti et al. 1995). Expression of WT *E. coli* RecA in *C. reinhardtii* chloroplasts increased the frequency of homologous recombination between 216 bp direct repeats by over 15-fold. This indicates that recombination between direct repeats appears to be limited by RecA-mediated strand exchange. In contrast, overexpressed WT *E. coli* RecA did not enhance survival of cells exposed to DNA damaging agents indicating that chloroplast DNA repair pathways are not limited by RecA activity but by processing of DNA substrates generated by DNA damaging agents. An *E. coli* dominant-negative RecA protein reduced recombination between direct repeats and DNA repair in chloroplasts consistent with a negative influence on a RecA-mediated DNA-RRR pathway in plastids (Cerutti et al. 1995).

13.4 DNA topoisomerases

Topoisomerases regulate DNA supercoiling, DNA catenation and knotting, and are important enzymes required for DNA replication, recombination and repair. Type I topoisomerases relax supercoiled DNA whereas type II topoisomerases, such as DNA gyrase, not only relax supercoiled DNA but also introduce supercoils using ATP (Singh et al. 2004). Several reports have identified type I topoisomerases in chloroplasts including a 115 kDa protein in *S. oleracea* (Siedlecki et al. 1983), a 54 kDa protein in *Brassica oleracea* (cauliflower; Fukata et al. 1991), a 69 kDa protein in *P. sativum* (Mukherjee et al. 1994), and a 70 kDa type IB topoisomerase in *Sinapsis alba* (white mustard; Belkina et al. 2004). Evidence for type II topoisomerases in plastids include the observation that a *P. sativum* chloroplast transcription extract contained a gyrase-like activity which was sensitive to the gyrase inhibitor novobiocin (Lam and Chua 1987). Furthermore, an an-

tibody raised against yeast topoisomerase II cross-reacted with 96 and 101 kDa proteins in *T. aestivum* chloroplasts (Pyke et al. 1989); *E. coli* gyrase is comprised of 95 kDa gyrase B and 105 kDa gyrase A subunits (Reece and Maxwell 1991). A more recent proteomic study identified gyrase A- and B-like subunits in protein-nucleic acid particles from purified *P. sativum* chloroplasts (Phinney and Thelen 2005).

The *A. thaliana* nucleus contains one gene encoding a gyrase A-like subunit and three genes for gyrase B-like subunits (Wall et al. 2004). The gyrase A subunit (At3g10690) coding sequence has alternative start sites giving rise to plastid and mitochondrial targeted forms. T-DNA knockouts of the gyrase A subunit were embryo-lethal. One gyrase B subunit appeared to be targeted to plastids, the second to mitochondria, whereas the location of the third was unclear but was possibly located in the nucleus/cytosol (Wall et al. 2004). Knockouts of either organelle-targeted gyrase B subunit were seedling-lethal rather than the more severe embryo-lethal phenotype of gyrase A mutants. This suggested the gyrase B subunits complement each other to a limited extent indicating their products might be targeted to both organelles (Wall et al. 2004). This appears to be the case. An alternative upstream non-AUG start codon (most probably CUG) in one gyrase B subunit (gene At3g10270) gives rise to an N-terminus that confers dual-targeting to mitochondria and plastids. Alternative translation start sites in the coding sequence for the second organelle-targeted gyrase B subunit (gene At5g04130) give rise to either mitochondrial or plastid targeted proteins (Christensen et al. 2005). In summary, one gyrase A subunit and two gyrase B subunits appear to be targeted to both plastids and mitochondria in *A. thaliana* (Table 2). Presequences that confer dual-targeting of gyrase A and B subunits to mitochondria and chloroplasts have been found in *Nicotiana benthamiana* (Cho et al. 2004).

The effects of transient downregulation of organelle-targeted gyrase A and B subunit expression were studied in *N. benthamiana* by virus-induced gene silencing (Cho et al. 2004). Downregulation of gyrase A or B subunits prevents chloroplast development giving rise to white or yellow leaf sectors. Larger nucleoids and a mixture of heterogeneous high MW DNA molecules in plastids, possibly representing tangled DNA and their breakage products, are consistent with a crucial role for gyrase in untangling plastid DNA following replication and recombination (Cho et al. 2004). A role for gyrase in plastid DNA maintenance is supported by an earlier study where the gyrase inhibitors novobiocin and naladixic acid were shown to reduce the copy number of plastid DNA in *Solanum nigrum* suspension cultures (Ye and Sayre 1990).

In addition to roles in DNA-RRR pathways gyrase activity can also influence transcription (Chapter 5) through changes in supercoiling (Reece and Maxwell 1991). Mutations in gyrase activity might therefore impact on plastid gene expression as well as genome maintenance. The gyrase inhibitors novobiocin and naladixic acid were found to alter the accumulation of plastid transcripts in *C. reinhardtii* (Thompson and Mosig 1985). Addition of novobiocin to a *P. sativum* chloroplast transcription system containing cloned plastid genes inhibited the expression of the *atpB* gene to a larger extent than the *rbcL* gene. This raised the

possibility that template topology may enable differential regulation of plastid genes (Lam and Chua 1987).

13.5 DNA helicases

The DNA unwinding steps of DNA-RRR pathways include the generation of single-stranded recombination substrates (Fig. 15) and are carried out by ATP-dependent helicases. In an early study, a helicase fraction containing 6-8 protein bands was purified from *G. max* chloroplasts and shown to remove a 28 base oligomer from a single-stranded circular M13 template (Cannon and Heinhorst 1990). A similar biochemical approach identified 68 and 78 kDa helicases in purified *P. sativum* chloroplasts (Tuteja 2003). The 78 kDa helicase was stimulated by DNA fork structures indicating a role in replication (Tuteja and Phan 1998). Unwinding was inhibited by nogalamycin and ATPase activity by daunorubicin (Tuteja and Phan 1998). Both nogalamycin and daunorubicin are major groove intercalating agents. Chloroplast helicases appear to be sensitive to actinomycin C1 and resistant to ellipticine whereas the converse is true of nuclear helicases (Tuteja 2003).

A genomics based study has identified two *O. sativa* RecQ helicase homologues that are likely to be present in plastids (Saotome et al. 2006). Transient expression of GFP-fusions in onion epidermal cells indicated one 588 amino acid RecQ-like protein (OsRecQ1, Nucleotide Acc No. AK101124) was targeted to nuclei and plastids, whereas the second 844 amino acid Rec Q-like protein (OsRecQsim, Nucleotide Acc No. AK072977) was targeted predominantly to plastids. RNA levels for these RecQ-like proteins appeared highest in meristematic tissues containing immature plastids and did not increase in response to light and chloroplast development. A role for these proteins in repair was suggested by the observations that RNA encoding the 588 amino acid RecQ-like protein increased in levels in response to the four DNA damaging agents, mitomycin C, H_2O_2, MMS and bleomycin. Expression of the RNA encoding the 844 amino acid protein increased following treatment with mitomycin C and bleomycin and increased slightly with MMS (Saotome et al. 2006).

14 Identifying DNA-RRR proteins by complementation of *E. coli* mutants

Most of our knowledge of eubacterial DNA-RRR pathways is based on *E. coli* (CameriniOtero and Hsieh 1995; Kowalczykowski 2000). Plastids are probably descendants of an ancient cyanobacterium (Martin et al. 2002; Chapter 1) and have retained components of eubacterial DNA-RRR pathways including homologues of DNA polymerase, gyrase, and RecA (Table 2). In some cases, these plastid proteins have been shown to complement mutations in the homologous

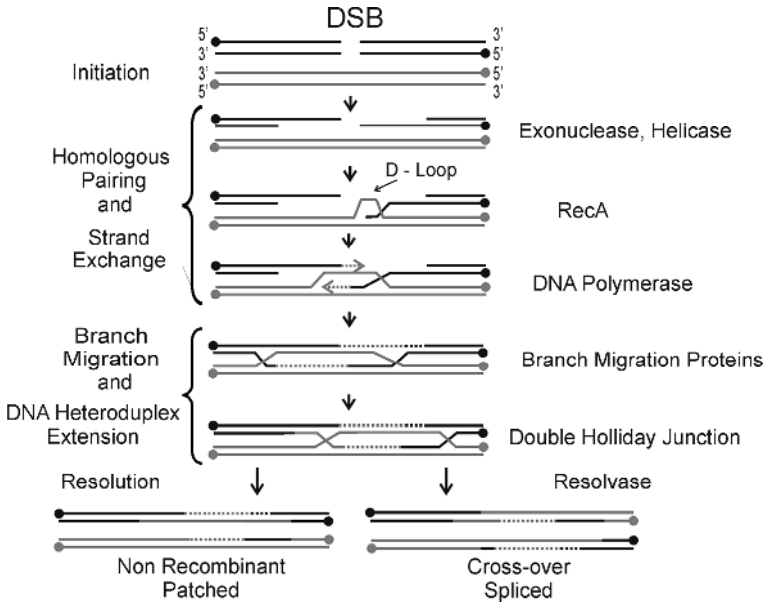

Fig. 15. Double-strand-break model of homologous recombination showing main proteins involved (Szostak et al. 1983; Kowalczykowski 2000).

E. coli proteins (Cho et al. 2004). Complementation of *E. coli* mutations provides a functional assay for identifying cDNAs encoding plastid-DNA RRR proteins andwould appear to be an attractive method for isolating plant homologues of *E. coli* DNA-RRR proteins. Using cDNA libraries, plant cDNAs that complement a number of *E. coli* mutations in DNA-RRR genes were isolated (Pang et al. 1993b) including *RecA* (Pang et al. 1992), and *ruvC* and *recG* mutants (Pang et al. 1993a). Unfortunately, the isolated cDNAs did not encode proteins related to characterized DNA-RRR proteins hindering progress. One of the isolated cDNAs complementing *ruvC* and *recG* double mutants was subsequently identified as plastocyanin (acc number P42699) raising questions on the validity of these library screening experiments. Screening cDNA libraries for functional complementation of *E. coli* mutants is technically difficult especially if complementation is not strong. Pitfalls include transformation or transfection of the tiny number of *E. coli* cells in which the mutation has reverted or been suppressed, which will give rise to viable cells on antibiotic medium. Alternatively, the cDNA may encode a protein that rescues the mutation indirectly, for example, by stabilising a temperature sensitive *E. coli* mutant protein.

15 Conclusions and outlook

Plastid transformation experiments have demonstrated an efficient homologous recombination pathway in plastids mediated by a RecA-homologue that appears to be active throughout shoot development. The presence of this pathway is consistent with a new emerging view of plastid DNA maintenance in which recombination plays a predominant role. WT plastid DNA is comprised of a mixture of circular and linear DNA molecules, which form a multimeric series from monomer to at least the octomer, and high MW DNA complexes (Section 3). Deleted plastid genomes in grasses contain sub-genomic circles and linear hairpin DNA molecules (Section 3.1, 6). The relationship between the mechanisms responsible for the maintenance of WT plastid DNA and the formation and replication of small linear DNA molecules in plastids is not understood. Replication models to account for the different topological forms of WT plastid DNA (circular DNA, linear DNA, branched complex DNA) have been proposed (Fig. 5). To identify which of these models are correct requires experimental confirmation beyond further descriptions of topological forms in WT chloroplasts. Progress in this research area requires the identification of proteins involved and mutants to determine the impact of loss or downregulation of these DNA-RRR proteins on plastid DNA levels and topological forms.

Whilst several approaches have localised putative replication origins in plastid DNA from flowering plants they have been mapped to different positions (Fig. 6) hindering the application of a universal model. Multiple locations for replication origins might reflect multiple origins in plastid DNA and differential usage of replication origins in different cells or differences in the accuracies of the methods used. The possibility of alternative modes of replication in different plastid types (Wang et al. 2003) increases the complexity of studying plastid genome maintenance. Distinguishing between these replication pathways might require the isolation and analysis of mutants affecting specific pathways. Recombination-dependent DNA replication plays an important role in genome maintenance in bacteria and has been suggested to be active in plastids to account for the complex branched DNA structures found in *Z. mays* plastids (Bendich 2004; Oldenburg and Bendich 2004b). Linear DNA molecules with heterodisperse or defined ends could invade template DNA molecules to prime DNA replication by recombination (Fig. 5d). The identification of linear DNA molecules with fixed ends that map to potential origins in the large inverted repeat (Oldenburg and Bendich 2004b; Scharff and Koop 2006) is interesting and characterisation of the structures of these ends might reveal the mechanisms involved in their formation. Whether recombination-dependent DNA replication is limited to a relatively small number of specific sites in plastids, possibly corresponding to the natural ends mapped in *Z. mays* (Oldenburg and Bendich 2004b) and *N. tabacum* (Scharff and Koop 2006), can be addressed by mutating these sites in recombinant plastid genomes (Scharff and Koop 2007). The finding that a high proportion of plastid DNA (50%) is comprised of complex branched DNA molecules in *Z. mays* seedlings (Oldenburg and Bendich 2004b) warrants further investigation in other species us-

ing additional techniques such as DNA fibre-based FISH with plastid DNA probes (Lilly et al. 2001) to study the organisation of these complexes.

A highly active homologous recombination pathway in plastids is consistent with recombination-dependent DNA replication. Widespread inter-molecular and intra-molecular recombination between large inverted repeat sequences or between repeated copies of the unit genome would be expected to produce a large number of isomers. If the molecules are linked by strand-invasion and recombination-dependent DNA replication this will give rise to a complex mixture of interconnected high-molecular-weight complexes (Oldenburg and Bendich 2004). The organisation of plastid DNA as high MW multi-genome complexes (containing linear and branched forms) has been suggested to underlie the packaging of plastid DNA into nucleoids (Bendich 2004). Random replication and recombination events are thought to contribute to the random segregation patterns observed for plastid genomes (Birky 1994, 2001). This raises the question of whether all genomes and topological forms have an equal chance of being replicated? Our current knowledge is too limited to address such a question. Other interesting areas worth exploring in future work include the relationship between topological forms and transcription, and the maintenance of heteroplasmic states. Distinct plastid genomes in heteroplasmic plants, where both genomes are required for survival, might be expected to segregate to different high MW DNA complexes within a plastid. However, the maintenance of different plastid genomes in the same high MW complexes might be possible and shed light on the dynamics of plastid genome maintenance. In normal WT plants copy-correction involving DNA repair pathways would be expected to ensure the maintenance of a uniform population of plastid DNA molecules.

Plastids have been evolving in the cytoplasm of their eukaryotic hosts for several billion years and have acquired proteins of nuclear or mitochondrial origin that were not present in the original cyanobacterial symbiont (see for example Wagner and Pfannschmidt 2006). Elucidation of DNA-RRR pathways in plastids should confirm roles for eukaryotic proteins (Mukherjee et al. 1994) in addition to roles for homologues of well known prokaryotic DNA-RRR proteins. It seems likely that the proteins, mechanisms, and regulation of plastid DNA-RRR pathways will have diverged substantially from the eubacterial DNA-RRR model. The availability of whole plant genome sequences allow genomic approaches to identify genes encoding proteins of prokaryotic (Table 2) and eukaryotic origin involved in plastid genome maintenance. A major problem is the prediction of plastid location due to the difficulty in identifying the N-termini of proteins from gene and cDNA sequences, and because computer programs (Emanuelsson et al. 2000) are only partially successful in predicting plastid-targeted proteins. Approximately 30% of chloroplast proteins do not contain recognisable plastid targeting signals (Kleffmann et al. 2004). Proteomics provides an alternative method to identify plastid DNA-RRR proteins. However, proteomic studies on chloroplasts (Kleffmann et al. 2004) have not uncovered the suite of DNA-RRR present in plastids possibly because of their limited abundance. More success has been achieved by proteomics of purified subfractions enriched in nucleoids (Sakai et al. 1999; Phinney and Thelen 2005). Alternatively, shoot tissues with actively dividing cells ex-

press a number of DNA-RRR proteins (Sakai et al. 1999; Saotome et al. 2006) and might provide better material for proteomic studies on plastid DNA-RRR proteins.

Reverse genetics provides a powerful tool to elucidate the roles of candidate DNA-RRR proteins in plastid genome maintenance. Knockout, using T-DNA insertions or transposons, and knockdown approaches, using RNAi, can be used to identify genes with important roles in plastid DNA maintenance. Knockdowns are particularly suitable for studying essential genes by allowing the isolation of viable plants. This provides the plant material in which to study the impact of DNA-RRR protein deficiencies on plastid genome maintenance. Plastid transformation allows the DNA substrates of plastid DNA-RRR pathways to be manipulated. Combining plastid transformation technologies with knockouts and knockdowns in nuclear genes is a particularly attractive method for studying plastid DNA-RRR pathways. Functional assays (see for example Fig. 11c) based on recombinant plastid genomes (Mühlbauer et al. 2002; Khakhlova and Bock 2006; Kode et al. 2006) will enable the impact of plastid DNA-RRR deficiencies on plastid DNA maintenance to be monitored. These new experimental approaches involving genomics, reverse genetics and transplastomic technologies, where both the trans-acting proteins and cis-acting DNA sequences can be manipulated, are likely to provide the functional studies needed to improve our understanding of the DNA-RRR pathways responsible for the maintenance of plastid genomes.

Acknowledgement

We are grateful to Professors Hans-Ulrich Koop (Munich) and Madeline Wu (Hong Kong) for providing information prior to publication and improvements to the text suggested by two anonymous reviewers. Work in the authors' laboratory was supported by the EU-FP6 Plastomics STREP consortium and the Biotechnology and Biological Sciences Research Council (UK).

References

Asai T, Bates DB, Kogoma T (1994) DNA replication triggered by double stranded breaks in *Escherichia coli*: dependence on homologous recombination functions. Cell 78:1051-1061

Backert S, Dorfel P, Börner T (1995) Investigation of plant organellar DNAs by pulsed-field gel-electrophoresis. Curr Genet 28:390-399

Bailey JC, Heinhorst S, Cannon GC (1995) Accuracy of deoxynucleotide incorporation by soybean chloroplast DNA polymerases is independent of the presence of a 3' to 5' exonuclease. Plant Physiol 107:1277-1284

Baumgartner BJ, Rapp JC, Mullet JE (1989) Plastid transcription activity and DNA copy number increase early in barley chloroplast development. Plant Physiol 89:1011-1018

Belkina GG, Pogul'skaya EV, Yurina NP (2004) Isolation and partial characterization of DNA topoisomerase I from the nucleoids of white mustard chloroplasts. Appl Biochem Microbiol 40:231-235

Bendich AJ (1987) Why do chloroplasts and mitochondria contain so many copies of their genome? Bioessays 6:279-282

Bendich AJ (2004) Circular chloroplast chromosomes: the grand illusion. Plant Cell 16:1661-1666

Bendich AJ, Smith SB (1990) Moving-pictures and pulsed-field gel-electrophoresis show linear DNA-molecules from chloroplasts and mitochondria. Curr Genet 17:421-425

Bensen RJ, Warner HR (1987) Partial purification and characterization of uracil-DNA glycosylase activity from chloroplasts of *Zea mays* Seedlings. Plant Physiol 84:1102-1106

Bi X, Liu LF (1996) Replicational model for DNA recombination between direct repeats. J Mol Biol 256:849-858

Birky CW (1994) Relaxed and stringent genomes: why cytoplasmic genes don't obey Mendel's Laws. J Heredity 85:355-365

Birky CW (2001) The inheritance of genes in mitochondria and chloroplasts: Laws, mechanisms, and models. Ann Rev Genet 35:125-148

Birky CW, Walsh JB (1992) Biased gene conversion, copy number, and apparent mutation rate differences within chloroplast and bacterial genomes. Genetics 130:677-683

Boynton JE, Gillham NW, Harris EH, Hosler JP, Johnson AM, Jones AR, Randolph-Anderson BL, Robertson D, Klein TM, Shark KB, Sanford JC (1988) Chloroplast transformation in *Chlamydomonas* with high velocity microprojectiles. Science 240:1534-1538

Cahoon AB, Cunningham KA, Bollenbach TJ, Stern DB (2003) Maize BMS cultured cell lines survive with massive plastid gene loss. Curr Genet 44:104-113

CameriniOtero RD, Hsieh P (1995) Homologous recombination proteins in prokaryotes and eukaryotes. Ann Rev Genet 29:509-552

Cannon GC, Heinhorst S (1990) Partial purification and characterization of a DNA helicase from chloroplasts of *Glycine max*. Plant Mol Biol 15:457-464

Cannon GC, Hedrick LA, Heinhorst S (1995) Repair mechanisms of UV-induced DNA damage in soybean chloroplasts. Plant Mol Biol 29:1267-1277

Cao J, Combs C, Jagendorf AT (1997) The chloroplast located homolog of bacterial DNA recombinase. Plant Cell Physiol 38:1319-1325

Carrillo N, Bogorad L (1988) Chloroplast DNA replication *in vitro*: site-specific initiation from preferred templates. Nucl Acids Res 16:5603-5620

Casjens S (1999) Evolution of the linear DNA replicons of the *Borrelia* spirochetes. Curr Opin Microbiol 2:529-534

Cavalier-Smith T (1974) Palindromic base sequences and replication of eukaryote chromosome ends. Nature 250:467-470

Cerutti H, Ibrahim HZ, Jagendorf AT (1993) Treatment of pea (*Pisum sativum* L) Protoplasts with DNA-damaging agents induces a 39-kilodalton chloroplast protein immunologically related to *Escherichia coli* RecA. Plant Physiol 102:155-163

Cerutti H, Johnson AM, Boynton JE, Gillham NW (1995) Inhibition of chloroplast DNA recombination and repair by dominant negative mutants of *Escherichia coli* RecA. Mol Cell Biol 15:3003-3011

Cerutti H, Osman M, Grandoni P, Jagendorf AT (1992) A homolog of *Escherichia coli* RecA protein in plastids of higher plants. Proc Natl Acad Sci USA 89:8068-8072

Chang CH, Wu M (2000) The effects of transcription and RNA processing on the initiation of chloroplast DNA replication in *Chlamydomonas reinhardtii*. Mol Gen Genet 263:320-327

Chen WL, Gaikwad A, Mukherjee SK, Choudhary NR, Kumar D, Tewari KK (1996) A 43 kDa DNA binding protein from the pea chloroplast interacts with and stimulates the cognate DNA polymerase. Nucl Acids Res 24:3953-3961

Chiang KS, Sueoka N (1967) Replication of chloroplast DNA in *Chlamydomonas reinhardi* during vegetative cell cycle: its mode and regulation. Proc Natl Acad Sci USA 57:1506-1513

Chiu WL, Sears BB (1992) Electron microscopic localization of replication origins in *Oenothera* chloroplast DNA. Mol Gen Genet 232:33-39

Cho HS, Lee SS, Kim KD, Hwang I, Lim JS, Park YI, Pai HS (2004) DNA gyrase is involved in chloroplast nucleoid partitioning. Plant Cell 16:2665-2682

Christensen AC, Lyznik A, Mohammed S, Elowsky CG, Elo A, Yule R, Mackenzie SA (2005) Dual-domain, dual-targeting organellar protein presequences in *Arabidopsis* can use non-AUG start codons. Plant Cell 17:2805-2816

Chumley TW, Palmer JD, Mower JP, Fourcade HM, Calie PJ, Boore JL, Jansen RK (2006) The complete chloroplast genome sequence of *Pelargonium* x hortorum: organization and evolution of the largest and most highly rearranged chloroplast genome of land plants. Mol Biol Evol 23:2175-2190

Collin S, Ellis THN (1991) Evidence for the presence of hairpin chloroplast DNA molecules in barley cultivars. Curr Genet 20:253-258

Corneille S, Lutz K, Svab Z, Maliga P (2001) Efficient elimination of selectable marker genes from the plastid genome by the CRE-lox site-specific recombination system. Plant J 27:171-178

Corneille S, Lutz KA, Azhagiri AK, Maliga P (2003) Identification of functional lox sites in the plastid genome. Plant J 35:753-762

Corriveau JL, Coleman AW (1988) Rapid screening method to detect potential biparental inheritance of plastid DNA and results for over 200 angiosperm species. Amer J Bot 75:1443-1458

Cox MM, Goodman MF, Kreuzer KN, Sherratt DJ, Sandler SJ, Marians KJ (2000) The importance of repairing stalled replication forks. Nature 404:37-41

Cuzzoni E, Giordani C, Stampacchia O, Bolchi A, Malcevschi A, Ottonello S, Ferretti L, Sala F (1995) Presence of a chloroplast DNA sequence in an autonomous circular DNA molecule in cultures rice cell (*Oryza sativa*). Plant Cell Physiol 36:717-72

Day A, Ellis THN (1984) Chloroplast DNA deletions associated with wheat plants regenerated from pollen: possible basis for maternal inheritance of chloroplasts. Cell 39:359-368

Day A, Ellis THN (1985) Deleted forms of plastid DNA in albino plants from cereal anther culture. Curr Genet 9:671-678

Dehaas JM, Boot KJM, Haring MA, Kool AJ, Nijkamp HJJ (1986) A *Petunia hybrida* chloroplast DNA region, close to one of the inverted repeats, shows sequence homology with the *Euglena gracilis* chloroplast DNA region that carries the putative replication origin. Mol Gen Genet 202:48-54

Deng XW, Wing RA, Gruissem W (1989) The chloroplast genome exists in multimeric forms. Proc Natl Acad Sci USA 86:4156-4160

Douglas SE, Penny SL (1999) The plastid genome of the cryptophyte alga, *Guillardia theta*: complete sequence and conserved synteny groups confirm its common ancestry with red algae. J Mol Evol 48:236-244

Doyle JJ, Davis JI, Soreng RJ, Garvin D, Anderson MJ (1992) Chloroplast DNA inversions and the origin of the grass family (*Poaceae*). Proc Natl Acad Sci USA 89:7722-7726

Draper CK, Hays JB (2000) Replication of chloroplast, mitochondrial and nuclear DNA during growth of unirradiated and UVB-irradiated *Arabidopsis* leaves. Plant J 23:255-265

Drescher A, Ruf S, Calsa T, Carrer H, Bock R (2000) The two largest chloroplast genome-encoded open reading frames of higher plants are essential genes. Plant J 22:97-104

Dufourmantel N, Dubald M, Matringe M, Canard H, Garcon F, Job C, Kay E, Wisniewski JP, Ferullo JM, Pelissier B, Sailland A, Tissot G (2007) Generation and characterization of soybean and marker-free tobacco plastid transformants over-expressing a bacterial 4-hydroxyphenylpyruvate dioxygenase which provides strong herbicide tolerance. Plant Biotechnol J 5:118-133

Dufourmantel N, Tissot G, Goutorbe F, Garcon F, Muhr C, Jansens S, Pelissier B, Peltier G, Dubald M (2005) Generation and analysis of soybean plastid transformants expressing *Bacillus thuringiensis* Cry1Ab protoxin. Plant Mol Biol 58:659-668

Ellis THN, Day A (1986) A hairpin plastid genome in barley. EMBO J 5:2769-2774

Emanuelsson O, Nielsen H, Brunak S, von Heijne G (2000) Predicting sub-cellular localization of proteins based on their N-terminal amino acid sequence. J Mol Biol 300:1005-1016

Epp MD (1973) Nuclear gene-induced plastome mutations in *Oenothera hookeri*. I. Genetic analysis. Genetics 75:465-483

Evans E, Alani E (2000) Roles for mismatch repair factors in regulating genetic recombination. Mol Cell Biol 20:7839-7844

Fischer N, Stampacchia O, Redding K, Rochaix JD (1996) Selectable marker recycling in the chloroplast. Mol Gen Genet 251:373-380

Fujie M, Kuroiwa H, Kawano S, Mutoh S, Kuroiwa T (1994) Behavior of organelles and their nucleoids in the shoot apical meristem during leaf development in *Arabidopsis thaliana* L. Planta 194:395-405

Fukata H, Mochida A, Maruyama N, Fukasawa H (1991) Chloroplast DNA topoisomerase I from cauliflower. J Biochem 109:127-131

Futcher AB (1986) Copy number amplification of the 2-μm circle plasmid of *Saccharomyces cerevisiae*. J Theor Biol 119:197-204

Gaikwad A, Hop DV, Mukherjee SK (2002) A 70-kDa chloroplast DNA polymerase from pea (*Pisum sativum*) that shows high processivity and displays moderate fidelity. Mol Genet Genomics 267:45-56

Gilbert DM (2004) In search of the holy replicator. Nat Rev Mol Cell Biol 5:848-854

Gillham NW (1974) Genetic analysis of chloroplast and mitochondrial genomes. Ann Rev Genet 8:347-391

Gold B, Carrillo N, Tewari KK, Bogorad L (1987) Nucleotide sequence of a preferred maize chloroplast genome template for *in vitro* DNA synthesis. Proc Natl Acad Sci USA 84:194-198

Goodman MF (2000) Coping with replication 'train wrecks' in *Escherichia coli* using Pol V, Pol II and RecA proteins. Trends Biochem Sci 25:189-195

Grant D, Swinton DC, Chiang KS (1978) Differential patterns of mitochondrial, chloroplastic and nuclear-DNA synthesis in synchronous cell-cycle of *Chlamydomonas reinhardtii*. Planta 141:259-267

Hada M, Hino K, Buchholz G, Goss J, Wellmann E, Shin M (2000) Assay of DNA photolyase activity in spinach leaves in relation to cell compartmentation: evidence for lack of DNA photolyase in chloroplasts. Biosci Biotechnol Biochem 64:1288-1291

Hagemann R (1976) Plastid distribution and plastid competition in higher plants and the induction of plastom mutations in higher plants by nitroso-urea compounts. In: Bücher T, Neupert W, Sebald W, Werner S (eds) Genetics and biogenesis of chloroplasts and mitochondria. Amsterdam: Elsevier, pp 331-338

Hajdukiewicz PTJ, Gilbertson L, Staub JM (2001) Multiple pathways for Cre/lox-mediated recombination in plastids. Plant J 27:161-170

Harada T, Ishikawa R, Niizeki M, Saito KI (1992) Pollen derived rice calli that have large deletions in plastid DNA do not require protein-synthesis in plastids for growth. Mol Gen Genet 233:145-150

Hartung F, Puchta H (2006) The Rec Q gene family in plants. J Plant Physiol 163:287-296

Hedrick LA, Heinhorst S, White MA, Cannon GC (1993) Analysis of soybean chloroplast DNA replication by 2-dimensional gel electrophoresis. Plant Mol Biol 23:779-792

Heinhorst S, Cannon G, Weissbach A (1985) Plastid and nuclear-DNA synthesis are not coupled in suspension cells of *Nicotiana tabacum*. Plant Mol Biol 4:3-12

Heinhorst S, Cannon GC, Weissbach A (1990) Chloroplast and mitochondrial DNA polymerases from cultured soybean cells. Plant Physiol 92:939-945

Herrmann RG, Bohnert HJ, Kowallik KV, Schmitt JM (1975) Size, conformation and purity of chloroplast DNA of some higher plants. Biochim Biophys Acta 378:305-317

Hess WR, Hubschmann T, Börner T (1994) Ribosome deficient plastids of albostrians barley: extreme representatives of nonphotosynthetic plastids. Endocytobiosis Cell Res 10:65-80

Hiratsuka J, Shimada H, Whittier R, Ishibashi T, Sakamoto M, Mori M, Kondo C, Honji Y, Sun CR, Meng BY, Li YQ, Kanno A, Nishizawa Y, Hirai A, Shinozaki K, Sugiura M (1989) The complete sequence of the rice (*Oryza sativa*) chloroplast genome: Intermolecular recombination between distinct transfer RNA genes accounts for a major plastid DNA inversion during the evolution of the cereals. Mol Gen Genet 217:185-194

Howe CJ, Smith AG (1991) Plants without chlorophyll. Nature 349:109-109

Hubscher U, Nasheuer HP, Syvaoja JE (2000) Eukaryotic DNA polymerases, a growing family. Trends Biochem Sci 25:143-147

Iamtham S, Day A (2000) Removal of antibiotic resistance genes from transgenic tobacco plastids. Nature Biotechnol 18:1172-1176

Jacob F, Cuzin F, Brenner S (1963) On regulation of DNA replication in bacteria. Cold Spring Harbor Symp Quant Biol 28:329-348

Kaneko T, Nakamura Y, Wolk CP, Kuritz T, Sasamoto S, Watanabe A, Iriguchi M, Ishikawa A, Kawashima K, Kimura T, Kishida Y, Kohara M, Matsumoto M, Matsuno A, Muraki A, Nakazaki N, Shimpo S, Sugimoto M, Takazawa M, Yamada M, Yasuda M, Tabata S (2001) Complete genomic sequence of the filamentous nitrogen-fixing *Cyanobacterium anabaena* sp strain PCC 7120. DNA Res 8:205-213

Kavanagh TA, Thanh ND, Lao NT, McGrath N, Peter SO, Horvath EM, Dix PJ, Medgyesy P (1999) Homeologous plastid DNA transformation in tobacco is mediated by multiple recombination events. Genetics 152:1111-1122

Kawata M, Harada T, Shimamoto Y, Oono K, Takaiwa F (1997) Short inverted repeats function as hotspots of intermolecular recombination giving rise to oligomers of deleted plastid DNAs (ptDNAs). Curr Genet 31:179-184

Keim CA, Mosbaugh DW (1991) Identification and characterization of a 3' to 5' exonuclease Associated with spinach chloroplast DNA-polymerase. Biochemistry 30:11109-11118

Kelchner SA, Wendel JF (1996) Hairpins create minute inversions in non-coding regions of chloroplast DNA. Curr Genet 30:259-262

Khakhlova O, Bock R (2006) Elimination of deleterious mutations in plastid genomes by gene conversion. Plant J 46:85-94

Khazi FR, Edmondson AC, Nielsen BL (2003) An *Arabidopsis* homologue of bacterial RecA that complements an *E. coli* RecA deletion is targeted to plant mitochondria. Mol Genet Genomics 269:454-463

Kimura S, Sakaguchi K (2006) DNA repair in plants. Chem Rev 106:753-766

Kimura S, Uchiyama Y, Kasai N, Namekawa S, Saotome A, Ueda T, Ando T, Ishibashi T, Oshige M, Furukawa T, Yamamoto T, Hashimoto J, Sakaguchi K (2002) A novel DNA polymerase homologous to *Escherichia coli* DNA polymerase I from a higher plant, rice (*Oryza sativa* L.). Nucl Acids Res 30:1585-1592

Klaus SMJ, Huang FC, Golds TJ, Koop HU (2004) Generation of marker-free plastid transformants using a transiently cointegrated selection gene. Nature Biotechnol 22:225-229

Kleffmann T, Russenberger D, von Zychlinski A, Christopher W, Sjolander K, Gruissem W, Baginsky S (2004) The *Arabidopsis thaliana* chloroplast proteome reveals pathway abundance and novel protein functions. Curr Biol 14:354-362

Kode V, Mudd EA, Iamtham S, Day A (2005) The tobacco *acc*D gene is essential and is required for leaf development. Plant J 44:237-244

Kode V, Mudd EA, Iamtham S, Day A (2006) Isolation of precise plastid deletion mutants by homology-based excision: a resource for site-directed mutagenesis, multi-gene changes and high-throughput plastid transformation. Plant J 46:901-909

Kolodner R, Tewari KK (1972) Molecular-size and conformation of chloroplast deoxyribonucleic acid from pea leaves. J Biol Chem 247:6355-6364

Kolodner RD, Tewari KK (1975) Chloroplast DNA from higher plants replicates by both Cairns and rolling circle mechanism. Nature 256:708-711

Kolodner R, Tewari KK (1979) Inverted repeats in chloroplast DNA from higher plants. Proc Natl Acad Sci USA 76:41-45

Koumandou VL, Nisbet RER, Barbrook AC, Howe CJ (2004) Dinoflagellate chloroplasts: where have all the genes gone? Trends Genet 20:261-267

Kowalczykowski SC (2000) Initiation of genetic recombination and recombination-dependent replication. Trends Biochem Sci 25:156-165

Kowallik KV, Stoebe B, Schaffran I, KrothPancic P, Freier U (1995) The chloroplast genome of a chlorophyll a+c-containing alga, *Odontella sinensis*. Plant Mol Biol Rep 13:336-342

Kreuzer KN (2000) Recombination-dependent DNA replication in phage T4. Trends Biochem Sci 25:165-173

Kreuzer KN (2005) Interplay between DNA replication and recombination in prokaryotes. Ann Rev Microbiol 59:43-67

Kunnimalaiyaan M, Nielsen BL (1997a) Fine mapping of replication origins (oriA and oriB) in *Nicotiana tabacum* chloroplast DNA. Nucl Acids Res 25:3681-3686

Kunnimalaiyaan M, Shi F, Nielsen BL (1997b) Analysis of the tobacco chloroplast DNA replication origin (oriB) downstream of the 23 S rRNA gene. J Mol Biol 268:273-283

Kuroiwa T (1991) The replication, differentiation, and inheritance of plastids with emphasis on the concept of organelle nuclei. Int Rev Cytol 128:1-62

Lam E, Chua NH (1987) Chloroplast DNA gyrase and *in vitro* regulation of transcription by template topology and novobiocin. Plant Mol Biol 8:415-424

Lamppa GK, Bendich AJ (1979) Changes in chloroplast DNA levels during development of pea (*Pisum sativum*). Plant Physiol 64:126-130

Lawrence ME, Possingham JV (1986) Microspectrofluorometric measurement of chloroplast DNA in dividing and expanding leaf cells of *Spinacia oleracea*. Plant Physiol 81:708-710

Lelivelt CLC, McCabe MS, Newell CA, deSnoo CB, van Dun KMP, Birch-Machin I, Gray JC, Mills KHG, Nugent JM (2005) Stable plastid transformation in lettuce (*Lactuca sativa* L.). Plant Mol Biol 58:763-774

Li WM, Ruf S, Bock R (2006) Constancy of organellar genome copy numbers during leaf development and senescence in higher plants. Mol Genet Genomics 275:185-192

Lilly JW, Havey MJ, Jackson SA, Jiang JM (2001) Cytogenomic analyses reveal the structural plasticity of the chloroplast genome in higher plants. Plant Cell 13:245-254

Liu NT, Jane WN, Tsay HS, Wu H, Chang WC, Lin CS (2007) Chloroplast genome aberration in micropropagation-derived albino *Bambusa edulis* mutants, ab1 and ab2. Plant Cell Tiss Org Culture 88:147-156

Lovett ST (2004) Encoded errors: mutations and rearrangements mediated by misalignment at repetitive DNA sequences. Mol Microbiol 52:1243-1253

Lu Z, Kunnimalaiyaan M, Nielsen BL (1996) Characterization of replication origins flanking the 23S rRNA gene in tobacco chloroplast DNA. Plant Mol Biol 32:693-706

Martin W, Rujan T, Richly E, Hansen A, Cornelsen S, Lins T, Leister D, Stoebe B, Hasegawa M, Penny D (2002) Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. Proc Natl Acad Sci USA 99:12246-12251

Maul JE, Lilly JW, Cui LY, dePamphilis CW, Miller W, Harris EH, Stern DB (2002) The *Chlamydomonas reinhardtti* plastid chromosome: Islands of genes in a sea of repeats. Plant Cell 14:2659-2679

McKown RL, Tewari KK (1984) Purification and properties of a pea chloroplast DNA polymerase. Proc Natl Acad Sci USA 81:2354-2358

Medgyesy P, Fejes E, Maliga P (1985) Interspecific chloroplast recombination in a *Nicotiana* somatic hybrid. Proc Natl Acad Sci USA 82:6960-6964

Meeker R, Nielsen B, Tewari KK (1988) Localization of replication origins in pea chloroplast DNA. Mol Cell Biol 8:1216-1223

Miyamura S, Kuroiwa T, Nagata T (1990) Multiplication and differentiation of plastid nucleoids during development of chloroplasts and etioplasts from proplastids in *Triticum aestivum*. Plant Cell Physiol 31:597-602

Miyamura S, Nagata T, Kuroiwa T (1986) Quantitative fluorescence microscopy on dynamic changes of plastid nucleoids during wheat development. Protoplasma 133:66-72

Morgan A, Maliga P (1987) Rapid chloroplast segregation and recombination of mitochondrial DNA in *Brassica* cybrids. Mol Gen Genet 209:240-246

Mori Y, Kimura S, Saotome A, Kasai N, Sakaguchi N, Uchiyama Y, Ishibashi T, Yamamoto T, Chiku H, Sakaguchi K (2005) Plastid DNA polymerases from higher plants, *Arabidopsis thaliana*. Biochem Biophys Res Comm 334:43-50

Mühlbauer SK, Lossl A, Tzekova L, Zou ZR, Koop HU (2002) Functional analysis of plastid DNA replication origins in tobacco by targeted inactivation. Plant J 32:175-184

Mukherjee SK, Reddy MK, Kumar D, Tewari KK (1994) Purification and characterization of a eukaryotic type I topoisomerase from pea chloroplast. J Biol Chem 269:3793-3801

Muller HJ (1964) The relation of recombination to mutational advance. Mutation Res 1:2-9

Nakayama N, Arai N, Bond MW, Kaziro Y, Arai K (1984) Nucleotide sequence of *dna* B and the primary structure of the dna B protein from *Escherichia coli*. J Biol Chem 259:97-101

Nakazato E, Fukuzawa H, Tabata S, Takahashi H, Tanaka K (2003) Identification and expression analysis of cDNA encoding a chloroplast recombination protein REC1, the chloroplast RecA homologue in *Chlamydomonas reinhardtii*. Biosci Biotechnol Biochem 67:2608-2613

Nelson MJ, Green BR (2005) Double hairpin elements and tandem repeats in the noncoding region of *Adenoides eludens* chloroplast gene minicircles. Gene 358:102-110

Newman SM, Boynton JE, Gillham NW, Randolphanderson BL, Johnson AM, Harris EH (1990) Transformation of chloroplast ribosomal RNA genes in *Chlamydomonas*: molecular and genetic-characterization of integration events. Genetics 126:875-888

Newman SM, Harris EH, Johnson AM, Boynton JE, Gillham NW (1992) Nonrandom distribution of chloroplast recombination events in *Chlamydomonas reinhardtii*: evidence for a hotspot and an adjacent cold region. Genetics 132:413-429

Nielsen BL, Rajasekhar VK, Tewari KK (1991) Pea chloroplast DNA primase: characterization and role in initiation of replication. Plant Mol Biol 16:1019-1034

Ogihara Y, Isono K, Kojima T, Endo A, Hanaoka M, Shiina T, Terachi T, Utsugi S, Murata M, Mori N, Takumi S, Ikeo K, Gojobori T, Murai R, Murai K, Matsuoka Y, Ohnishi Y, Tajiri H, Tsunewaki K (2000) Chinese spring wheat (*Triticum aestivum* L.) chloroplast genome: Complete sequence and contig clones. Plant Mol Biol Rep 18:243-253

Ogihara Y, Isono K, Kojima T, Endo A, Hanaoka M, Shiina T, Terachi T, Utsugi S, Murata M, Mori N, Takumi S, Ikeo K, Gojobori T, Murai R, Murai K, Matsuoka Y, Ohnishi Y, Tajiri H, Tsunewaki K (2002) Structural features of a wheat plastome as revealed by complete sequencing of chloroplast DNA. Mol Genet Genomics 266:740-746

Ohtani T, Uchimiya H, Kato A, Harada H, Sugita M, Sugiura M (1984) Location and nucleotide-sequence of a tobacco chloroplast DNA segment capable of replication in yeast. Mol Gen Genet 195:1-4

Oldenburg DJ, Bendich AJ (1998) The structure of mitochondrial DNA from the liverwort, *Marchantia polymorpha*. J Mol Biol 276:745-758

Oldenburg DJ, Bendich AJ (2001) Mitochondrial DNA from the liverwort *Marchantia polymorpha*: Circularly permuted linear molecules, head-to-tail concatemers, and a 5 ' protein. J Mol Biol 310:549-56

Oldenburg DJ, Bendich AJ (2004a) Changes in the structure of DNA molecules and the amount of DNA per plastid during chloroplast development in maize. J Mol Biol 344:1311-1330

Oldenburg DJ, Bendich AJ (2004b) Most chloroplast DNA of maize seedlings in linear molecules with defined ends and branched forms. J Mol Biol 335:953-970

Oldenburg DJ, Rowan BA, Zhao L, Watcher CL, Schleh M, Bendich AJ (2006) Loss or retention of chloroplast DNA in maize seedlings is affected by both light and genotype. Planta 225:41-55

Palmer JD (1983) Chloroplast DNA exists in 2 orientations. Nature 301: 92-93

Palmer JD (1990) Contrasting modes and tempos of genome evolution in land plant organelles. Trends Genet 6:115-120

Palmer JD, Thompson WF (1982) Chloroplast DNA rearrangements are more frequent when a large inverted repeat sequence is lost. Cell 29:537-550

Pang QS, Hays JB, Rajagopal I (1992) A plant cDNA that partially complements *Escherichia coli* RecA mutations predicts a polypeptide not strongly homologous to RecA proteins. Proc Natl Acad Sci USA 89:8073-8077

Pang QS, Hays JB, Rajagopal I (1993a) Two cDNAs from the plant *Arabidopsis thaliana* that partially restore recombination proficiency and DNA damage resistance to *Escherichia coli* mutants lacking recombination intermediate resolution activities. Nucl Acids Res 21:1647-1653

Pang QS, Hays JB, Rajagopal I, Schaefer TS (1993b) Selection of *Arabidopsis* cDNAs that partially correct phenotypes of *Escherichia coli* DNA damage sensitive mutants and analysis of 2 plant cDNAs that appear to express UV specific dark repair activities. Plant Mol Biol 22:411-426

Perry AS, Wolfe KH (2002) Nucleotide substitution rates in legume chloroplast DNA depend on the presence of the inverted repeat. J Mol Evol 55:501-508

Phinney BS, Thelen JJ (2005) Proteomic characterization of a triton-insoluble fraction from chloroplasts defines a novel group of proteins associated with macromolecular structures. J Proteome Res 4:497-506

Przykorska A, Solecka K, Olszak K, Keith G, Nawrot B, Kuligowska E (2004) Wheat (*Triticum vulgare*) chloroplast nuclease ChSI exhibits 5 ' flap structure-specific endonuclease activity. Biochemistry 43:11283-11294

Pyke KA, Marrison J, Leech RM (1989) Evidence for a Type-II topoisomerase in wheat chloroplasts. FEBS Letts 242:305-308

Qin ZJ, Cohen SN (2000) Long palindromes formed in *Streptomyces* by nonrecombinational intra-strand annealing. Genes Develop 14:1789-1796

Raynaud C, Perennes C, Reuzeau C, Catrice O, Brown S, Bergounioux C (2005) Cell and plastid division are coordinated through the prereplication factor AtCDT1. Proc Natl Acad Sci USA 102:8216-8221

Reboud X, Zeyl C (1994) Organelle inheritance in plants. Heredity 72:132-140

Reece RJ, Maxwell A (1991) DNA gyrase - structure and function. Crit Rev Biochem Mol Biol 26:335-375

Reith M, Munholland J (1995) Complete nucleotide sequence of the *Porphyra purpurea* chloroplast genome. Plant Mol Biol Rep 13:333-335

Rogalski M, Ruf S, Bock R (2006) Tobacco plastid ribosomal protein S18 is essential for cell survival. Nucl Acids Res 34:4537-4545

Rowan BA, Oldenburg DJ, Bendich AJ (2004) The demise of chloroplast DNA in *Arabidopsis*. Curr Genet 46:176-181

Ruf S, Hermann M, Berger IJ, Carrer H, Bock R (2001) Stable genetic transformation of tomato plastids and expression of a foreign protein in fruit. Nat Biotechnol 19:870-875

Rybchin VN, Svarchevsky AN (1999) The plasmid prophage N15: a linear DNA with covalently closed ends. Mol Microbiol 33:895-903

Sakai A, Suzuki T, Nagata N, Sasaki N, Miyazawa Y, Saito C, Inada N, Nishimura Y, Kuroiwa T (1999) Comparative analysis of DNA synthesis activity in plastid-nuclei and mitochondrial-nuclei simultaneously isolated from cultured tobacco cells. Plant Sci 140:9-19

Sakai A, Takano H, Kuroiwa T (2004) Organelle nuclei in higher plants: Structure, composition, function, and evolution. Int Rev Cytol 238:59-118

Sala F, Amileni AR, Parisi B, Spadari S (1980) A gamma-like DNA polymerase in spinach chloroplasts. Eur J Biochem 112:211-217

Saotome A, Kimura S, Mori Y, Uchiyama Y, Morohashi K, Sakaguchi K (2006) Characterization of four Rec Q homologues from rice (*Oryza sativa* L. cv. Nipponbare). Biochem Biophys Res Commun 345:1283-1291

Saski C, Lee SB, Daniell H, Wood TC, Tomkins J, Kim HG, Jansen RK (2005) Complete chloroplast genome sequence of *Glycine max* and comparative analyses with other legume genomes. Plant Mol Biol 59:309-322

Sato N, Terasawa K, Miyajima K, Kabeya Y (2003) Organization, developmental dynamics, and evolution of plastid nucleoids. Int Rev Cytol 232:217-262

Scharff LB, Koop HU (2006) Linear molecules of tobacco ptDNA end at known replication origins and additional loci. Plant Mol Biol 62:611-621

Scharff LB, Koop HU (2007) Targeted inactivation of the tobacco plastome origins of replication A and B. Plant J: in press

Sears BB (1998) Replication, recombination and repair in the chloroplast genetic system of *Chlamydomonas*. In: Rochaix JD, Goldschmidt-Clermont M, Merchant S (eds) The molecular biology of chloroplasts and mitochondria in *Chlamydomonas*. Dordrecht-Boston-London: Kluwer, 115-138

Sears BB, Stoike LL, Chiu WL (1996) Proliferation of direct repeats near the *Oenothera* chloroplast DNA origin of replication. Mol Biol Evol 13:850-863

Seow F, Sato S, Janssen CS, Riehle MO, Mukhopadhyay A, Phillips RS, Wilson RJM, Barrett MP (2005) The plastidic DNA replication enzyme complex of *Plasmodium falciparum*. Mol Biochem Parasitol 141:145-153

Shaver JM, Oldenburg DJ, Bendich AJ (2006) Changes in chloroplast DNA during development in tobacco, *Medicago truncatula*, pea, and maize. Planta 224:72-82

Shen P, Huang HV (1986) Homologous recombination in *Escherichia coli* dependence on substrate length and homology. Genetics 112:441-457

Shikanai T, Shimizu K, Ueda K, Nishimura Y, Kuroiwa T, Hashimoto T (2001) The chloroplast *clp* P gene, encoding a proteolytic subunit of ATP-dependent protease, is indispensable for chloroplast development in tobacco. Plant Cell Physiol 42:264-273

Shinozaki K, Ohme M, Tanaka M, Wakasugi T, Hayashida N, Matsubayashi T, Zaita N, Chunwongse J, Obokata J, Yamaguchishinozaki K, Ohto C, Torazawa K, Meng BY, Sugita M, Deno H, Kamogashira T, Yamada K, Kusuda J, Takaiwa F, Kato A, Tohdoh N, Shimada H, Sugiura M (1986) The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. EMBO J 5:2043-2049

Sidorov VA, Kasten D, Pang SZ, Hajdukiewicz PTJ, Staub JM, Nehra NS (1999) Stable chloroplast transformation in potato: use of green fluorescent protein as a plastid marker. Plant J 19:209-216

Siedlecki J, Zimmermann W, Weissbach A (1983) Characterization of a prokaryotic topoisomerase I activity in chloroplast cxtracts from spinach. Nucl Acids Res 11:1523-1536

Singh BN, Sopory SK, Reddy MK (2004) Plant DNA topoisomerases: Structure, function, and cellular roles in plant development. Crit Rev Plant Sci 23:251-269

Soltis PS, Soltis DE, Chase MW (1999) Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology. Nature 402:402-404

Spencer D, Whitfeld PR (1969) Characteristics of spinach chloroplast DNA polymerase. Arch Biochem Biophy 132:477-488

Staub JM, Maliga P (1992) Long regions of homologous DNA are incorporated into the tobacco plastid genome by transformation. Plant Cell 4:39-45
Staub JM, Maliga P (1994) Extrachromosomal elements in tobacco plastids. Proc Natl Acad Sci USA 91:7468-7472
Sternberg N, Hamilton D (1981) Bacteriophage-P1 site-specific recombination. I. Recombination between LoxP sites. J Mol Biol 150:467-486
Stoike LL, Sears BB (1998) Plastome mutator-induced alterations arise in Oenothera chloroplast DNA through template slippage. Genetics 149:347-353
Streisinger G, Edgar RS, Denhardt GH (1964) Chromosome structure in phage T4, I. circularity of linkage map. Proc Natl Acad Sci USA 51:775-779
Suzuki H, Ingersoll J, Stern DB, Kindle KL (1997) Generation and maintenance of tandemly repeated extrachromosomal plasmid DNA in *Chlamydomonas* chloroplasts. Plant J 11:635-648
Svab Z, Hajdukiewicz P, Maliga P (1990) Stable transformation of plastids in higher plants. Proc Natl Acad Sci USA 87:8526-8530
Svab Z, Maliga P (1993) High frequency plastid transformation in tobacco by selection for a chimeric *aad*A Gene. Proc Natl Acad Sci USA 90:913-917
Szostak JW, Orrweaver TL, Rothstein RJ, Stahl FW (1983) The double strand break repair model for recombination. Cell 33:25-35
Takeda Y, Hirokawa H, Nagata T (1992) The replication origin of proplastid DNA in cultured cells of tobacco. Mol Gen Genet 232:191-198
Thompson RJ, Mosig G (1985) An ATP dependent supercoiling topoisomerase of *Chlamydomonas reinhardtii* affects accumulation of specific chloroplast transcripts. Nucleic Acids Res 13:873-891
Tuteja N (2003) Plant DNA helicases: the long unwinding road. J Exp Bot 54:2201-2214
Tuteja N, Phan TN (1998) A chloroplast DNA helicase II from pea that prefers fork-like replication structures. Plant Physiol 118:1029-1039
VanWinkle-Swift KP (1980) A model for the rapid vegetative segregation of multiple chloroplast genomes in *Chlamydomonas* - assumptions and predictions of the model. Curr Genet 1:113-125
Velemínský J, Švachulová J, Šatava J (1980) Endonucleases for UV-irradiated and depurinated DNA in barley chloroplasts. Nucl Acids Res 8:1373-1381
Waddell J, Wang XM, Wu M (1984) Electron microscopic localization of the chloroplast DNA replicative origins in *Chlamydomonas reinhardii*. Nucl Acids Res 12:3843-3856
Wagner R, Pfannschmidt T (2006) Eukaryotic transcription factors in plastids: bioinformatic assessment and implications for the evolution of gene expression machineries in plants. Gene 381:62-70
Wall MK, Mitchenall LA, Maxwell A (2004) *Arabidopsis thaliana* DNA gyrase is targeted to chloroplasts and mitochondria. Proc Natl Acad Sci USA 101:7821-7826
Wang TSF (1991) Eukaryotic DNA polymerases. Ann Rev Biochem 60:513-552
Wang Y, Saitoh Y, Sato T, Hidaka S, Tsutsumi K (2003) Comparison of plastid DNA replication in different cells and tissues of the rice plant. Plant Mol Biol 52:905-913
Wang ZF, Yang JM, Nie ZQ, Wu M (1991) Purification and characterization of a gamma-like DNA-polymerase from *Chlamydomonas reinhardtii*. Biochemistry 30:1127-1131
Williamson D (2002) Timeline-The curious history of yeast mitochondrial DNA. Nat Rev Genet 3:475-481

Woelfle MA, Thompson RJ, Mosig G (1993) Roles of novobiocin sensitive topoisomerases in chloroplast DNA replication in *Chlamydomonas reinhardtii*. Nucl Acids Res 21:4231-4238

Wolfe KH, Li WH, Sharp PM (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proc Natl Acad Sci USA 84:9054-9058

Wolfe KH, Morden CW, Palmer JD (1992) Function and evolution of a minimal plastid genome from a nonphotosynthetic parasitic plant. Proc Natl Acad Sci USA 89:10648-10652

Wycliffe P, Sitbon F, Wernersson J, Ezcurra I, Ellerstrom M, Rask L (2005) Continuous expression in tobacco leaves of a *Brassica napus* PEND homologue blocks differentiation of plastids and development of palisade cells. Plant J 44:1-15

Ye JS, Sayre RT (1990) Reduction of chloroplast DNA content in *Solanum nigrum* suspension cells by treatment with chloroplast DNA synthesis inhibitors. Plant Physiol 94:1477-1483

Zoschke R, Liere K, Börner T (2007) From seedling to mature plant: *Arabidopsis* plastidial genome copy number, RNA accumulation and transcription are differentially regulated during leaf development. Plant J 50:710-722

Zoubenko OV, Allison LA, Svab Z, Maliga P (1994) Efficient targeting of foreign genes into the tobacco plastid genome. Nucl Acids Res 22:3819-3824

Zubko MK, Day A (1998) Stable albinism induced without mutagenesis: a model for ribosome-free plastid inheritance. Plant J 15:265-271

Zubko MK, Day A (2002) Differential regulation of genes transcribed by nucleus-encoded plastid RNA polymerase, and DNA amplification, within ribosome- deficient plastids in stable phenocopies of cereal albino mutants. Mol Gen Genomics 267:27-37

Zubko MK, Zubko EI, van Zuilen K, Meyer P, Day A (2004) Stable transformation of petunia plastids. Transgenic Res 13:523-530

Day, Anil
Faculty of Life Sciences, The University of Manchester, Oxford Road, Manchester M13 9PT, UK
anil.day@manchester.ac.uk

Madesis, Panagiotis
Faculty of Life Sciences, The University of Manchester, Oxford Road, Manchester M13 9PT, UK

List of abbreviations

DAPI: 4',6-diamidino-2-phenylindole
DSB: Double strand break
D-loop: Displacement loop
DNA-RRR: DNA replication, recombination, and repair
FISH: Fluorescent *in situ* hybridization
G. max: *Glycine max* (soybean)
H. vulgare: Hordeum vulgare (barley)
IR: Inverted repeat
L. sativa: *Lactuca sativa* (lettuce)
MMS: methyl methane sulfonate
MW: molecular weight
N. tabacum: Nicotiana tabacum (tobacco)
O. hookeri: Oenothera hookeri (evening primrose)
O. sativa: *Oryza sativa* (rice)
Ori: origin of replication
P. sativum: *Pisum sativum* (pea)
RBS: Ribosome binding site
S. oleracea: Spinacia oleracea (spinach)
T. aestivum: Triticum aestivum (wheat)
UTR: Untranslated region
WT: wild type
Z. mays: Zea mays (maize)