

Volker Knoop

The mitochondrial DNA of land plants: peculiarities in phylogenetic perspective

Received: 4 April 2004 / Revised: 6 July 2004 / Accepted: 7 July 2004 / Published online: 6 August 2004
© Springer-Verlag 2004

Abstract Land plants exhibit a significant evolutionary plasticity in their mitochondrial DNA (mtDNA), which contrasts with the more conservative evolution of their chloroplast genomes. Frequent genomic rearrangements, the incorporation of foreign DNA from the nuclear and chloroplast genomes, an ongoing transfer of genes to the nucleus in recent evolutionary times and the disruption of gene continuity in introns or exons are the hallmarks of plant mtDNA, at least in flowering plants. Peculiarities of gene expression, most notably RNA editing and *trans*-splicing, are significantly more pronounced in land plant mitochondria than in chloroplasts. At the same time, mtDNA is generally the most slowly evolving of the three plant cell genomes on the sequence level, with unique exceptions in only some plant lineages. The slow sequence evolution and a variable occurrence of introns in plant mtDNA provide an attractive reservoir of phylogenetic information to trace the phylogeny of older land plant clades, which is as yet not fully resolved. This review attempts to summarize the unique aspects of land plant mitochondrial evolution from a phylogenetic perspective.

Keywords Group I introns · Group II introns · RNA editing · *Trans*-splicing · Gene transfer · Land plant phylogeny

Complexity of mitochondrial genomes

Today there is no remaining doubt that the organelle genomes of mitochondria (mt) and chloroplasts (cp) are

reduced forms of the genomes of α -Proteobacteria and Cyanobacteria, respectively, which as endosymbionts gave rise to these organelles in the emerging eukaryotic cell more than 10^9 years ago (Dyall et al. 2004; Gray 1999; Gray et al. 1999). The mainstream of organelle genomic evolution has been the massive loss and transfer of genes to the eukaryotic nucleus, thus reducing the bacterial genome of some 4 mbp and 3,000 genes to an organelle genome of 150 kbp and 120 genes, as for example in land plant cpDNAs (Timmis et al. 2004).

Investigation of plant mitochondrial DNA (mtDNA) has continually revealed surprises that one would not have expected to see in the supposedly simple and reduced genome of a former endosymbiotic bacterium (Fig. 1). The mitochondrial genome (the chondriome) of plants contrasts with its generally simple, streamlined, circular and compact counterpart of about 16 kbp in most animals in nearly every aspect. Even the smallest as yet known chondriome of a land plant (embryophyte), that of the liverwort *Marchantia polymorpha* (Oda et al. 1992), is 186.6 kbp, more than ten times as large as that of humans; and chondriome sizes can exceed 2,000 kbp in flowering plants of the Cucurbitaceae family (Ward et al. 1981). Instead of the compact, gene-dense and simple circular mitochondrial genome of animals (metazoa), a complex pool of frequently recombining molecules, the stoichiometry of which is controlled by nuclear genes (Abdelnoor et al. 2003), typically constitute flowering plant (angiosperm) chondriomes.

Deviations from the gene-dense, small-circle chondriome architecture characteristic of the animal (metazoan) lineage were very recently described in phylogenetically early-branching protists, *Monosiga brevicollis* and *Amoebidium parasiticum* (Burger et al. 2003a), including genes that were transferred to the nucleus before the emergence of the metazoan lineage. The exceptional discovery of introns in the mtDNA of the sea anemone *Metridium senile* (Beagley et al. 1996) was a first hint that peculiar chondriome features can exist in early branches of the animal lineage.

Communicated by A. Brennicke

V. Knoop
IZMB—Institut für Zelluläre und Molekulare Botanik,
Universität Bonn, Kirschallee 1, 53115 Bonn,
Germany
E-mail: volker.knoop@uni-bonn.de
Tel.: +49-228-736466
Fax: +49-228-736467

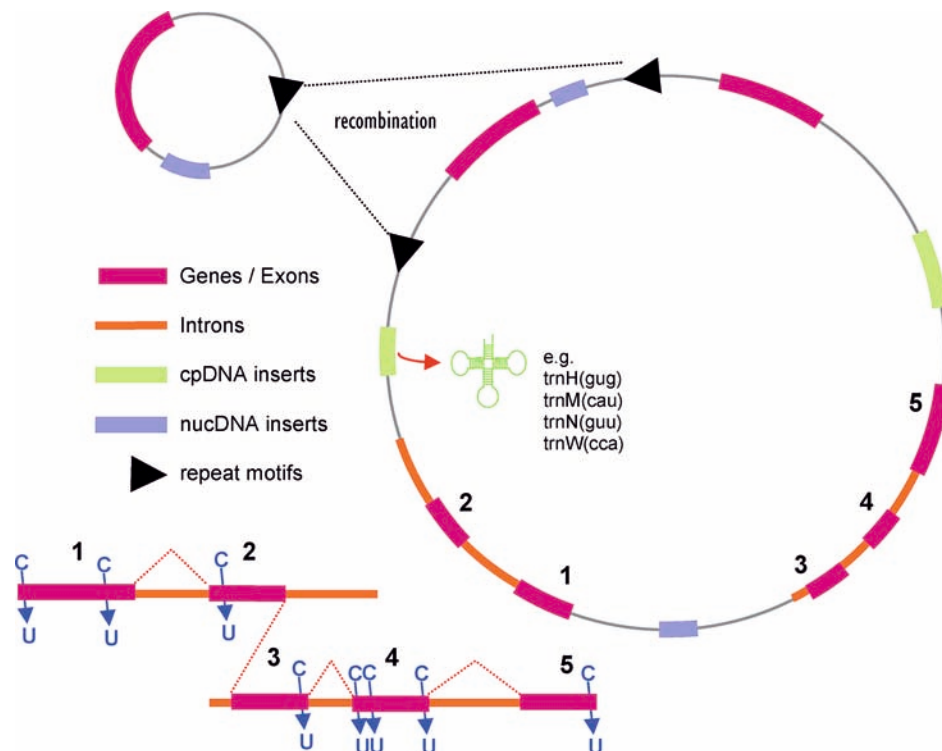
The mtDNA of the liverwort *Marchantia polymorpha* was the first embryophyte chondriome sequence to be completely determined, later followed by those of angiosperms—the widely investigated model species *Arabidopsis thaliana* with a chondriome of 367 kbp (Unsel'd et al. 1997), sugar beet, *Beta vulgaris*, with an mtDNA of 369 kbp (Kubo et al. 2000) and rice, *Oryza sativa*, with a mitochondrial genome of 491 kbp (Notsu et al. 2002). The most recently completed chondriome sequence of rapeseed (*Brassica napus*) has a size of “only” 222 kbp, which immediately illustrates the puzzling increase in mtDNA size in *Arabidopsis*, another member of the family Brassicaceae and thus a closely related species (Handa 2003). The completely determined land plant chondriome sequences are now ideally complemented by those of the charophyte alga *Chaetosphaeridium globosum* of the order Coleochaetales (Turmel et al. 2002a) and *Chara vulgaris* (Turmel et al. 2003), a member of the Charales. Together with the land plants, the charophyte algae are placed in the clade Streptophyta. Common features of the *Chara* and *Marchantia* mtDNA sequence

have strongly confirmed the placement of the Charales as an extant phylogenetic sister group to the land plants.

The circular mtDNAs of *Chaetosphaeridium* and *Chara* are 56 kbp and 67 kbp, respectively, in size. These chondriome sequences clearly show that most idiosyncratic features of land plant chondriomes—size increase, high recombinational activity, gain of several introns, high recombinational activity, gain of several introns and *trans*-splicing, RNA editing and foreign DNA insertions (Fig. 1)—are derived features acquired with or after the establishment of the land plant lineage. The transition to the land plant life form extended chondriome size by a factor of three in the liverwort *Marchantia*, without a gain of genes or apparent functionality, but mainly through the increase of spacer DNA between genes.

Although not the subject of this review, it must be kept in mind that some independent, peculiar trends of mitochondrial genome evolution do exist in the chlorophyte algal lineages, not closely related to land plants. Ancestral, circular and gene-rich chondriomes are found in the genera *Prototheca* (Wolff et al. 1994) and *Nephroselmis* (Turmel et al. 1999), whereas a circular, yet highly gene-reduced chondriome is found in *Pedinomonas* (Turmel et al. 1999; Wolff et al. 1994). *Chlorogonium* (Kroymann and Zetsche 1998) also has a circular chondriome but carries fragmented ribosomal RNA (rRNA) genes that are typical for mtDNA in the genus *Chlamydomonas* (Nedelcu 1997), from which both circular and linear chondriomes are known. Fragmented rRNA genes are also found in *Scenedesmus* mtDNA, which additionally exhibits a deviant genetic code (Nedelcu et al. 2000). A compact circular chondriome of 45 kbp also exists in *Mesostigma*, an early branching lineage in algal phylog-

Fig. 1 Typical idiosyncrasies in an angiosperm mitochondrial genome. Recombination commonly fragments the circularly mapping complete genomic entity of 200–2,000 kbp (the “master-circle”) into multipartite structures with sub-genomic circles. Some mitochondrial genes are continuous, but many others carry introns. Some of the organelle group II introns are disrupted, resulting in *trans*-arrangements of genes *nad1*, *nad2* and *nad5*, which require *trans*-splicing for maturation. Virtually all plant mitochondrial transcripts are subject to RNA editing, which convert cytidines to uridines to reconstitute conserved codon identities. The origin of some intergenic sequences can be identified, given their homologies with nuclear DNA (*nucDNA*) or cpDNA. The latter, when expressed, can give rise to functional transfer RNA (tRNA) in angiosperm mitochondria, replacing their native counterparts



eny; and in this alga two group II introns, so far without orthologues in other species, have attained a *trans*-splicing configuration (Turmel et al. 2002b).

With its evolutionary plasticity in the land plant lineage, mtDNA stands in contrast to the other organelle genome in the plant cell, that of the chloroplast (Palmer 1990). The chloroplast genome (the plastome) with its circular structure encompassing a large and a small single copy region separated by a pair of inverted repeats (IR) is structurally conserved, even in the ancient algal lineages such as *Mesosigma* (Lemieux et al. 2000). Most gene clusters and orders are conserved and expansions and contractions of the IR are the most significant differentiating feature of the plastome, which was fixed in structure and appearance early in plant evolution. Extensive structural rearrangements, such as in subclover (Milligan et al. 1989) or in *Trachelium* (Cosner et al. 1997), appear to be rare in plants. However, unique structural changes in the plastome can very well convey significant phylogenetic information. A 30-kbp genomic cpDNA inversion documents the earliest dichotomy in vascular plant phylogeny between lycophytes and other tracheophytes (Raubeson and Jansen 1992); and the identification of two introns common in embryophytes and charophyte algae identified the streptophyte clade over 10 years ago (Manhart and Palmer 1990).

Hence, 500×10^6 years of land plant evolution rarely resulted in significant alterations to the chloroplast genome, but significant changes are found in the mtDNA, even on very narrow taxonomic scales. Whereas the plant mitochondrial genome itself is structurally dynamic in evolution, the sequence drift in coding regions is generally very low in plant mtDNA (Wolfe et al. 1987) and this can offer advantages for phylogenetic analysis of deeper branches in phylogeny. In the marchantiid subclass of liverworts, this slow sequence drift appears to be yet further retarded (Beckert et al. 1999), but the opposite trends exist in other lineages. Mitochondrial sequence evolution is strikingly accelerated in the angiosperm genera *Pelargonium* and *Plantago* (Palmer et al. 2000) and the fern genus *Lygodium* (Vangerow et al. 1999). Thus, contradictory trends exist in embryophyte mtDNAs: whereas gene sequences mostly evolve slowly, chondriome structures evolve quickly and gene orders are soon rearranged (Palmer and Herbon 1988). Genes are continuously transferred from the mtDNA of angiosperms to the nuclear genome in recent evolutionary times, while at the same time their chondriomes readily accept DNA from the nucleus or the chloroplast genome for incorporation.

The phylogenetic perspective

Several land plant groups circumscribed by conventional taxonomy are now well corroborated as monophyletic clades by the overwhelming majority of molecular phylogenies (Pryer et al. 2002; Qiu and Palmer 1999): the flowering plants (angiosperms), the seed plants

(spermatophytes), the vascular plants (tracheophytes), the lycophytes, the liverworts, the mosses and the hornworts (Fig. 2). Other superior land plant clades suggested more recently are now likewise clearly defined and resolve the issues of “fern allies” (Kolukisaoglu et al. 1995; Raubeson and Jansen 1992). The “moniliformopses” (Kenrick and Crane 1997a; Pryer et al. 2001) include the true ferns, the horsetails (*Equisetum*) and the whisk ferns (Psilotales), but exclude the lycophytes (Fig. 2). Other groups as yet retain a less clear status, such as the gymnosperms. The anthophyte hypothesis assuming the proximity of Gnetopsida to angiosperms (Crepet 2000; Donoghue and Doyle 2000; Rydin et al. 2002) is not recovered by molecular analyses, which instead frequently identify the two sister groups Ginkgopsida–Cycadopsida and Coniferopsida–Gnetopsida in monophyletic gymnosperms (Bowe et al. 2000; Chaw et al. 1997, 2000; Goremykin et al. 1996; Malek et al. 1996). While the first sister group relation is recovered by most molecular phylogenies, the peculiar evolution of the latter two classes, long branch attraction problems and taxon sampling seem to affect molecular phylogenetic analyses in that case in a particularly strong manner (Magallon and Sanderson 2002; Rydin et al. 2002). Hence, it may as yet be more reasonable to consider the gymnosperm classes as unresolved at the base of seed plants as a whole (Pryer et al. 2002).

Most importantly, the deepest branches of the plant phylogenetic backbone are unresolved. The origin of land plants (embryophytes) supposedly dates to Ordovician times (Graham et al. 2000; Kenrick and Crane 1997b), but the notorious lack of macrofossils has so far not allowed a consensus on the first land plants (Wellman and Gray 2000). Molecular data make it convincingly clear that, among extant algae, the Charales share a most recent common ancestor with land plants (embryophytes) and that this joint group is sister to the Coleochaetales algae (Karol et al. 2001; Malek et al. 1996; Turmel et al. 2003). The division bryophytes, including the base-most land plant classes liverworts, mosses and hornworts, is most likely not a monophyletic group. The three earliest bifurcations of land plant evolution seem to separate the three bryophyte clades, possibly placing only one of them as sister group to the vascular plants. As discussed in the next section, molecular support for the idea of liverworts as the base-most group of land plants was found in the identification of three mitochondrial introns present in all land plant clades except the liverworts (Qiu et al. 1998). However, alternative views place hornworts as the base-most land plant clade, based both on molecular (Nickrent et al. 2000) and morphological insights (Renzaglia et al. 2000).

Evolution of gene content: nuclear transfer and replacement

The small “core” set of 13 genes in animal mitochondria coding for protein subunits of respiratory chain

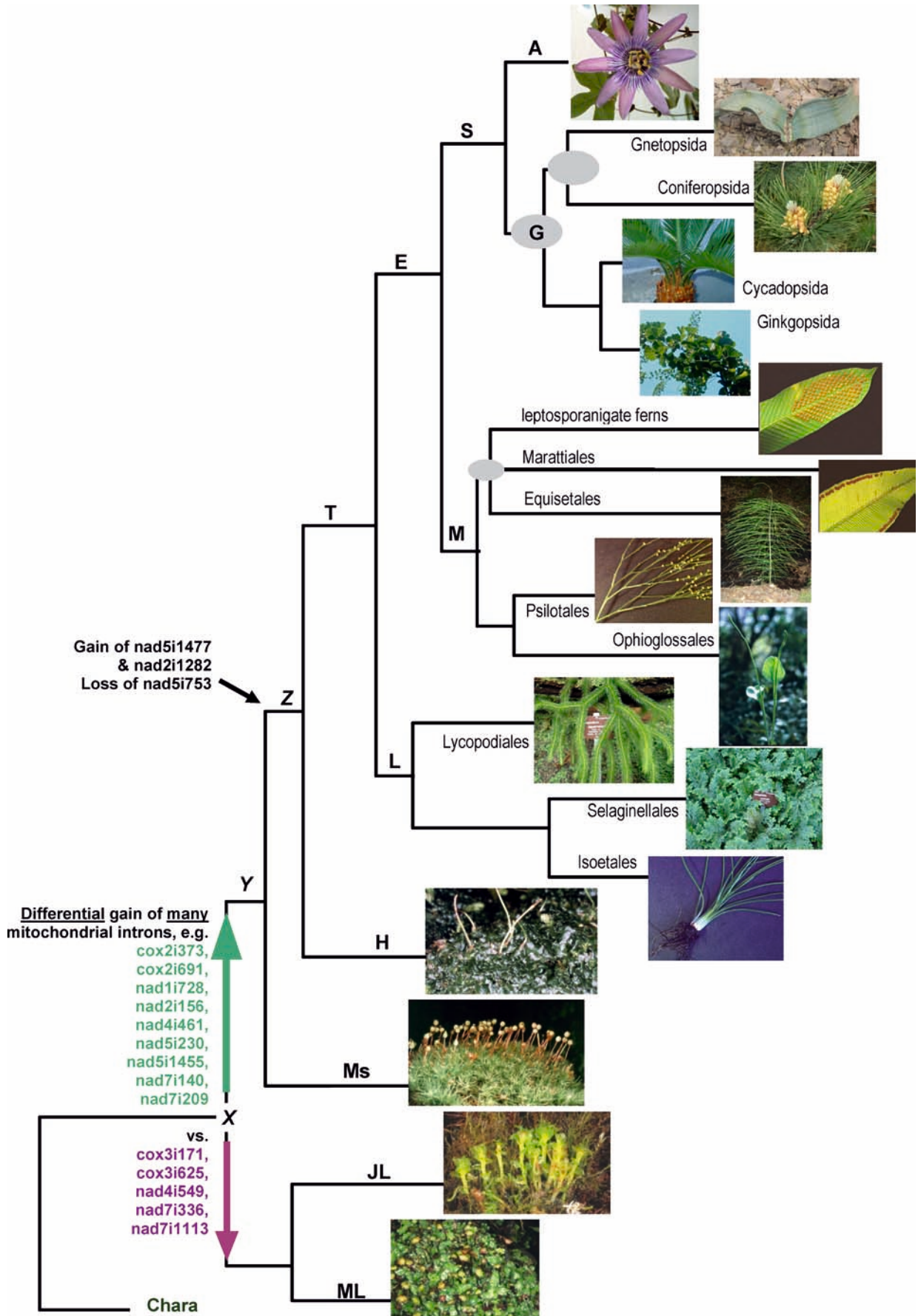


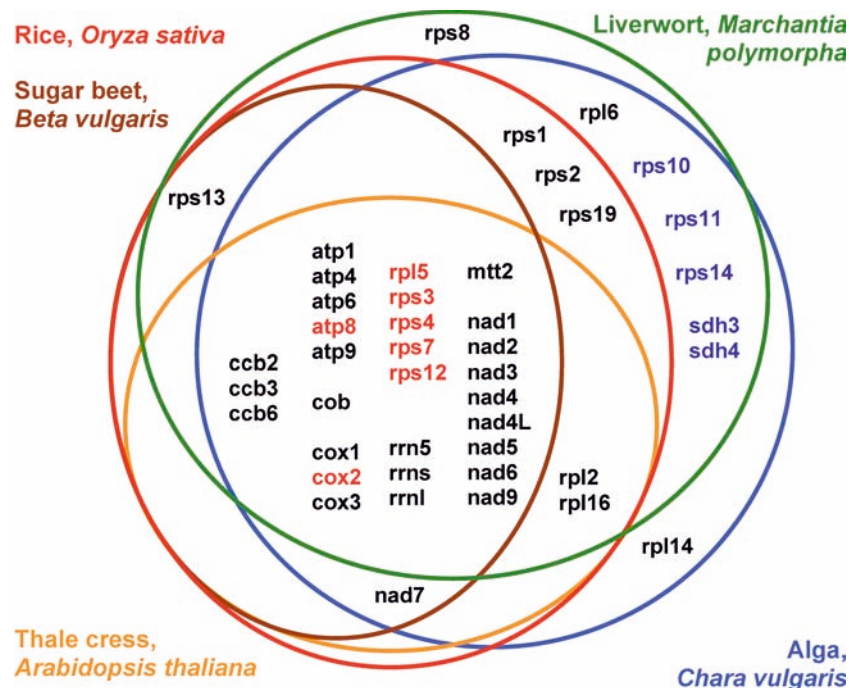


Fig. 2 Several clades in the land plant phylogeny identified and/or confirmed by molecular data as monophyletic groups are: *A* angiosperms, *S* seed plants, *M* moniliformopses, *E* euphyllophytes, *L* lycophytes, *T* tracheophytes. The monophyly of gymnosperms (*G*) as a whole and that of a clade comprising two of its classes, Gnetopsida and Coniferopsida, is somewhat less well supported. Likewise, more phylogenetic resolution is needed for a clade comprising the eusporangiate ferns of the order Marattiales, the horsetails (Equisetales) and the leptosporangiate ferns vs a clade comprising the Ophioglossales and whisk ferns (Psilotales). The tree shown is topologically consistent, albeit not always statistically supported. Several phylogenetic analyses of organelle genes (unpublished observations) place hornworts (*H*) as sister group to the tracheophytes (node *Z*), mosses (*Ms*) as a sister group to the joint clade (*Y*) and confirm marchantiid (*ML*) and jungermanniid (*JL*) liverworts jointly as a sister group to all other embryophytes (node *X*)

Fig. 3 Gene content in five completely sequenced streptophyte chondriomes. Figure design is adapted from Gray et al. 1999. Intron-encoded open reading frames (orfs) and tRNA genes, which are frequently lost and replaced by chloroplast counterparts in angiosperms, are not considered but documented in the supplementary table under <http://www.izmb.de/knoop/chondriomes.htm>. Genes encode subunits of respiratory chain complexes I (nad), II (sdh), III (cob), IV (cox) and V (atp), proteins of the small and large ribosomal subunit (rps, rpl) and proteins for cytochrome c biogenesis (ccb). The latter have also been referred to as yej, ccm or hel genes, respectively. The ccb6 (= yejR) gene is split into separate orfs in plant mitochondria (see text; Fig. 4). Designations atp4 and atp8 are new nomenclature proposals (Burger et al. 2003b; Gray et al. 1998; Heazlewood et al. 2003; Sabar et al. 2003) for genes that until recently were referred to as orf25 (or ymf39) and orfB, respectively, in the literature and in database entries. The mtt2 gene is the homologue of the secY-independent membrane transport protein, which was referred to as orfX, tatC or ymf16, respectively. Genes that are not present in the completely sequenced angiosperm chondriomes but known to be present in other angiosperm mtDNAs are shown in blue. Genes of the core set for which nuclear transfer in other angiosperms is demonstrated are given in red

complexes I, III, IV and V (atp6, atp9, cob, cox1, cox2, cox3, nad1, nad2, nad3, nad4, nad4L, nad5, nad6) is extended by extra genes for additional subunits of complex I (nad genes) and complex V (atp genes) and genes encoding ribosomal proteins (rpl, rps genes), subunits of complex II (sdh genes) and proteins involved in cytochrome c biogenesis (ccb genes) in plants. A comprehensive table of genes and introns in the completely sequenced land plant chondriomes and other supplementary information, intended to be continuously updated, has been assembled for the purpose of this review and is available on the internet at <http://izmb.de/knoop/chondriomes>. An overview of the protein-encoding genes in five completely sequenced chondriomes of plants is shown in Fig. 3. The liverwort *Marchantia* has the most gene-rich (and ancestral) embryophyte chondriome known so far, with the exception of nad7, which is degenerated to a pseudogene here. The unknown land plant ancestor may also have carried rpl14 on its chondriome, which still exists in the mtDNA of the alga *Chara*. The rice chondriome is the most gene-rich angiosperm chondriome so far known. The two most closely related plants, the Brassicaceae *Arabidopsis thaliana* and *Brassica napus*, differ with respect to the gene for rps14, which is a pseudogene in the *Arabidopsis* mtDNA but a functional gene in the mitochondrion of *B. napus*. At the same time, the *Arabidopsis* mtDNA is significantly larger (367 kbp vs 222 kbp), proving that intergenic DNA without obvious function can accumulate in short time intervals in a flowering plant chondriome.

When a gene is missing in the chondriome of one plant species but present in another, one can reasonably assume that a copy has been transferred to the nucleus, gained function through appropriate regulatory



elements and a proper amino-terminal targeting sequence and that its product is imported into mitochondria after synthesis in the cytosol (Adams and Palmer 2003). The as yet best documented example of such a “gene transfer in progress” in land plant evolution is the *cox2* gene (Nugent and Palmer 1991). The different steps of gene-copying to the nucleus, the activation of the nuclear copy and the degeneration of its mitochondrial counterpart can be traced for the *cox2* genes of leguminous plants (Adams et al. 1999; Daley et al. 2002). However, being a respiratory chain complex subunit belonging to the core set of genes, which is also encoded in mtDNA of animals, *cox2* is clearly an exception. In fact, the *atp8* gene of *Allium* (onion) is the only other example of a gene encoding a protein in the respiratory chain complexes I, III, IV or V that has been identified as missing from the mtDNA in a wide survey of angiosperms (Adams et al. 2002b). However, gene transfer events are significantly more often observed for the *sdh* genes encoding complex II subunits and those encoding ribosomal proteins (Adams et al. 2002b). A prominent example is the *rps10* gene encoding ribosomal protein S10, where many gene transfers took place independently (Adams et al. 2000; Knoop et al. 1995). Most of the ribosomal protein genes were lost from plant chondriomes independently in a patchy manner; and only the *rps2* and *rps12* genes were lost a few times deep in angiosperm phylogeny. Gene transfer frequencies to the nucleus are highly skewed, with some taking place frequently and independently (e.g. *rps7*, lost at least 42 times among 280 sampled flowering plants) and others much more rarely, such as *rps12* (six times; Adams and Palmer 2003). All ribosomal protein genes were lost from the chondriome in *Lachnocaulon* (bog buttons), but none in the basal-most angiosperms (Qiu et al. 2001; Zanis et al. 2002); and the latter observation may indicate that the set of mitochondrial genes is conserved in phylogeny through vascular plant evolution up to the root of angiosperms. Obviously though, gene transfer activity could be increased independently in the other major land plant clades (Fig. 2). Taxonomically extensive studies like those by Adams and colleagues on angiosperms (Adams et al. 2002b; Adams and Palmer 2003) are as yet not available for older land plant groups—the gymnosperms, the pteridophytes (i.e. lycophytes, moniliformopses) or the bryophytes. So far, the *nad7* gene in the nucleus of *Marchantia* is the only example of a functional gene transfer from the mitochondrion documented in the early embryophytes (Kobayashi et al. 1997). The degeneration of *nad7* to a mitochondrial pseudo-gene by a gain of stop codons is not restricted to *Marchantia* among the liverworts. Stop codons are gained independently at different positions in *nad7* in other liverwort genera (unpublished observations). However, the *nad7* gene is functional in moss mitochondria (Hashimoto and Sato 2001; Pruchner et al. 2001).

Mitochondrial re-targeting peptide sequences can be acquired in different ways: duplication of existing

pre-sequences in the nucleus (e.g. mitochondrial heat-shock proteins), altered use of other coding sequences, exon shuffling, insertion into an intron for a mitochondrially targeted protein gene and alternative splicing or de novo acquisition of a targeting sequence at the insertion site (Adams et al. 2000, 2001b; Figueroa et al. 1999; Kadowaki et al. 1996; Sanchez et al. 1996; Sandoval et al. 2004; Wischmann and Schuster 1995). A further alternative is that no cleavable targeting peptide sequence may be required at all when the nuclear copy fortuitously contains targeting information, as in the case of *rps10* (Adams et al. 2000; Kubo et al. 2003).

It must be noted that an observed gene loss from the chondriome does not necessarily implicate that the mitochondrial gene was functionally transferred to the nucleus. It was suggested that an extension of the *rps19* gene gained after transfer to the nucleus could functionally replace the missing *rps13* gene in *Arabidopsis* mitochondria (Sanchez et al. 1996). Instead, a diverged nuclear copy of *rps13* of chloroplast origin was found to replace the function of its mitochondrial counterpart (Adams et al. 2002a). The *rps8* gene, which is not identified in angiosperm mitochondria (Fig. 3), is functionally replaced by the adapted copy of a nuclear gene originally coding for *rps15A* in cytosolic ribosomes (Adams et al. 2002a).

The most common form of functional gene replacement in the plant mitochondrion involves tRNA genes derived from the plastome. Frequent insertion of promiscuous cpDNA sequences into plant chondriomes is common in flowering plants. Not all of these transferred sequences are rendered useless as intergenic DNA in their new location—some tRNA encoding genes of chloroplast origin functionally replace the original native mitochondrial tRNA species in the plant mitochondrial genome (Marechal-Drouard et al. 1990). For example, the plastid genes *trnH(gug)*, *trnM(cau)*, *trnN(guu)* and *trnW(cca)* are present in all completely sequenced angiosperm chondriomes (Fig. 1), whereas the native mitochondrial genes are (still) present in the liverwort chondriome. However, a chloroplast-derived gene known to function in one plant species may be inactive in another, indicating the import of a tRNA encoded in the nucleus (Duchene and Marechal-Drouard 2001; Kumar et al. 1996).

The presence of *ccb* genes encoding proteins participating in cytochrome c biogenesis in embryophytes and *Chara*, but their absence in other green algae, must be explained by several independent losses (at least in *Mesostigma*, *Chaetosphaeridium* and an early chlorophyte ancestor). An alternative gain of *ccb* genes in the common ancestor of *Chara* and the embryophytes involving horizontal gene transfer may sound far-fetched but would evidently be a more parsimonious explanation. Actually, two recent reports claim horizontal gene transfer of mtDNA sequences into several angiosperm genera: *atp1* into *Amborella*, *rps2* into *Actinidia* and *rps11* into *Betula* and *Sanguinaria* (Bergthorsson et al. 2003) and an angiosperm-like intron in

nad1 into the gymnosperm *Gnetum* (Won and Renner 2003). If true, such events would naturally add a highly complicating factor to gain, loss and transfer scenarios. So far, the reported observations are restricted to individual mitochondrial loci in the suspected acceptor plants and a more extensive characterization of their chondriomes will be very helpful to further substantiate the conclusions.

Whereas the functional organelle-to-nucleus transfer of a gene can proceed via RNA and cDNA (Nugent and Palmer 1991), inter-genomic sequence transfer can also proceed via DNA and is not necessarily functional or successful. Even at pre-genomic times, it was possible to identify mtDNA, such as intron fragments, in the nuclear DNA of angiosperm species (Blanchard and Schmidt 1995; Knoop and Brennicke 1991, 1994). Sequencing the *Arabidopsis thaliana* nuclear genome revealed big stretches of mtDNA (Stupar et al. 2001) and it is not surprising that organelle sequences that are derived from DNA, not RNA, show up in the now-completed genome of rice. A surprisingly frequent random transfer of cpDNA to the nucleus that can even be traced within generation time is now experimentally documented in tobacco (Huang et al. 2003; Stegemann et al. 2003).

Plant mitochondrial introns

A recent proposal to designate organelle introns with the name of their host gene and a number indicating the respective upstream exon nucleotide of the homologous *Marchantia polymorpha* orf (Dombrowska and Qiu 2004) is adapted here by adding the classifications into group I and group II introns, respectively. It is important to distinguish these two different types of “organelle” introns (with group II introns increasingly being identified also in eubacteria; Dai and Zimmerly 2002), which are classified according to their entirely different secondary structures and which have different modes of splicing and invading new loci. Both types are frequently referred to as self-splicing—a somewhat misleading simplification because autocatalytic activity is documented only for a minority of fungal and bacterial species, but not for the plant organelle introns; and protein cofactors are known to be essential in vivo in any case. Group I introns can carry orfs encoding homing endonucleases of the LAGLIDADG type which take part in invading intron-less loci; and some group II introns can carry maturase-type orfs with reverse transcriptase sequence motifs. A new level of complexity to intron evolution and the modes of intron propagation comes from the discovery of group II introns with LAGLIDADG type orfs (Toor and Zimmerly 2002).

Table 1 attempts to summarize the knowledge on intron insertion positions in those plant mitochondrial genes for which there is at least some taxonomic sampling in non-seed plants other than the liverwort *Marchantia* alone, i.e. for moss, hornwort or

pteridophyte species. All introns in the completely sequenced streptophyte chondriomes, together with information on sizes, similarity groupings and intron-borne orfs are compiled in the complementary table at <http://izmb.de/knoop/chondriomes.htm>.

Comparing the chondriome sequences of the flowering plants with that of the liverwort *Marchantia*, one observation is immediately evident: of 25 known group II intron sites in angiosperm mitochondria and an equal number in *Marchantia*, only a single intron, nad2i709, is conserved in position (Table 1). This contrasts with the situation in chloroplast genomes, where most introns (a total of 17) are highly conserved among land plants of all groups and were obviously gained “in the water”, even before the separation of the Coleochaetales alga *Chaetosphaeridium* and the land plant lineage (Turmel et al. 2002a). Only four other plastome introns may have been gained after this separation “on land”.

The *Marchantia* chondriome carries group I introns, one in nad5 and six in the cox1 gene—none are present in the completely sequenced angiosperm chondriomes. Group I introns were unknown in angiosperm mitochondrial genes until the discovery of cox1i729 in *Peperomia* (Vaughn et al. 1995). This intron turned out to be a highly invasive species with a patchy distribution among angiosperms (Cho et al. 1998). Strikingly, the cox1 gene generally appears to be a prime target for group I intron invasion in algae and *Marchantia* (Table 1), with orthologous introns being observed in identical positions in the mtDNAs of fungi and chlorophyte algae (Ohta et al. 1993; Ohya et al. 1993; Turmel et al. 2002b; Wolff et al. 1993). A horizontal inter-species transfer is thus very likely for most of them. The nad5 group I intron nad5i753 (Table 1) even has a positional homologue in the mtDNA of *Monosiga brevicollis* (Burger et al. 2003a), very likely representing a very early branch of the animal lineage. However, this intron is present (so far without exception) in all mosses and liverworts investigated (Beckert et al. 1999), indicating a highly stable vertical transmission once land plants began to evolve. It may have been lost in a common ancestor of hornworts and tracheophytes under the phylogenetic scenario shown in Fig. 2.

The single mitochondrial intron conserved in position in angiosperms and *Marchantia*, nad2i709 (Table 1), appeared as an attractive candidate for an intron potentially conserved in all land plant clades but surprisingly turned out to be absent in all mosses (Beckert et al. 2001). In contrast, moss nad2 genes (again, so far without any exception) exclusively carry the “angiosperm-type” intron nad2i156. Similar observations are made for the nad4 and nad7 genes (Table 1), where mosses likewise only feature angiosperm-type introns nad4i461, nad7i140 and nad7i209 (Pruchner et al. 2001). Group I intron nad5i753 found to be present both in all mosses and in all liverworts therefore turns out to be exceptional; and it is actually so far the only documented example of an intron shared between mosses and liverworts. Recently extended investigations of the nad5

Table 1 Variable intron occurrence as so far documented in mitochondrial genes for which there is at least some taxon sampling exceeding *Marchantia*, algae and angiosperms, i.e. for mosses, hornworts, lycophytes or moniliformopses: *cox1* (Bowe et al. 2000; Sperwhitis et al. 1996), *cox2* (Qiu et al. 1998; Sperwhitis et al. 1994), *cox3* (Malek et al. 1996), *nad1* (Dombrowska and Qiu 2004), *nad2* (Beckert et al. 2001; Pruchner et al. 2002), *nad5* (Beckert et al. 1999; Vangerow et al. 1999), *nad4* and *nad7* (Pruchner et al. 2001). *Plus/*

minus signs Presence/absence of a given intron documented for at least one member of the respective clade, *E*, *M* endonuclease-type and maturase-type intron-borne orfs in group I and group II introns, respectively, *T* disruption into a *trans*-arrangement, *brackets* within-clade variability, *question marks* a lack of information for the clade or gene region, respectively, *intron names in italics* orthologues in non-streptophyte groups (fungi, chlorophytes)

Intron	<i>Chara</i>	Liverworts	Mosses	Hornworts	Lycophytes	Moniliformopses	Spermatophytes
<i>cox1</i>							
<i>cox1i44g2</i>	–	+M	?	?	?	?	?
<i>cox1i178g2</i>	–	+M	?	?	?	–	–
<i>cox1i211g2</i>	+M	–	?	?	?	–	–
<i>cox1i375g1</i>	–	+	?	?	?	–	–
<i>cox1i395g1</i>	–	+E	?	?	?	–	–
<i>cox1i511g2</i>	–	+M	?	?	?	–	–
<i>cox1i624g1</i>	–	+	?	?	–	–	–
<i>cox1i729g1</i>	+	(+)	–	–	–	–	(+)
<i>cox1i740g1</i>	+	–	–	–	–	–	–
<i>cox1i835g2</i>	+M	–	–	–	–	–	–
<i>cox1i876g1</i>	+	–	–	–	–	–	–
<i>cox1i909g1</i>	+	–	–	–	–	–	–
<i>cox1i1116g1</i>	–	+E	?	?	–	–	–
<i>cox1i1303g1</i>	–	+	?	?	?	–	–
<i>cox2</i>							
<i>cox2i97g2</i>	–	+	?	?	?	–	–
<i>cox2i104g2</i>	+	–	?	?	?	–	–
<i>cox2i250g2</i>	–	+	?	?	?	–	–
<i>cox2i373g2</i>	–	–	+	+	(+)	(+)	(+)
<i>cox2i691g2</i>	–	–	+	(+)	(+)	–	(+)
<i>cox3</i>							
<i>cox3i171g2</i>	–	+	–	–	(+)	–	–
<i>cox3i625g2</i>	–	+	–	–	–	–	–
<i>nad1</i>							
<i>nad1i258g2</i>	–	–	–	–	–	+	–
<i>nad1i285g2</i>	–	–	+	+	–	–	–
<i>nad1i348g2</i>	–	–	–	+	–	–	–
<i>nad1i394g2</i>	–	–	–	–	+	(+)	+T
<i>nad1i477g2</i>	–	–	–	–	(+)	(+)	(+)
<i>nad1i669g2</i>	–	–	–	–	(+)	(+)	+T
<i>nad1i728g2</i>	–	–	+(M)	+M	(+M)	+M	(+) M (T)
<i>nad2</i>							
<i>nad2i156g2</i>	–	–	+	–	+	?	+
<i>nad2i542g2</i>	–	–	–	–	(+)	+	+T
<i>nad2i709g2</i>	–	+	–	+	(+)	?	+
<i>nad2i830g2</i>	–	–	–	–	+	?	–
<i>nad2i1282g2</i>	–	–	–	+	–	?	+
<i>nad4</i>							
<i>nad4i461g2</i>	–	–	+	+	?	?	+
<i>nad4i549g2</i>	–	+	–	–	?	?	–
<i>nad4i976g2</i>	+	–	–	+	?	?	(+)
<i>nad4i1399g2</i>	–	–	–	–	?	?	(+)
<i>nad5</i>							
<i>nad5i230g2</i>	–	–	+	+	?	+	+
<i>nad5i391g2</i>	–	–	–	–	(+)	–	–
<i>nad5i753g1</i>	–	+	+	–	–	–	–
<i>nad5i881g2</i>	–	–	–	(+)	–	–	–
<i>nad5i1242g2</i>	–	–	–	–	+	(+)	–
<i>nad5i1455g2</i>	–	–	+	+	+	+	+T
<i>nad5i1477g2</i>	–	–	–	+	+	(+)	+T
<i>nad5i1872g2</i>	–	–	–	–	+	+	+
<i>nad7</i>							
<i>nad7i140g2</i>	–	–	+	?	?	+	+
<i>nad7i209g2</i>	–	–	+	?	?	+	+
<i>nad7i336g2</i>	–	+	–	?	?	–	–
<i>nad7i676g2</i>	–	–	–	?	?	?	(+)
<i>nad7i917g2</i>	–	–	–	?	?	?	+
<i>nad7i1113g2</i>	–	+	–	?	?	?	–

gene, which is rich in angiosperm introns, find this observation extended: angiosperm type introns nad5i230 and nad5i1455 are absent in liverworts but present in mosses, the latter being *trans*-disrupted in flowering plants (M. Groth-Malonek, D. Pruchter, F. Grewe, V. Knoop, submitted).

Only some mitochondrial group II introns are of narrower taxonomic distribution and confirm classic taxonomic groups: nad2i830 in lycophytes (Pruchner et al. 2002), nad1i348 in hornworts and nad1i258 in moniliformopses (Dombrovskaya and Qiu 2004). Others link two clades, which are, however, not suggested as sister groups in phylogenetic analyses (Fig. 2): nad5i1242 in lycophytes and moniliformopses (Vangerow et al. 1999), nad5i285 in mosses and hornworts and nad5i753 in mosses and liverworts. One deep gain and one deep loss along the backbone of plant phylogeny would need to be assumed for those. Introns that so far have been identified in single genera only are rare: nad5i391 in the lycophyte *Huperzia* (Vangerow et al. 1999) and intron nad5i881 in a hornwort (Beckert et al. 1999). The *Huperzia* intron nad5i391 is particularly interesting as it is closely related to the common “pteridophyte” (moniliformopses, lycophytes) intron nad5i1242 and is thus a first example of what may be a within-gene group II intron-copying mechanism.

None of the known angiosperm group II introns is a characteristic “monocot” or “dicot” intron but introns can be lost independently in genera of both angiosperm groups. Variability in intron occurrence among angiosperms that can most consistently be explained by independent losses is reported for cox2i373 and cox2i691 (Kudla et al. 2002; Qiu et al. 1998), nad4i976 and nad4i1399 (Geiss et al. 1994; Itchoda et al. 2002), nad7i676 (Pla et al. 1995), nad1i477 (Bakker et al. 2000) and rps3i74 (Kubo et al. 2000). An extensive study investigating the loss of cox2i373 documents many independent losses among angiosperms (Joly et al. 2001). Interestingly, this intron has a positional homologue in a strain of the yeast *Schizosaccharomyces pombe* (Schafer et al. 1998) which may, apart from independent losses, also hint at independent gains on wider taxonomic scales. The independent loss of group II introns is also evident for non-flowering vascular plants: introns cox2i373 and cox2i691 are lost in some pteridophytes and gymnosperms (Qiu et al. 1998), nad5i1242 is lost in *Equisetum* and *Ophioglossum* (Vangerow et al. 1999), nad1i394 in *Ophioglossum*, nad1i477 in *Isoetes*, *Asplenium*, *Equisetum*, among gymnosperms in *Welwitschia* and in a conifer clade and nad1i669 is lost in *Isoetes* and *Asplenium* (Dombrovskaya and Qiu 2004; Gugerli et al. 2001).

Taking the *Chara* chondriome as a reference, it appears that invasion of mitochondrial genes with introns largely took place during establishment of the embryophytes. *Chara* shares only two introns with *Marchantia*: nad3i140 (absent in angiosperms) and cox1i729—the “rampant invader” found in some angiosperm cox1 genes. Likewise, *Chara* shares two

introns with angiosperms, absent in *Marchantia*: nad4i976 and rps3i74.

The earliest gains of new mitochondrial introns coinciding with the emergence of land plants after separation from the *Chara* lineage could be nad5i753 and nad2i709 (Fig. 2). After that, introns may have invaded the chondriomes in a differential manner after an early split of land plants into liverworts and non-liverwort plants. Certainly, intron transposition and differential loss scenarios are an alternative, although the less parsimonious explanation. Given that nine mitochondrial group II introns in *Marchantia* encode maturase orfs, mobility may have been significantly higher in early land plant evolution. It may have ceased somewhere in the evolution of the angiosperm lineage, possibly reflected in the fact that the maturase-like orf in nad1i728 is the only one of its kind in angiosperms. It is noteworthy that, while intron nad1i728 is retained in mosses, the maturase orf is lost from the intron in derived moss genera (Dombrovskaya and Qiu 2004).

There is no doubt that some land plant lineages will yield further surprises with respect to the occurrence of mitochondrial introns, possibly comparable with those of suddenly accelerated mitochondrial sequence evolution in single angiosperm genera (Adams et al. 1998). For example, very small mitochondrial group II introns have been found in the quillwort *Isoetes* (Malek and Knoop 1998) and new mitochondrial sequences of this genus reveal additional introns so far without orthologues in other species (unpublished observations). At least one group II intron is a strong candidate for independent gains over a large evolutionary distance: cox3i171, which is present in liverworts and the lycophyte *Lycopodium* (Hiesel et al. 1994b), and the introns nad4i976 and rps3i74 shared between *Chara* and some angiosperms.

Given (1) that some mitochondrial intron positions are shared between plants and fungi, (2) that mycorrhizal fungi were already present on earth in Ordovician times when the evolution of embryophytes began (Redecker et al. 2000), (3) that not only vascular plants but also bryophytes can be partners in such symbioses (Read et al. 2000) and (4) that very diverse non-mycorrhizal fungi also inhabit aboveground organs of all plants as endophytes (Gamboa et al. 2003; Ligrone et al. 1993), it is conceivable that different fungi may have acted differentially as donors of different mitochondrial introns in the plant lineages.

Recombinational activity disrupting gene continuity

Mono-circular chondriomes, such as in *Brassica hirta* (Palmer and Herbon 1987) and *Marchantia*, represent the ancestral state of mitochondrial genome organization in the plant lineage, but are probably the exception rather than the rule in angiosperms. Active recombination via a repeated sequence motif producing two sub-genomic circles is evident for other *Brassica* species

(Palmer and Herbon 1988; Palmer and Shields 1984); and the completed *B. napus* chondriome sequence has identified this sequence as a directly repeated motif of 2,427 bp (Handa 2003). This sequence repeat is not related to those found in *Arabidopsis*, where two pairs of repeated sequences produce a more complex population of sub-genomic molecules (Klein et al. 1994). Maize (*Zea mays*) is an important model for the significantly higher chondriome complexity through coexisting sub-genomic molecules found in other angiosperms (Andre and Walbot 1995; Small et al. 1989). Additionally, the plant mtDNA complexity is yet further enhanced in some plant species by the presence of plasmids, for example a linear 11-kbp DNA molecule with terminal IR sequences in *B. napus* (Handa et al. 2002), which does not share significant homology with the now completed main chondriome sequence (Handa 2003).

Whereas the active homologous recombination via larger sequence repeats produces coexisting sub-genomic molecules in a plant mitochondrion, recombinations via shorter sequence motifs appear to be the major force rearranging plant chondriomes on evolutionary time-scales. Notably, even a supposedly simple mono-circular mtDNA, such as in the liverwort *Marchantia*, may in fact be physically represented (also) by circularly permuted, linear molecules (Bendich 1993; Oldenburg and Bendich 2001). Although there is no evidence for active recombination leading to coexisting molecules of *Marchantia* mtDNA, it should be noted that duplicated sequence stretches, potential substrates of homologous recombination, are present in the liverwort chondriome (Ohyama et al. 1993). These include duplication of a cob gene intron sequence contributing to a size increase in the nad5-nad4 spacer region for which, however, homologous arrangements exist in jungermanniid liverworts (unpublished observations), indicating a recombination event on evolutionary time-scales.

Among angiosperms, a variation in mitochondrial genome arrangements is common in plants, even between isolates of the same species (Scotti et al. 2004; Ullrich et al. 1997). Mitochondrial dysfunctions in mutant lines of flowering plants, typified by cytoplasmic male sterility (Budar and Pelletier 2001), are frequently connected to such recombination events when they affect conserved orfs or their expression.

Comparing the *Marchantia* and *Chara* chondriomes, the conservation of some gene linkages is obvious (Turmel et al. 2003). The larger of these arrangements are the nad5-nad4-nad2- and the trnC-trnF-rps2-ccb2-ccb3-ccb6-trnQ gene orders and the nad7-rps10-rpl2-rps19-rps3-rpl16-[rpl14]-rpl5-rps14-(rps8)-rpl6-(rps13)-rps11 cluster, the latter of which includes the ancient bacterial rps10 operon, from which rps8 and rps13 are lost in *Chara*, whereas rpl14 is inserted at this position. Gene orders generally are not found conserved in angiosperm mitochondria, with the exception of the nad3-rps12 linkage and the (overlapping) rps3-rpl16 genes (Nakazono et al. 1995), preceded by rps19 in some species.

In some cases, the frequent recombinational activity not only affected intergenic regions but also disrupted the continuity of genes, either within introns or even in the coding region. On the gene expression level, one consequence of frequent recombination in land plant chondriomes is the necessity of messenger RNA (mRNA) maturation by *trans*-splicing in some plant mitochondrial genes. Three mitochondrial genes in angiosperms carry *trans*-splicing introns: nad1 (Chapdelaine and Bonen 1991; Wissinger et al. 1991), nad2 (Binder et al. 1992) and nad5 (Knoop et al. 1991; Pereira et al. 1991). In these cases, group II introns are disrupted in their continuous sequence and the 5' half of the affected intron follows the upstream exon in one genomic location, while the 3' half precedes the downstream exon in another genomic location. The disruptions of the introns in a large and variable domain IV of their secondary structures (Michel and Ferat 1995) do not affect maturation of the mRNAs, which have to be assembled from independent precursor transcripts. Therefore, this process was named *trans*-splicing. Among angiosperms, five *trans*-splicing events (nad1i394T, nad1i669T, nad2i542T, nad5i1455T, nad5i1477T) are highly conserved (Table 1). As shown for nad5i1477 in *Oenothera*, an intron can even be split twice (Knoop et al. 1997). None of these *trans*-splicing introns has an algal or liverwort homologue. Tracing the origins of *trans*-splicing introns in plant phylogeny (Fig. 2), it was possible to identify *cis*-arranged homologues and thus presumable ancestors in pteridophytes (Malek et al. 1997; Malek and Knoop 1998). In the case of nad5i1477T, a *cis*-arranged homologue was identified in the hornwort *Anthoceros*, demonstrating that gene invasion by ancestral introns could predate the bryophyte-vascular plant separation. Exploring the complete nad5 gene structure in mosses, we recently found that this intron, nad5i1477, is absent in mosses. However, mosses do show the presence of a *cis*-arranged counterpart for nad5i1455T (M. Groth-Maloney, D. Pruchter, F. Grewe, V. Knoop, submitted). Hence, serial intron entries beginning in the bryophyte clades for nad5i1455 and nad5i1477 are the likely evolutionary scenario (Fig. 2). As an alternative to evolution towards *trans*-splicing (in angiosperms), the introns can get lost secondarily in pteridophytes, as shown for nad1i669 and nad5i1477 (Dombrowska and Qiu 2004; Malek and Knoop 1998). An additional disruption splits intron nad1i728 in some angiosperm species (Chapdelaine and Bonen 1991; Conklin et al. 1991). This intron, which is omnipresent in the non-liverwort lineage (Qiu et al. 1998), has now been shown to evolve a *trans*-splicing status several times independently and differently among angiosperms (Qiu et al. 2004).

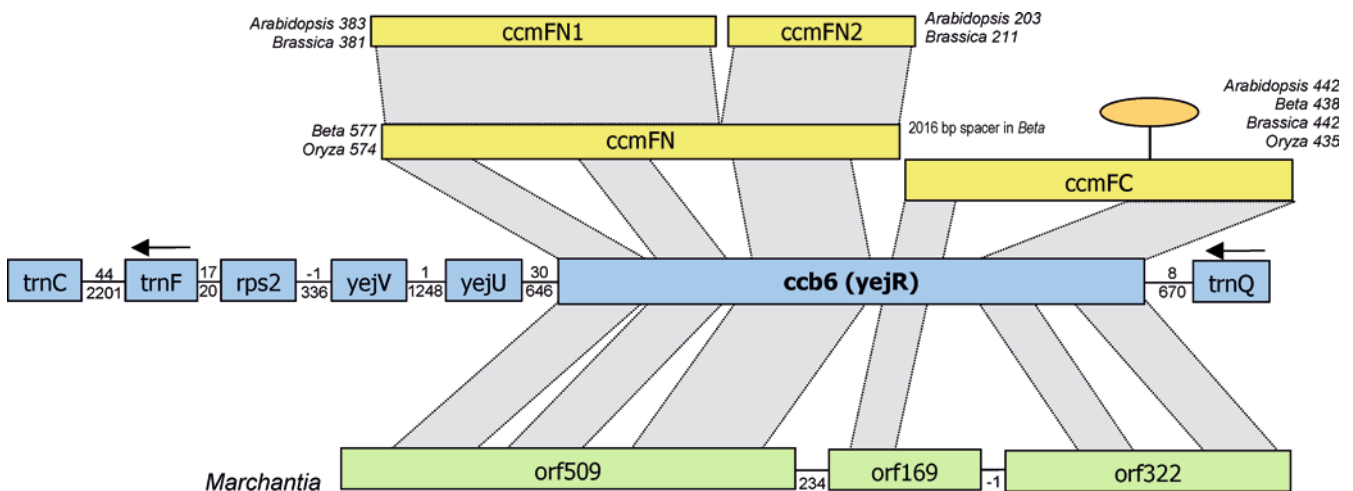
In line with a lack of conserved gene order in its chondriome, the alga *Mesostigma* features *trans*-splicing of group II introns, indicating an independent increase in recombinational activity in this early algal lineage. The complete chondriome sequence of this species

allowed the identification of two *trans*-splicing introns in the *nad3* gene, for which no orthologues have been reported so far (Turmel et al. 2002b).

In cpDNA of the streptophyte lineage, one example of a *trans*-splicing intron is known, which disrupts the continuity of the *rps12* gene but, as with most genomic features in the plastome, the emergence of this *trans*-splicing arrangement has also preceded the split of *Chaetosphaeridium* from the land plant lineage (Turmel et al. 2002a). A second example of chloroplast *trans*-splicing, taxonomically rather restricted, affects two introns in the *psaA* gene of the chlorophyte alga *Chlamydomonas reinhardtii* (Choquet et al. 1988; Turmel et al. 1995).

Gene disruptions are not restricted to introns, but can also affect coding regions in plant mitochondria. The *ccb* genes encoding components of cytochrome *c* biogenesis are embedded in the gene order *trnC-trnF-rps2-ccb2-ccb3-ccb6-trnQ* conserved in *Chara* and *Marchantia*. Gene spacers are extended significantly in the liverwort in comparison to the alga (Fig. 4), whereas the cluster is completely disintegrated in angiosperms. More notably, the *ccb6* orf itself is disrupted in embryophytes. The conservation of peptide motifs suggest that a first disruption in the middle of the *ccb6* gene is shared by *Marchantia* and the flowering plants. Non-coding spacers of 234 bp and 2,016 bp are found in *Marchantia* and sugar beet, respectively, at the corresponding positions, whereas linkage of the two resulting parts is completely lost in other angiosperms. A second disruption of the downstream part resulted in two orfs overlapping by one nucleotide in *Marchantia* and another independent disruption of the upstream orf (termed *ccmFN*) is found in Brassicaceae chondriomes (Fig. 4).

Fig. 4 Disintegration of the *ccb* gene cluster in embryophyte evolution. The gene order *trnC-trnF-rps2-ccb2(yejV)-ccb3(yejU)-ccb6(yejR)-trnQ* is conserved in the chondriomes of *Chara vulgaris* and *Marchantia polymorpha*, but not in angiosperms. Arrows indicate transcription from the opposite strand. Numbers above and below the lines indicate spacer sizes in *Chara* and *Marchantia*, respectively. The *ccb6* gene is split differently into separate orfs in the embryophytes. The ellipse indicates a group II intron in angiosperms



A second example of a discontinuity in a coding region, which additionally involves transfer to the nucleus, is the *rpl2* gene. The situation is particularly complex, since transfer of *rpl2* to the nucleus can occur in the whole or in parts after fission of the mitochondrial gene (Adams et al. 2001a). As an intermediate state, after transfer, the nucleus can provide the 3' part, with the 5' part coming from the mitochondrion.

RNA editing

Evolutionary and phylogenetic interest in plant mtDNA originated in part in the peculiar phenomenon of RNA editing (Hiesel et al. 1994a, 1994b). Plant mitochondrial mRNAs are subject to RNA editing, which site-specifically replaces specific cytidines by uridines to restore conserved codon identities (Covello and Gray 1989; Gualberto et al. 1989; Hiesel et al. 1989). Editing a protein orf is the rule in plant mitochondrial transcripts; and examples of a lack of editing in a mRNA, such as *rps12* in rice, are rare (Notsu et al. 2002). Non-coding RNA leader and trailer sequences, pseudo-genes, introns and tRNAs are affected much more rarely (Binder et al. 1992, 1994). In the latter, the editing events sometimes reconstitute base-pairing and can be a prerequisite for correct processing of the affected tRNA (Marchfelder et al. 1996; Marechal-Drouard et al. 1996a, 1996b). In other instances, similar base mismatches in tRNA stem regions are not corrected by RNA editing and this remains without consequence for correct tRNA processing (Schock et al. 1998). Similar observations are made for base-pairing regions in group II introns (Carrillo et al. 2001; Carrillo and Bonen 1997).

In chloroplasts, the same type of RNA editing process, exchanging pyrimidines, is active, although at a significantly lower rate. Whereas totals of 441 and 491 such sites are reported for transcripts of the *Arabidopsis* and *Oryza* chondriomes, respectively (Giege and Brennicke 1999; Notsu et al. 2002), only some 30 sites are observed in angiosperm chloroplast transcripts (Tsudzuki et al. 2001). Given this difference in

frequency between the organelles is a factor of at least ten, yet more striking are the differences between land plant groups. Initially assumed to be absent in bryophytes, RNA editing in the organelles has been identified in all classes of bryophytes (Freyer et al. 1997; Malek et al. 1996). However, a so far puzzling observation shows that RNA editing is present in the jungermanniid subclass (leafy, simple thallus) of liverworts to varying degrees, but is apparently lacking in the marchantiid subclass (complex thallus; Steinhauser et al. 1999). RNA editing is likewise not reported for algae. Very high frequencies of mtRNA editing are observed in the lycophyte *Isoetes* (Malek et al. 1996) and yet more surprising is RNA editing in hornworts and ferns. Only very rare exceptions of reversing the standard direction of C-to-U editing have been identified in angiosperms (Schuster et al. 1990), but “reverse” editing exchanging U-to-C is found frequently in hornworts (Steinhauser et al. 1999; Yoshinaga et al. 1996) and ferns (Vangerow et al. 1999). With the full plastome sequence of the hornwort *Anthoceros formosae* now available (Kugita et al. 2003a), a total of 509 C-to-U and 433 U-to-C edits have been identified in its chloroplast transcripts (Kugita et al. 2003b).

Foreign sequence insertions

The land plant chondriomes carry large stretches of intergenic sequences and most of this extra DNA has no significant homology to other sequences which would allow identifying their origin or potential function. In some cases however, the identity and origin of such sequences are obvious. The initial identification of a 12-kbp cpDNA sequence inserted into the maize chondriome (Stern and Lonsdale 1982) was followed by many other examples in many other species. For example, a systematic survey showed that coding sequences of the chloroplast *rbcL* gene (encoding ribulose-bisphosphate carboxylase) have been transferred independently several times (Cummings et al. 2003). The inter-organellar transfer of sequences is not restricted to cpDNA. Some “promiscuous” sequences were identified as degenerated retrotransposon sequences of all types that normally reside in the nucleus (Knoop et al. 1996). A total of 13.4% of sequences were nuclear and 6.3% were of plastid origin, respectively, in the 490.5-kbp rice mtDNA (Notsu et al. 2002). Hence, at least flowering plant mitochondria readily accept DNA from the chloroplast and the nuclear genome. Interestingly, a recent report demonstrated that the import of DNA into potato mitochondria is an active process (Koulintchenko et al. 2003).

No recognizable homology to sequences from the chloroplast or nuclear genome is present in the *Marchantia* chondriome. Thus, it is an open question when and why during evolution did plants accept foreign DNA inserted into their chondriomes. Clearly though, it is not the import of DNA alone that results

in the extraordinary size increase of plant chondriomes. In other cases such as the Cucurbitaceae family, which shows a striking diversity of chondriome sizes, extensive sequence duplications are among the major reasons for size increase (Lilly and Havey 2001).

Conclusions and outlook

Problems in establishing a robust land plant phylogenetic tree for the base-most (bryophyte) groups are at least two-fold: first, to identify (molecular) markers that have conserved very ancient phylogenetic information among the majority of non-informative exchanges, to document the very short internodes presumably resulting from a very early separation of the bryophyte lineages and, second, to identify the long branch attraction artifacts as a result of the evolutionary distance to even the most closely related outgroup algae (Charales, Coleochaetales). Most single-gene and multigene analyses of organelle genes at low taxon sampling result in liverwort-basal topologies with hornworts sister to tracheophytes (Fig. 2), albeit mostly with low statistical support (unpublished observations). How do features of the mitochondrial genome evolution fit into such a view of land plant phylogeny?

Parsimonious explanations of intron occurrence are complicated in cases where frequent independent gains are obvious, such as for the “rampant invader” group I intron *cox1i729*. Horizontal transfer, possibly from fungal sources, is an obvious explanation for the occurrence of most group I introns in the basal genera. *Chaetosphaeridium*, *Chara* and *Marchantia* combined have a total of 26 mitochondrial group I intron locations in *cob*, *cox1*, *cox2*, *rrn1* and *rrns*. Whereas only four of these are shared among them (two between the algae, one each between *Marchantia* and one of the algae), a full 19 of them have orthologues that can be identified in the available mitochondrial sequences of fungi or chlorophyte algae.

In contrast, the occurrence of group II introns in land plants seems to be largely explained by vertical transmission and occasional loss. Strong support for the liverwort basal topology originally deduced from *cox2i373*, *cox2i691* and *nad1i728* absent in liverworts (Qiu et al. 1998) is further corroborated by *nad2i156*, *nad4i461*, *nad7i140*, *nad7i209*, *nad5i230* and *nad5i1455* in the non-liverwort branch. Notably, the topology is consistent with serial entries of the introns later developing a *trans*-splicing status in angiosperms: *nad5i1455* at node Y, *nad5i1477* at Z and *nad2i542*, *nad1i394* and *nad1i669* in the tracheophyte ancestor (Fig. 2). A sister group relation of hornworts and tracheophytes would be further supported by the gain of *nad2i1282* and loss of *nad5i753*, the group I intron which is vertically inherited in mosses and liverworts. However, some intron losses on the backbone of land plant phylogeny need to be postulated in any topology and, in the topology shown in Fig. 2, this would be the

loss of nad2i709 in the last common ancestor of mosses and nad2i156 in hornworts.

The suite of genes present in the common ancestor of embryophytes was most likely that of *Marchantia*, including functional nad7 and rpl14 genes (Fig. 1). Surprises for additional genes in mosses, the potential base-most non-liverwort clade (Fig. 2) are not to be expected. Given its presence in the chondriomes of *Chara* and *Marchantia* and its general absence in angiosperms, the rpl6 gene is an interesting candidate for a gene transfer early in the non-liverwort embryophytes (NLE) lineage.

No conservation of gene order is discernible in angiosperms, but the gene clusters conserved between *Marchantia* and *Chara* could well be conserved in the basal NLE groups. At least the nad5-nad4-nad2 gene arrangement is indeed conserved in both mosses and hornworts but is only partially retained in *Isoetes* (unpublished data).

A significant increase in chondriome size occurred early in embryophyte establishment. The total amount of intron sequences doubled in *Marchantia*, compared with *Chara* (56 kbp vs 28 kbp, respectively), but the contribution of 100 kbp of intergenic DNA to the non-coding sequences in the *Marchantia* chondriome was considerably larger. The absence of evolutionary pressure on these sequences may have obliterated the origins of early sequence acquisitions, whereas the source of the likely more recently transferred DNA in angiosperms is clearly recognizable as the chloroplast or nucleus.

There is only scarce information on other non-flowering plant chondriomes. The mtDNA of the fern *Oncoclea sensibilis* and the horsetail *Equisetum arvense* are approximately 300 kbp and 200 kbp in size, respectively (Palmer et al. 1992), and thus well within the size range of *Marchantia* and the smaller of the known angiosperm mtDNAs. Partial sequence information of clones from lycophyte mtDNAs of *Selaginella* and *Isoetes* show cpDNA homologies in the chondriomes of both genera (unpublished observations), indicating that mechanisms to stably integrate cpDNA into plant chondriomes evolved in the common ancestor of the tracheophytes (Fig. 2).

The occurrence of RNA editing remains puzzling. The high ratio of U-to-C in addition to C-to-U editing could be a transitory phenomenon arising in the common ancestor of hornworts and tracheophytes, retained in lycophytes and moniliformopses and on its way out in the ancestor of seed plants. While at first glance a gain in RNA editing may have coincided with the emergence of land plants, its apparent absence in the marchantiid liverworts is problematic, given that recent phylogenetic analyses support the monophyly of liverworts as a whole (Fig. 2), including the conservation of introns in both liverwort subclasses (for a divergent phylogenetic assessment, see Capesius and Bopp 1997). For now, a loss of RNA editing in marchantiid or alternatively a later separate gain in jungermanniid liverworts needs to be postulated.

Chondriome sequences of mosses, hornworts, lycophytes, moniliformopses and gymnosperms may reveal further surprises but will in any case most likely continue to provide valuable information on plant phylogeny.

Acknowledgements The author gratefully acknowledges continuous collegial exchange with Y.-L. Qiu (Ann Arbor, Mich., USA) and welcomes his laboratory's initiative to standardize organelle intron nomenclature. I wish to thank Jan-Peter Frahm (Bonn, Germany) for the bryophyte photographs used in Fig. 2 and the Deutsche Forschungsgemeinschaft (DFG) for their support of research in my laboratory.

References

- Abdelnoor RV, Yule R, Elo A, Christensen AC, Meyer-Gauen G, Mackenzie SA (2003) Substoichiometric shifting in the plant mitochondrial genome is influenced by a gene homologous to MutS. *Proc Natl Acad Sci USA* 100:5968–5973
- Adams KL, Palmer JD (2003) Evolution of mitochondrial gene content: gene loss and transfer to the nucleus. *Mol Phylogenet Evol* 29:380–395
- Adams KL, et al (1998) Evolution of flowering plant mitochondrial genomes: gene content, gene transfer to the nucleus, and highly accelerated mutation rates. In: Moller IM, Gardestrom P, Glimelius K, Glaser E (eds) *Plant mitochondria: from gene to function*. Backhuys, Leiden, pp 13–18
- Adams KL, Song KM, Roessler PG, Nugent JM, Doyle JL, Doyle JJ, Palmer JD (1999) Intracellular gene transfer in action: dual transcription and multiple silencings of nuclear and mitochondrial *cox2* genes in legumes. *Proc Natl Acad Sci USA* 96:13863–13868
- Adams KL, Daley DO, Qiu YL, Whelan J, Palmer JD (2000) Repeated, recent and diverse transfers of a mitochondrial gene to the nucleus in flowering plants. *Nature* 408:354–357
- Adams KL, Ong HC, Palmer JD (2001a) Mitochondrial gene transfer in pieces: fission of the ribosomal protein gene *rpl2* and partial or complete gene transfer to the nucleus. *Mol Biol Evol* 18:2289–2297
- Adams KL, Rosenblueth M, Qiu YL, Palmer JD (2001b) Multiple losses and transfers to the nucleus of two mitochondrial succinate dehydrogenase genes during angiosperm evolution. *Genetics* 158:1289–1300
- Adams KL, Daley DO, Whelan J, Palmer JD (2002a) Genes for two mitochondrial ribosomal proteins in flowering plants are derived from their chloroplast or cytosolic counterparts. *Plant Cell* 14:931–943
- Adams KL, Qiu YL, Stoutemyer M, Palmer JD (2002b) Punctuated evolution of mitochondrial gene content: High and variable rates of mitochondrial gene loss and transfer to the nucleus during angiosperm evolution. *Proc Natl Acad Sci USA* 99:9905–9912
- Andre CP, Walbot V (1995) Pulsed-field gel mapping of maize mitochondrial chromosomes. *Mol Gen Genet* 247:255–263
- Bakker FT, Culham A, Pankhurst CE, Gibby M (2000) Mitochondrial and chloroplast DNA-based phylogeny of Pelargonium (Geraniaceae). *Am J Bot* 87:727–734
- Beagley CT, Okada NA, Wolstenholme DR (1996) Two mitochondrial group I introns in a metazoan, the sea anemone *Metridium senile*: one intron contains genes for subunits 1 and 3 of NADH dehydrogenase. *Proc Natl Acad Sci USA* 93:5619–5623
- Beckert S, Steinhauser S, Muhle H, Knoop V (1999) A molecular phylogeny of bryophytes based on nucleotide sequences of the mitochondrial *nad5* gene. *Plant Syst Evol* 218:179–192
- Beckert S, Muhle H, Pruchner D, Knoop V (2001) The mitochondrial *nad2* gene as a novel marker locus for phylogenetic analysis of early land plants: a comparative analysis in mosses. *Mol Phylogenet Evol* 18:117–126

- Bendich AJ (1993) Reaching for the ring: the study of mitochondrial genome structure. *Curr Genet* 24:279–290
- Bergthorsson U, Adams KL, Thomason B, Palmer JD (2003) Widespread horizontal transfer of mitochondrial genes in flowering plants. *Nature* 424:197–201
- Binder S, Marchfelder A, Brennicke A (1994) RNA editing of tRNA(Phe) and tRNA(Cys) in mitochondria of *Oenothera berteriana* is initiated in precursor molecules. *Mol Gen Genet* 244:67–74
- Binder S, Marchfelder A, Brennicke A, Wissinger B (1992) RNA editing in trans-splicing intron sequences of nad2 mRNAs in *Oenothera* mitochondria. *J Biol Chem* 267:7615–7623
- Blanchard JL, Schmidt GW (1995) Pervasive migration of organellar DNA to the nucleus in plants. *J Mol Evol* 41:397–406
- Bowe LM, Coat G, dePamphilis CW (2000) Phylogeny of seed plants based on all three genomic compartments: extant gymnosperms are monophyletic and Gnetales' closest relatives are conifers. *Proc Natl Acad Sci USA* 97:4092–4097
- Budar F, Pelletier G (2001) Male sterility in plants: occurrence, determinism, significance and use. *C R Acad Sci III* 324:543–550
- Burger G, Forget L, Zhu Y, Gray MW, Lang BF (2003a) Unique mitochondrial genome architecture in unicellular relatives of animals. *Proc Natl Acad Sci USA* 100:892–897
- Burger G, Lang BF, Braun HP, Marx S (2003b) The enigmatic mitochondrial ORF ymf39 codes for ATP synthase chain b. *Nucleic Acids Res* 31:2353–2360
- Capesius I, Bopp M (1997) New classification of liverworts based on molecular and morphological data. *Plant Syst Evol* 207:87–97
- Carrillo C, Bonen L (1997) RNA editing status of nad7 intron domains in wheat mitochondria. *Nucleic Acids Res* 25:403–409
- Carrillo C, Chapdelaine Y, Bonen L (2001) Variation in sequence and RNA editing within core domains of mitochondrial group II introns among plants. *Mol Gen Genet* 264:595–603
- Chapdelaine Y, Bonen L (1991) The wheat mitochondrial gene for subunit I of the NADH dehydrogenase complex: a trans-splicing model for this gene-in-pieces. *Cell* 65:465–472
- Chaw SM, Zharkikh A, Sung HM, Lau TC, Li WH (1997) Molecular phylogeny of extant gymnosperms and seed plant evolution: Analysis of nuclear 18S rRNA sequences. *Mol Biol Evol* 14:56–68
- Chaw SM, Parkinson CL, Cheng YC, Vincent TM, Palmer JD (2000) Seed plant phylogeny inferred from all three plant genomes: monophyly of extant gymnosperms and origin of Gnetales from conifers. *Proc Natl Acad Sci USA* 97:4086–4091
- Cho Y, Qiu YL, Kuhlman P, Palmer JD (1998) Explosive invasion of plant mitochondria by a group I intron. *Proc Natl Acad Sci USA* 95:14244–14249
- Choquet Y, Goldschmidt-Clermont M, Girard-Bascou J, Kuck U, Bannou P, Rochaix JD (1988) Mutant phenotypes support a trans-splicing mechanism for the expression of the tripartite *psaA* gene in the *C. reinhardtii* chloroplast. *Cell* 52:903–913
- Conklin PL, Wilson RK, Hanson MR (1991) Multiple trans-splicing events are required to produce a mature nad1 transcript in a plant mitochondrion. *Genes Dev* 5:1407–1415
- Cosner ME, Jansen RK, Palmer JD, Downie SR (1997) The highly rearranged chloroplast genome of *Trachelium caeruleum* (Campanulaceae): multiple inversions, inverted repeat expansion and contraction, transposition, insertions/deletions, and several repeat families. *Curr Genet* 31:419–429
- Covello PS, Gray MW (1989) RNA editing in plant mitochondria. *Nature* 341:662–666
- Crepet WL (2000) Progress in understanding angiosperm history, success, and relationships: Darwin's abominably "perplexing phenomenon". *Proc Natl Acad Sci USA* 97:12939–12941
- Cummings MP, Nugent JM, Olmstead RG, Palmer JD (2003) Phylogenetic analysis reveals five independent transfers of the chloroplast gene *rbcL* to the mitochondrial genome in angiosperms. *Curr Genet* 43:131–138
- Dai L, Zimmerly S (2002) Compilation and analysis of group II intron insertions in bacterial genomes: evidence for retroelement behavior. *Nucleic Acids Res* 30:1091–1102
- Daley DO, Adams KL, Clifton R, Qualmann S, Millar AH, Palmer JD, Pratje E, Whelan J (2002) Gene transfer from mitochondrion to nucleus: novel mechanisms for gene activation from *Cox2*. *Plant J* 30:11–21
- Dombrovska O, Qiu YL (2004) Distribution of Introns in the mitochondrial gene *nad1* in land plants: phylogenetic and molecular evolutionary implication. *Mol Phylogenet Evol* 32:246–263
- Donoghue MJ, Doyle JA (2000) Seed plant phylogeny: demise of the anthophyte hypothesis? *Curr Biol* 10:106–109
- Duchene AM, Marechal-Drouard L (2001) The chloroplast-derived *trnW* and *trnM-e* genes are not expressed in *Arabidopsis* mitochondria. *Biochem Biophys Res Commun* 285:1213–1216
- Dyall SD, Brown MT, Johnson PJ (2004) Ancient invasions: from endosymbionts to organelles. *Science* 304:253–257
- Figueroa P, Gomez I, Holuigue L, Araya A, Jordana X (1999) Transfer of *rps14* from the mitochondrion to the nucleus in maize implied integration within a gene encoding the iron-sulphur subunit of succinate dehydrogenase and expression by alternative splicing. *Plant J* 18:601–609
- Freyer R, Kiefer-Meyer MC, Kossel H (1997) Occurrence of plastid RNA editing in all major lineages of land plants. *Proc Natl Acad Sci USA* 94:6285–6290
- Gamboa MA, Laureano S, Bayman P (2003) Measuring diversity of endophytic fungi in leaf fragments: does size matter? *Mycopathologia* 156:41–45
- Geiss KT, Abbas GM, Makaroff CA (1994) Intron loss from the *nadh* dehydrogenase subunit 4 gene of lettuce mitochondrial DNA—evidence for homologous recombination of a cDNA intermediate. *Mol Gen Genet* 243:97–105
- Giege P, Brennicke A (1999) RNA editing in *Arabidopsis* mitochondria effects 441 C to U changes in ORFs. *Proc Natl Acad Sci USA* 96:15324–15329
- Goremykin V, Bobrova V, Pahnke J, Troitsky A, Antonov A, Martin W (1996) Noncoding sequences from the slowly evolving chloroplast inverted repeat in addition to *rbcL* data do not support Gnetalean affinities of angiosperms. *Mol Biol Evol* 13:383–396
- Graham LE, Cook ME, Busse JS (2000) The origin of plants: body plan changes contributing to a major evolutionary radiation. *Proc Natl Acad Sci USA* 97:4535–4540
- Gray MW (1999) Evolution of organellar genomes. *Curr Opin Genet Dev* 9:678–687
- Gray MW, Lang BF, Cedergren R, Golding GB, Lemieux C, Sankoff D, Turmel M, Brossard N, Delage E, Littlejohn TG, Plante I, Rioux P, Saint-Louis D, Zhu Y, Burger G (1998) Genome structure and gene content in protist mitochondrial DNAs. *Nucleic Acids Res* 26:865–878
- Gray MW, Burger G, Lang BF (1999) Mitochondrial evolution. *Science* 283:1476–1481
- Gualberto JM, Lamattina L, Bonnard G, Weil JH, Grienenberger JM (1989) RNA editing in wheat mitochondria results in the conservation of protein sequences. *Nature* 341:660–662
- Gugerli F, Sperisen C, Buchler U, Brunner I, Brodbeck S, Palmer JD, Qiu YL (2001) The evolutionary split of Pinaceae from other conifers: evidence from an intron loss and a multigene phylogeny. *Mol Phylogenet Evol* 21:167–175
- Handa H (2003) The complete nucleotide sequence and RNA editing content of the mitochondrial genome of rapeseed (*Brassica napus* L.): comparative analysis of the mitochondrial genomes of rapeseed and *Arabidopsis thaliana*. *Nucleic Acids Res* 31:5907–5916
- Handa H, Itani K, Sato H (2002) Structural features and expression analysis of a linear mitochondrial plasmid in rapeseed (*Brassica napus* L.). *Mol Genet Genomics* 267:797–805
- Hashimoto K, Sato N (2001) Characterization of the mitochondrial *nad7* gene in *Physcomitrella patens*: similarity with angiosperm *nad7* genes. *Plant Sci* 160:807–815
- Heazlewood JL, Whelan J, Millar AH (2003) The products of the mitochondrial *orf25* and *orfB* genes are FO components in the plant F1FO ATP synthase. *FEBS Lett* 540:201–205

- Hiesel R, Wissinger B, Schuster W, Brennicke A (1989) RNA editing in plant mitochondria. *Science* 246:1632–1634
- Hiesel R, Combettes B, Brennicke A (1994a) Evidence for RNA editing in mitochondria of all major groups of land plants except the Bryophyta. *Proc Natl Acad Sci USA* 91:629–633
- Hiesel R, Haeseler A von, Brennicke A (1994b) Plant mitochondrial nucleic acid sequences as a tool for phylogenetic analysis. *Proc Natl Acad Sci USA* 91:634–638
- Huang CY, Ayliffe MA, Timmis JN (2003) Direct measurement of the transfer rate of chloroplast DNA into the nucleus. *Nature* 422:72–76
- Itchoda N, Nishizawa S, Nagano H, Kubo T, Mikami T (2002) The sugar beet mitochondrial nad4 gene: an intron loss and its phylogenetic implication in the Caryophyllales. *Theor Appl Genet* 104:209–213
- Joly S, Brouillet L, Bruneau A (2001) Phylogenetic implications of the multiple losses of the mitochondrial coxII.i3 intron in the angiosperms. *Int J Plant Sci* 162:359–373
- Kadowaki KI, Kubo N, Ozawa K, Hirai A (1996) Targeting pre-sequence acquisition after mitochondrial gene transfer to the nucleus occurs by duplication of existing targeting signals. *EMBO J* 15:6652–6661
- Karol KG, McCourt RM, Cimino MT, Delwiche CF (2001) The closest living relatives of land plants. *Science* 294:2351–2353
- Kenrick P, Crane PR (1997a) A cladistic study. Smithsonian Institution, Washington, D.C.
- Kenrick P, Crane PR (1997b) The origin and early evolution of plants on land. *Nature* 389:33–39
- Klein M, Eckert-Ossenkopp U, Schmiedeberg I, Brandt P, Unseld M, Brennicke A, Schuster W (1994) Physical mapping of the mitochondrial genome of *Arabidopsis thaliana* by cosmid and YAC clones. *Plant J* 6:447–455
- Knoop V, Brennicke A (1991) A mitochondrial intron sequence in the 5'-flanking region of a plant nuclear lectin gene. *Curr Genet* 20:423–425
- Knoop V, Brennicke A (1994) Promiscuous mitochondrial group II intron sequences in plant nuclear genomes. *J Mol Evol* 39:144–150
- Knoop V, Schuster W, Wissinger B, Brennicke A (1991) Trans splicing integrates an exon of 22 nucleotides into the nad5 mRNA in higher plant mitochondria. *EMBO J* 10:3483–3493
- Knoop V, Ehrhardt T, Lüttig K, Brennicke A (1995) The gene for ribosomal protein S10 is present in mitochondria of pea and potato but absent from those of *Arabidopsis* and *Oenothera*. *Curr Genet* 27:559–564
- Knoop V, Unseld M, Marienfeld J, Brandt P, Sunkel S, Ullrich H, Brennicke A (1996) copia-, gypsy- and LINE-like retrotransposon fragments in the mitochondrial genome of *Arabidopsis thaliana*. *Genetics* 142:579–585
- Knoop V, Altwasser M, Brennicke A (1997) A tripartite group II intron in mitochondria of an angiosperm plant. *Mol Gen Genet* 255:269–276
- Kobayashi Y, Knoop V, Fukuzawa H, Brennicke A, Ohya Y (1997) Interorganellar gene transfer in bryophytes: the functional nad7 gene is nuclear encoded in *Marchantia polymorpha*. *Mol Gen Genet* 256:589–592
- Kolukisaoglu HU, Marx S, Wiegmann C, Hanelt S, Schneider-Poetsch HA (1995) Divergence of the phytochrome gene family predates angiosperm evolution and suggests that *Selaginella* and *Equisetum* arose prior to *Psilotum*. *J Mol Evol* 41:329–337
- Koulintchenko M, Konstantinov Y, Dietrich A (2003) Plant mitochondria actively import DNA via the permeability transition pore complex. *EMBO J* 22:1245–1254
- Kroymann J, Zetsche K (1998) The mitochondrial genome of *Chlorogonium elongatum* inferred from the complete sequence. *J Mol Evol* 47:431–440
- Kubo T, Nishizawa S, Sugawara A, Itchoda N, Estiati A, Mikami T (2000) The complete nucleotide sequence of the mitochondrial genome of sugar beet (*Beta vulgaris* L.) reveals a novel gene for tRNA(Cys)(GCA). *Nucleic Acids Res* 28:2571–2576
- Kubo N, Arimura S, Tsutsumi N, Hirai A, Kadowaki K (2003) Involvement of N-terminal region in mitochondrial targeting of rice RPS10 and RPS14 proteins. *Plant Sci* 164:1047–1055
- Kudla J, Albertazzi FJ, Blazevic D, Hermann M, Bock R (2002) Loss of the mitochondrial cox2 intron 1 in a family of monocotyledonous plants and utilization of mitochondrial intron sequences for the construction of a nuclear intron. *Mol Genet Genom* 267:223–230
- Kugita M, Kaneko A, Yamamoto Y, Takeya Y, Matsumoto T, Yoshinaga K (2003a) The complete nucleotide sequence of the hornwort (*Anthoceros formosae*) chloroplast genome: insight into the earliest land plants. *Nucleic Acids Res* 31:716–721
- Kugita M, Yamamoto Y, Fujikawa T, Matsumoto T, Yoshinaga K (2003b) RNA editing in hornwort chloroplasts makes more than half the genes functional. *Nucleic Acids Res* 31:2417–2423
- Kumar R, Marechal-Drouard L, Akama K, Small I (1996) Striking differences in mitochondrial tRNA import between different plant species. *Mol Gen Genet* 252:404–411
- Lemieux C, Otis C, Turmel M (2000) Ancestral chloroplast genome in *Mesostigma viride* reveals an early branch of green plant evolution. *Nature* 403:649–652
- Ligrone R, Pocock K, Duckett JG (1993) A comparative ultra-structural study of endophytic basidiomycetes in the parasitic achlorophyllous hepatic *Cryptothallus mirabilis* and the closely allied photosynthetic species *Aneura pinguis* (Metzgeriales). *Can J Bot* 71:666–679
- Lilly JW, Havey MJ (2001) Small, repetitive DNAs contribute significantly to the expanded mitochondrial genome of cucumber. *Genetics* 159:317–328
- Magallon S, Sanderson MJ (2002) Relationships among seed plants inferred from highly conserved genes: sorting conflicting phylogenetic signals among ancient lineages. *Am J Bot* 89:1991–2006
- Malek O, Knoop V (1998) Trans-splicing group II introns in plant mitochondria: the complete set of cis-arranged homologs in ferns, fern allies, and a hornwort. *RNA* 4:1599–1609
- Malek O, Lüttig K, Hiesel R, Brennicke A, Knoop V (1996) RNA editing in bryophytes and a molecular phylogeny of land plants. *EMBO J* 15:1403–1411
- Malek O, Brennicke A, Knoop V (1997) Evolution of trans-splicing plant mitochondrial introns in pre-Permian times. *Proc Natl Acad Sci USA* 94:553–558
- Manhart JR, Palmer JD (1990) The gain of two chloroplast tRNA introns marks the green algal ancestors of land plants. *Nature* 345:268–270
- Marchfelder A, Brennicke A, Binder S (1996) RNA editing is required for efficient excision of tRNA(Phe) from precursors in plant mitochondria. *J Biol Chem* 271:1898–1903
- Marechal-Drouard L, Guillemaut P, Cosset A, Arbogast M, Weber F, Weil JH, Dietrich A (1990) Transfer RNAs of potato (*Solanum tuberosum*) mitochondria have different genetic origins. *Nucleic Acids Res* 18:3689–3696
- Marechal-Drouard L, Cosset A, Remacle C, Ramamonjisoa D, Dietrich A (1996a) A single editing event is a prerequisite for efficient processing of potato mitochondrial phenylalanine tRNA. *Mol Cell Biol* 16:3504–3510
- Marechal-Drouard L, Kumar R, Remacle C, Small I (1996b) RNA editing of larch mitochondrial tRNA(His) precursors is a prerequisite for processing. *Nucleic Acids Res* 24:3229–3234
- Michel F, Ferat JL (1995) Structure and activities of group II introns. *Annu Rev Biochem* 64:435–461
- Milligan BG, Hampton JN, Palmer JD (1989) Dispersed repeats and structural reorganization in subcloroplast DNA. *Mol Biol Evol* 6:355–368
- Nakazono M, Itadani H, Wakasugi T, Tsutsumi N, Sugiura M, Hirai A (1995) The rps3-rpl16-nad3-rps12 gene cluster in rice mitochondrial DNA is transcribed from alternative promoters. *Curr Genet* 27:184–189
- Nedelcu AM (1997) Fragmented and scrambled mitochondrial ribosomal RNA coding regions among green algae: a model for their origin and evolution. *Mol Biol Evol* 14:506–517

- Nedelcu AM, Lee RW, Lemieux C, Gray MW, Burger G (2000) The complete mitochondrial DNA sequence of *Scenedesmus obliquus* reflects an intermediate stage in the evolution of the green algal mitochondrial genome. *Genome Res* 10:819–831
- Nickrent DL, Parkinson CL, Palmer JD, Duff RJ (2000) Multigene phylogeny of land plants with special reference to bryophytes and the earliest land plants. *Mol Biol Evol* 17:1885–1895
- Notsu Y, Masood S, Nishikawa T, Kubo N, Akiduki G, Nakazono M, Hirai A, Kadowaki K (2002) The complete sequence of the rice (*Oryza sativa* L.) mitochondrial genome: frequent DNA sequence acquisition and loss during the evolution of flowering plants. *Mol Genet Genomics* 268:434–445
- Nugent JM, Palmer JD (1991) RNA-mediated transfer of the gene *coxII* from the mitochondrion to the nucleus during flowering plant evolution. *Cell* 66:473–481
- Oda K, Yamato K, Ohta E, Nakamura Y, Takemura M, Nozato N, Akashi K, Kanegae T, Ogura Y, Kohchi T (1992) Gene organization deduced from the complete sequence of liverwort *Marchantia polymorpha* mitochondrial DNA. A primitive form of plant mitochondrial genome. *J Mol Biol* 223:1–7
- Ohta E, Oda K, Yamato K, Nakamura Y, Takemura M, Nozato N, Akashi K, Ohshima K, Michel F (1993) Group-I introns in the liverwort mitochondrial genome—the gene coding for subunit-1 of cytochrome-oxidase shares 5 intron positions with its fungal counterparts. *Nucleic Acids Res* 21:1297–1305
- Ohyama K, Oda K, Ohta E, Takemura M (1993). Plant mitochondria. VCH, Weinheim, pp 115–129
- Oldenburg DJ, Bendich AJ (2001) Mitochondrial DNA from the liverwort *Marchantia polymorpha*: circularly permuted linear molecules, head-to-tail concatamers, and a 5' protein. *J Mol Biol* 310:549–562
- Palmer JD (1990) Contrasting modes and tempos of genome evolution in land plant organelles. *Trends Genet* 6:115–120
- Palmer JD, Herbon LA (1987) Unicircular structure of the *Brassica hirta* mitochondrial genome. *Curr Genet* 11:565–570
- Palmer JD, Herbon LA (1988) Plant mitochondrial DNA evolves rapidly in structure, but slowly in sequence. *J Mol Evol* 28:87–97
- Palmer JD, Shields CR (1984) Tripartite structure of the *Brassica campestris* mitochondrial genome. *Nature* 307:437–440
- Palmer JD, Soltis D, Soltis P (1992) Large size and complex structure of mitochondrial DNA in two nonflowering land plants. *Curr Genet* 21:125–129
- Palmer JD, Adams KL, Cho Y, Parkinson CL, Qiu YL, Song K (2000) Dynamic evolution of plant mitochondrial genomes: mobile genes and introns and highly variable mutation rates. *Proc Natl Acad Sci USA* 97:6960–6966
- Pereira DS, Jubier MF, Delcher E, Lancelin D, Lejeune B (1991) A trans-splicing model for the expression of the tripartite *nad5* gene in wheat and maize mitochondria. *Plant Cell* 3:1363–1378
- Pla M, Mathieu C, DePaepe R, Chetrit P, Vedel F (1995) Deletion of the last 2 exons of the mitochondrial *Nad7* gene results in lack of the *nad7* polypeptide in a *Nicotiana sylvestris* Cms mutant. *Mol Gen Genet* 248:79–88
- Pruchner D, Nassal B, Schindler M, Knoop V (2001) Mosses share mitochondrial group II introns with flowering plants, not with liverworts. *Mol Genet Genom* 266:608–613
- Pruchner D, Beckert S, Muhle H, Knoop V (2002) Divergent intron conservation in the mitochondrial *nad2* gene: signatures for the three bryophyte classes (mosses, liverworts, and hornworts) and the lycophytes. *J Mol Evol* 55:265–271
- Pryer KM, Schneider H, Smith AR, Cranfill R, Wolf PG, Hunt JS, Sipes SD (2001) Horsetails and ferns are a monophyletic group and the closest living relatives to seed plants. *Nature* 409:618–622
- Pryer KM, Schneider H, Zimmer EA, Ann BJ (2002) Deciding among green plants for whole genome studies. *Trends Plant Sci* 7:550–554
- Qiu YL, Palmer JD (1999) Phylogeny of early land plants: insights from genes and genomes. *Trends Plant Sci* 4:26–30
- Qiu YL, Cho YR, Cox JC, Palmer JD (1998) The gain of three mitochondrial introns identifies liverworts as the earliest land plants. *Nature* 394:671–674
- Qiu YL, Lee J, Whitlock BA, Bernasconi-Quadroni F, Dombrowska O (2001) Was the ANITA rooting of the angiosperm phylogeny affected by long-branch attraction? Amborella, Nymphaeales, Illiciales, Trimeniaceae, and Austrobaileya. *Mol Biol Evol* 18:1745–1753
- Qiu YL, Dombrowska O, Palmer JD (2004) Many independent evolutions of trans-splicing of a plant mitochondrial group II intron. *J Mol Evol* (in press)
- Raubeson LA, Jansen RK (1992) Chloroplast DNA evidence on the ancient evolutionary split in vascular land plants. *Science* 255:1697–1699
- Read DJ, Duckett JG, Francis R, Ligrone R, Russell A (2000) Symbiotic fungal associations in 'lower' land plants. *Philos Trans R Soc Lond B Biol Sci* 355:815–830
- Redecker D, Kodner R, Graham LE (2000) Glomalean fungi from the Ordovician. *Science* 289:1920–1921
- Renzaglia KS, Duff RJ, Nickrent DL, Garbary DJ (2000) Vegetative and reproductive innovations of early land plants: implications for a unified phylogeny. *Philos Trans R Soc Lond B Biol Sci* 355:769–793
- Rydin C, Kallersjö M, Friest EM (2002) Seed plant relationships and the systematic position of Gnetales based on nuclear and chloroplast DNA: conflicting data, rooting problems, and the monophyly of conifers. *Int J Plant Sci* 163:197–214
- Sabar M, Gagliardi D, Balk J, Leaver CJ (2003) ORFB is a subunit of F(1)F(O)-ATP synthase: insight into the basis of cytoplasmic male sterility in sunflower. *EMBO Rep* 4:1–6
- Sanchez H, Fester T, Kloska S, Schroder W, Schuster W (1996) Transfer of *rps19* to the nucleus involves the gain of an RNP-binding motif which may functionally replace RPS13 in *Arabidopsis* mitochondria. *EMBO J* 15:2138–2149
- Sandoval P, Leon G, Gomez I, Carmona R, Figueroa P, Holuigue L, Araya A, Jordana X (2004) Transfer of RPS14 and RPL5 from the mitochondrion to the nucleus in grasses. *Gene* 324:139–147
- Schafer B, Kaulich K, Wolf K (1998) Mosaic structure of the *cox2* gene in the petite negative yeast *Schizosaccharomyces pombe*: a group II intron is inserted at the same location as the otherwise unrelated group II introns in the mitochondria of higher plants. *Gene* 214:101–112
- Schock I, Marechal-Drouard L, Marchfelder A, Binder S (1998) Processing of plant mitochondrial tRNAGly and tRNA-Ser(GCU) is independent of RNA editing. *Mol Gen Genet* 257:554–560
- Schuster W, Hiesel R, Wissinger B, Brennicke A (1990) RNA editing in the cytochrome b locus of the higher plant *Oenothera berteriana* includes a U-to-C transition. *Mol Cell Biol* 10:2428–2431
- Scotti N, Marechal-Drouard L, Cardi T (2004) The *rpl5-rps14* mitochondrial region: a hot spot for DNA rearrangements in *Solanum* spp somatic hybrids. *Curr Genet* 45:378–382
- Small I, Suffolk R, Leaver CJ (1989) Evolution of plant mitochondrial genomes via substoichiometric intermediates. *Cell* 58:69–76
- Sperwhitis GL, Russell AL, Vaughn JC (1994) Mitochondrial RNA editing of cytochrome-c-oxidase subunit II (CoxII) in the primitive vascular plant *Psilotum nudum*. *Biochem Biophys Acta Gene Struct Expr* 1218:218–220
- Sperwhitis GL, Moody JL, Vaughn JC (1996) Universality of mitochondrial RNA editing in cytochrome-c oxidase subunit I (*coxI*) among the land plants. *Biochem Biophys Acta Gene Struct Expr* 1307:301–308
- Stegemann S, Hartmann S, Ruf S, Bock R (2003) High-frequency gene transfer from the chloroplast genome to the nucleus. *Proc Natl Acad Sci USA* (in press)
- Steinhauser S, Beckert S, Capesius I, Malek O, Knoop V (1999) Plant mitochondrial RNA editing. *J Mol Evol* 48:303–312
- Stern DB, Lonsdale DM (1982) Mitochondrial and chloroplast genomes of maize have a 12-kilobase DNA sequence in common. *Nature* 299:698–702
- Stupar RM, Lilly JW, Town CD, Cheng Z, Kaul S, Buell CR, Jiang J (2001) Complex mtDNA constitutes an approximate 620-kb insertion on *Arabidopsis thaliana* chromosome 2: implication of

- potential sequencing errors caused by large-unit repeats. *Proc Natl Acad Sci USA* 98:5099–5103
- Timmis JN, Ayliffe MA, Huang CY, Martin W (2004) Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nat Rev Genet* 5:123–U16
- Toor N, Zimmerly S (2002) Identification of a family of group II introns encoding LAGLIDADG ORFs typical of group I introns. *RNA* 8:1373–1377
- Tsudzuki T, Wakasugi T, Sugiura M (2001) Comparative analysis of RNA editing sites in higher plant chloroplasts. *J Mol Evol* 53:327–332
- Turmel M, Choquet Y, Goldschmidt-Clermont M, Rochaix JD, Otis C, Lemieux C (1995) The trans-spliced intron 1 in the *psaA* gene of the *Chlamydomonas* chloroplast: a comparative analysis. *Curr Genet* 27:270–279
- Turmel M, Lemieux C, Burger G, Lang BF, Otis C, Plante I, Gray MW (1999) The complete mitochondrial DNA sequences of *Nephroselmis olivacea* and *Pedinomonas minor*. Two radically different evolutionary patterns within green algae. *Plant Cell* 11:1717–1730
- Turmel M, Otis C, Lemieux C (2002a) The chloroplast and mitochondrial genome sequences of the charophyte *Chaetosphaeridium globosum*: insights into the timing of the events that restructured organelle DNAs within the green algal lineage that led to land plants. *Proc Natl Acad Sci USA* 99:11275–11280
- Turmel M, Otis C, Lemieux C (2002b) The complete mitochondrial DNA sequence of *Mesostigma viride* identifies this green alga as the earliest green plant divergence and predicts a highly compact mitochondrial genome in the ancestor of all green plants. *Mol Biol Evol* 19:24–38
- Turmel M, Otis C, Lemieux C (2003) The mitochondrial genome of *Chara vulgaris*: insights into the mitochondrial DNA architecture of the last common ancestor of green algae and land plants. *Plant Cell* 15:1888–1903
- Ullrich H, Lattig K, Brennicke A, Knoop V (1997) Mitochondrial DNA variations and nuclear RFLPs reflect different genetic similarities among 23 *Arabidopsis thaliana* ecotypes. *Plant Mol Biol* 33:37–45
- Unsold M, Marienfeld JR, Brandt P, Brennicke A (1997) The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924 nucleotides. *Nat Genet* 15:57–61
- Vangerow S, Teerkorn T, Knoop V (1999) Phylogenetic information in the mitochondrial *nad5* gene of pteridophytes: RNA editing and intron sequences. *Plant Biol* 1:235–243
- Vaughn JC, Mason MT, Sperwhitis GL, Kuhlman P, Palmer JD (1995) Fungal origin by horizontal transfer of a plant mitochondrial group-I intron in the chimeric *coxi* gene of *peperomia*. *J Mol Evol* 41:563–572
- Ward BL, Anderson RS, Bendich AJ (1981) The mitochondrial genome is large and variable in a family of plants (Cucurbitaceae). *Cell* 25:793–803
- Wellman CH, Gray J (2000) The microfossil record of early land plants. *Philos Trans R Soc Lond B Biol Sci* 355:717–731
- Wischmann C, Schuster W (1995) Transfer of *Rps10* from the mitochondrion to the nucleus in *Arabidopsis thaliana*. Evidence for RNA-mediated transfer and exon shuffling at the integration site. *FEBS Lett* 374:152–156
- Wissinger B, Schuster W, Brennicke A (1991) Trans splicing in *Oenothera* mitochondria: *nad1* mRNAs are edited in exon and trans-splicing group II intron sequences. *Cell* 65:473–482
- 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
- Wolff G, Burger G, Lang BF, Kuck U (1993) Mitochondrial genes in the colorless alga *Prototheca wickerhamii* resemble plant genes in their exons but fungal genes in their introns. *Nucleic Acids Res* 21:719–726
- Wolff G, Plante I, Lang BF, Kuck U, Burger G (1994) Complete sequence of the mitochondrial DNA of the chlorophyte alga *Prototheca wickerhamii*. Gene content and genome organization. *J Mol Biol* 237:75–86
- Won H, Renner SS (2003) Horizontal gene transfer from flowering plants to *Gnetum*. *Proc Natl Acad Sci USA* 100:10824–10829
- Yoshinaga K, Iinuma H, Masuzawa T, Uedal K (1996) Extensive RNA editing of U to C in addition to C to U substitution in the *rbcl* transcripts of hornwort chloroplasts and the origin of RNA editing in green plants. *Nucleic Acids Res* 24:1008–1014
- Zanis MJ, Soltis DE, Soltis PS, Mathews S, Donoghue MJ (2002) The root of the angiosperms revisited. *Proc Natl Acad Sci USA* 99:6848–6853