Advances in Plant Biology 1

For other titles published in this series, go to
www.springer.com/series/8047

Frank Kempken
Editor

Plant Mitochondria

Springer

Editor
Prof. Dr. Frank Kempken
Christian-Albrechts-Universität zu Kiel; Botanisches Institut
Olshausenstr. 40
24098 Kiel
Germany
fkempken@bot.uni-kiel.de

ISBN 978-0-387-89780-6 e-ISBN 978-0-387-89781-3
DOI 10.1007/978-0-387-89781-3
Springer New York Dordrecht Heidelberg London

This book is dedicated to Prof. Daryl R. Pring for his lifetime contributions to the field of plant mitochondria

Preface

Mitochondria are the product of a long evolutionary history. It is now a well-established fact that mitochondria did evolve from free living bacteria being the common ancestor of both eukaryotic mitochondria and α-proteobacteria. Advances in genome sequencing, the establishment of in organello and in vitro assays to name only a few, contributed significantly to advances in plant mitochondrial research. Second generation sequencing and the ability to directly sequence and analyze the whole plant transcriptome certainly will help to develop the research on plant mitochondria to another level in the future. In this book the current knowledge about plant mitochondria is presented in a series of detailed chapters, which have been organized in five main parts: (1) dynamics, genes and genomes; (2) transcription and RNA processing; (3) translation and import; (4) biochemistry, regulation and function; and (5) mitochondrial dysfunction and repair. These parts consist of two to five chapters, each written by well-known specialists in the field. The 19 chapters cover the field very well.

In Part I (dynamics, genes and genomes) Volker Knoop (Bonn, Germany) and coworkers provide an insight to the evolution of plant mitochondria which is discussed in the framework of our modern understanding of plant phylogeny. David C. Logan (Saskatchewan, Canada) together with Iain Scott (Bethesda, USA) discuss mitochondrial division and fusion as primary processes controlling mitochondrial form, size, and number. Sally A. Mackenzie together with Maria P. Arrieta-Montiel (Lincoln, USA) reports on the emerging mitochondrial sequence data from early land plants and recent studies of nuclear influence on mitochondrial genome behavior which have provided important insight into the evolutionary trends and possible rationale for the genomic variability that is seen in plant mitochondria.

In Part II (transcription and RNA processing) Thomas Börner and Karsten Liere (Berlin, Germany) discuss the current knowledge about plant mitochondrial transcription, which is sustained by phage-type RNA polymerases that are encoded by a small nuclear encoded gene family. Transcription of most mitochondrial genes is driven by multiple promoters, which may ensure transcription despite possible mitochondrial genome rearrangements. Stefan Binder (Ulm, Germany) and coworkers summarize recent progress made in the understanding of RNA processing and RNA

degradation in mitochondria of higher plants. In the complex framework of plant mitochondria posttranscriptional processes play predominant roles. Linda Bonen (Ottawa, Canada) presents recent advances in our understanding of splicing mechanisms, the nature of splicing machinery, and the relationships among splicing and other RNA processing events in plant mitochondria. Anika Bruhs (Kiel, Germany) and myself present an overview of plant mitochondrial (and plastid) RNA editing, its consequences for translation, the current knowledge of its mechanism, and some ideas on its evolution.

In Part III (translation and import) Nicolas L. Taylor and coworkers (Crawley, Australia) compare the proteomes of mitochondria from monocots and dicots plants and highlight the conservation of the mitochondrial electron transfer chain protein complex I. They also provide important insights, directions, and methodology currently utilized in their laboratory. Laurence Maréchal-Drouard (Strasbourg, France) and coworkers address the basic questions on the tRNA mitochondrial import selectivity, regulation, targeting, and translocation in plants. These data are discussed and compared to what has been discovered in tRNA mitochondrial import in evolutionary divergent organisms. Elzbieta Glaser and James Whelan (Stockholm, Sweden) give an overview on mitochondrial protein import in higher plants. Interestingly, many components of the plant protein import apparatus appear to be different to those in yeast and mammalian systems.

In Part IV (biochemistry, regulation and function), Hans-Peter Braun (Hanover, Germany) and coworkers summarize recent insights into the assembly of the OXPHOS system, consisting of five large multisubunit complexes within the inner mitochondrial membrane, the soluble intermembrane space-localized protein cytochrome c and the lipid ubiquinone. Allan G. Rasmusson (Lund, Sweden) together with Ian M. Møller (Aarhus, Denmark) look at the mitochondrial electron transport chain, which can mediate major adjustments in cellular metabolism important for cellular function under a great variety of stress conditions such as low temperature and drought. Keisuke Yoshida together with Ko Noguchi (Tokyo, Japan) describe and discuss the interaction between chloroplasts and mitochondria and review recent advances of understanding about the activity, function, and regulation of the mitochondrial respiratory system during photosynthesis. David M. Rhoads (Tucson, Arizona) describes plant mitochondria as stress sensors that contribute to decisions regarding cell fate during stresses. These are conveyed to the nucleus by mitochondrial retrograde regulation. Paul F. McCabe (Dublin, Ireland) together with Mark Diamond (Piscataway, USA) discuss the plant mitochondrion as a crucial mediator of programmed cell death. While similarities between plant and animal programmed cell death systems have been discovered, current knowledge suggests there are also key differences.

In Part V (mitochondrial dysfunction and repair) Jenny Carlsson together with Kristina Glimelius (Uppsala, Sweden) report on the current knowledge on cytoplamic male sterility (CMS) which is caused by mutations, rearrangements, or recombinations in the plant mitochondrial genome. CMS has important applied aspects in plant breeding and is the prime example of mitochondrial dysfunction in higher plants.

However, this book is intended not only for the specialist in plant mitochondria, but also for colleagues from related fields. As such I have invited a few chapters from authors that work on plastids or non-plant systems. These additional chapters provide very important insights into related areas which are not yet developed in plants or are important to understand the level of differences between the plant mitochondrial machinery and other model systems. It seemed prudent to invite Christian Schmitz-Linneweber (Berlin, Germany) and coworkers to add information on RNA binding proteins in plant chloroplast, as there are clear links between plant plastids and mitochondria, e.g., in the field of plant RNA editing. Wolfgang Voos (Bonn, Germany) and coworkers contribution about animal and fungal protein import is a quite logical extension to plant mitochondrial import. Readers can now directly compare fungal or animal protein import with that of plant mitochondria (written by Elzbieta Glaser and James Whelan). Finally, Susan LeDoux (Mobile, USA) provides important insight in human mitochondrial mutations and repair, a field of research which is somewhat related to cytoplasmic male sterility (written by Jenny Carlsson and Kristina Glimelius).

Graduate student and post-docs who recently joined a new field of research often find it difficult to access the literature. Book chapters and reviews written by specialists may contain a tremendous amount of information which literally may overwhelm the inexperienced reader. For easier access to this book, each of its chapters contains a list of abbreviations and a short glossary with explanations for important keywords. Also each chapter contains at least one text box, where more detailed information or cross-reference is given to specific subjects related to the chapter in question. I am sure these measures will not only aid the newcomers to access a complex field but also make the chapters more comprehensive to all readers.

Finally, I would like to thank all the many authors and coauthors who provided great chapters. It was great fun to get this book together. I also would like to thank Hannah Schorr and the publisher for their help and input. I truly hope that this book will help to attract young scientists to the fascinating and exiting field of plant mitochondria.

July 2010

Contents

Part I
Dynamics, Genes & Genomes

Chapter 1
Mitochondrial Genome Evolution in the Plant Lineage

Volker Knoop, Ute Volkmar, Julia Hecht, and Felix Grewe

Abstract Land plants feature particularly complicated mitochondrial genomes. Plant mitochondrial DNAs may be more than 100 times larger than those of animals and are structurally much more complex due to frequent ongoing recombination. The significant increase of plant mitochondrial genome sizes results from a combination of several factors: more genes are encoded, many of these carry introns and, most importantly, the plant mitochondrial genome has a propensity to accept foreign DNA sequences from the chloroplast, the nucleus, or even from other mitochondrial genomes via horizontal gene transfer. Similarly, plant mitochondria are also more complex on the transcriptome level where processes such as frequent RNA editing or *trans*-splicing of disrupted introns contribute to RNA maturation. The evolution of these peculiar features is discussed in the framework of our modern understanding of plant phylogeny to which mitochondrial genome data have contributed significantly.

Keywords Introns • RNA editing • Plant phylogeny • Mitochondrial evolution

Abbreviations

CMS Cytoplasmic male sterility
EGT Endosymbiotic gene transfer
HGT Horizontal gene transfer
LGT Lateral gene transfer

1.1 Introduction

The mitochondrial genomes of land plants (embryophytes) are characterized by a multitude of peculiarities that are counterintuitive to a general understanding of increasingly compacted genomes in endosymbiotic organelles. Some 500 million

V. Knoop (✉)
Abteilung Molekulare Evolution, Institut für Zelluläre und Molekulare Botanik, Universität Bonn, Kirschallee 1, Bonn, 53115, Germany
e-mail: Volker.knoop@uni-bonn.de

F. Kempken (ed.), *Plant Mitochondria*, Advances in Plant Biology 1,
DOI 10.1007/978-0-387-89781-3_1,

years of evolution have created very streamlined, compact and economically organized mitochondrial DNAs (mtDNAs) in the animal (metazoa) lineage only rarely exceeding some 16–17 kbp in size (Lavrov 2007; Gissi et al. 2008). In contrast, it seems that opposite trends rather complicating than simplifying mitochondrial genome structures have taken over during evolution in the plant lineage.

Several features of plant mitochondrial gene arrangements and expression will be dealt with in much detail in other Chapters of this book, e.g., recombinational activity (Chap. 3), intron splicing (Chap. 6), or RNA editing (Chap. 7). We will here mainly give a phylogenetic perspective on several peculiar features in the evolution of plant mtDNAs since Ordovician times when first plants appeared on land and gave rise to one of the most significant evolutionary transitions of multicellular life on this planet.

1.2 Land Plant Mitochondrial DNAs and Their Peculiarities

The first completed sequence of a mitochondrial DNA of a land plant, the one of the liverwort *Marchantia polymorpha* (Oda et al. 1992a, b) at a size of 186 kbp turned out to be more than 10 times larger than the first completely determined mtDNA sequence of the animal lineage, the one of *Homo sapiens* of only 16.6 kbp (Anderson et al. 1981). The *Marchantia* mtDNA revealed several genes not present in the mitochondrial genomes of animals. Moreover, the liverwort mtDNA contains seven introns of group I and 25 of group II type, typical intervening sequences frequently found in the organelle DNAs of fungi, algae, and plants (see Box 1.1 and Chap. 6). Only exceptional occasional discoveries of introns in mtDNAs of animals were made later, beginning with a first group I intron identified in the mtDNA of the sea anemone *Metridium senile* (Beagley et al. 1996). By and large, however, introns are rare in animal mtDNAs, mostly identified in some early branches of the metazoa phylogeny.

By the time the *Marchantia* chondrome was completely determined (note that the term chondriome is alternatively used to describe all mitochondria in a cell, see Chap. 2), it was already clear that yet more surprising features exist in the mitochondrial DNAs of flowering plants (angiosperms) and several others were discovered subsequently:

1. Plant mitochondrial genome sizes of 2 Mbp (mega base pairs = million base pairs) and more, hence exceeding the ones of several free-living bacteria, are present in the angiosperm family of Cucurbitaceae.
2. Endosymbiotic gene transfer (EGT), the functional transfer of genes to the nucleus, is an ongoing process in plant mitochondria in recent times of plant evolution.
3. Simple circular mitochondrial genomes such as the ones found in algae or bryophytes may be more of an exception than a rule in land plants. Active recombination across repeated sequences produces multipartite structure of plant mitochondrial genomes, at least in vascular plants (tracheophytes).
4. The average plant mitochondrial genome is characterized by some 20–30 introns. With a single exception these introns are stable within well-defined monophyletic plant clades, but differ significantly between them and have helped to elucidate land plant phylogeny. Some disrupted genes need to be reassembled on

RNA level via *trans*-splicing group II introns and an example of a *trans*-splicing group I introns has recently been identified in a lycophyte.

5. The peculiar process of RNA editing apparently repairs DNA coding information posttranscriptionally on RNA level by precise pyrimidine exchanges in plant mitochondrial RNAs.
6. Foreign DNA derived from the chloroplast or the nuclear genome is inserted into the mtDNA of tracheophytes as so-called "promiscuous DNA."
7. Several instances of DNA sequences in plant mitochondria acquired by horizontal gene transfer (HGT) across species barriers and over wide phylogenetic distances have recently been reported.

Box 1.1 Group I and group II introns

These two types of introns are defined through their characteristic secondary structures and conserved RNA sequence motifs (see, e.g., Bonen 2008; Cech et al. 1994; Haugen et al. 2005; Pyle et al. 2007 for recent summary updates). Numerous group I and group II introns are typically present in the organelle genomes of plants and fungi. A nomenclature to label them according to their insertion sites has been proposed (Knoop 2004; Dombrovska and Qiu 2004) as detailed in the legend of Fig. 1.2. Group I and group II introns have also been identified in mitochondria of basal metazoan lineages (see, e.g., Fukami et al. 2007), protists, phages, and more frequently also bacteria (Tourasse and Kolstø 2008). Introns of both classes are often referred to as "autocatalytic," "self-splicing," "ribozymes," or "mobile," but this is in fact biochemically demonstrated only for some of them. Some intron species of both classes carry open reading frames (ORFs), which participate in intron mobility or splicing (see Chap. 6): endonucleases in group I introns and the so-called maturases in group II introns. Intron-encoded ORFs are absent in the overwhelming majority of seed plant chloroplast and mitochondrial introns with singular exceptions in both cases: nad1i782g2 in mitochondria carrying *mat*R and trnK37g2 in chloroplasts carrying *mat*K. Absence of other intron-encoded ORFs may explain both the obvious positional stasis of seed plant organelle introns since more than 200 million years of evolution as well as their reliance on nuclear encoded cofactors for splicing (see Chap. 8), which have been identified over the last years (recent examples include Asakura and Barkan 2006; Asakura and Barkan 2007; Keren et al. 2009; Kroeger et al. 2009; Ostheimer et al. 2006; Schmitz-Linneweber et al. 2006; Watkins et al. 2007). Several examples for disrupted group II introns, which require *trans*-splicing of independent primary transcripts for RNA maturation, have been identified in chloroplast and mitochondrial DNAs since quite some time (Bonen 2008; Glanz and Kück 2009). Examples for *trans*-splicing group I introns, however, have only been discovered very recently (Burger et al. 2009; Grewe et al. 2009). For more details see Chap. 6.

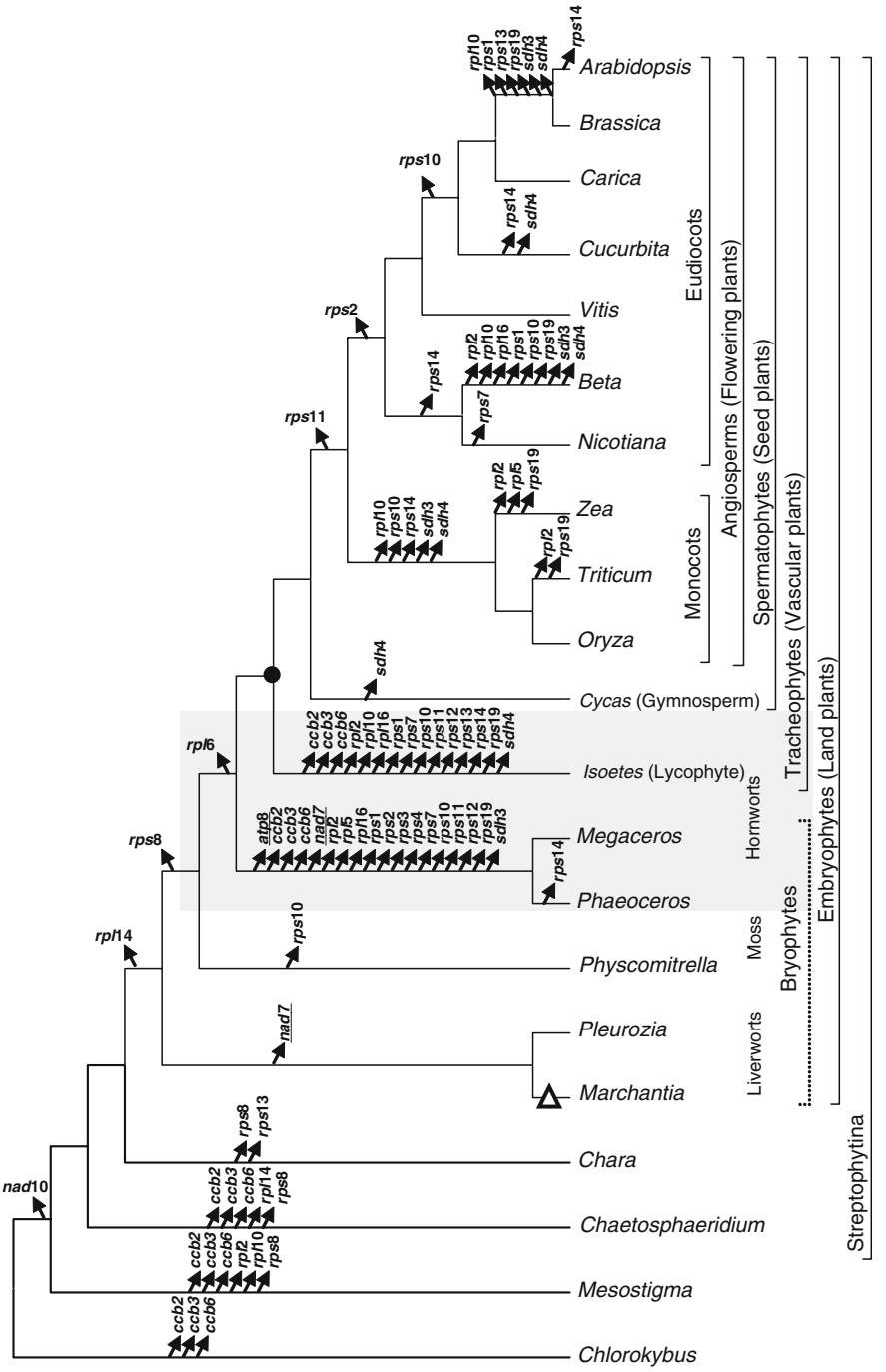

Fig. 1.1 Streptophyte genera with completed mitochondrial genome sequences currently available (references given in the text except for the *Carica papaya* mtDNA, as yet only available as database accession EU431224) are assembled into a cladogram based on modern insights of plant phylogeny (Qiu et al. 2006). The mtDNAs of taxa with gene complements identical to closely related genera included in the figure (e.g., *Citrullus, Tripsacum, Sorghum*) were omitted for clarity. No complete

1.3 Plant Mitochondrial Genomes: Completed MtDNA Sequences

The expanded sizes and genomic complexities of plant mitochondrial DNAs probably explain why only so few mtDNAs of plants have been determined in comparison to the numerous completely sequenced mtDNAs of animals. The first complete mtDNA of a flowering plant, the model angiosperm *Arabidopsis thaliana* (Unseld et al. 1997), was followed by those of more economically important plants such as sugar beet *Beta vulgaris* (Kubo et al. 2000), rice *Oryza sativa* (Notsu et al. 2002), rapeseed *Brassica napus* (Handa 2003), maize *Zea mays* (Clifton et al. 2004), tobacco *Nicotiana tabacum* (Sugiyama et al. 2005), wheat *Triticum aestivum* (Ogihara et al. 2005) and grape *Vitis vinifera* (Goremykin et al. 2009). Finally the mtDNAs from watermelon *Citrullus lanatus* and zucchini (*Cucurbita pepo*) have been determined recently (Alverson et al. 2010), two species of the Cucurbitaceae long known to harbor particularly large chondromes (possibly exceeding 2 Mbp in muskmelon Ward et al. 1981). The availability of complete mtDNAs from different strains or cultivars of the same species (e.g., sugar beet, Satoh et al. 2004; maize, Allen et al. 2007), allows insights on the mtDNA rearrangements accompanying important mitochondrially encoded traits such as cytoplasmic male sterility (CMS, see, e.g., Kubo and Newton 2008 see Chap. 18).

In contrast, only a single gymnosperm mtDNA sequence, the one of the cycad (fern palm) *Cycas taitungensis,* has so far been determined (Chaw et al. 2008). Similarly, it took more than a decade after determination of the *Marchantia polymorpha* mtDNA (Oda et al. 1992a, b) before a second bryophyte mtDNA became available, the one of the model moss *Physcomitrella patens* (Terasawa et al. 2006). The chondrome sequences of two representatives of the third clade of bryophytes, the hornworts, have been determined only very recently with the mtDNAs of *Megaceros aenigmaticus* (Li et al. 2009) and *Phaeoceros laevis* (Xue et al. 2010). Finally, the mtDNA of a second liverwort, *Pleurozia purpurea,* has recently been published (Wang et al. 2009). This currently leaves us with a considerable phylogenetic gap between the bryophytes and the seed plants for mitochondrial genome sequences of early branching (i.e., non-seed) tracheophyte lineages (Fig. 1.1).

Fig. 1.1 (continued) mtDNA sequence for a representative of the monilophytes (ferns *sensu lato,* including whisk ferns and horsetails) is currently available and only singular mitochondrial genome sequences are so far determined for mosses (*Physcomitrella*), lycophytes (*Isoetes*), and gymnosperms (*Cycas*). Taxonomic labeling is indicated to the right. Note that "bryophytes" are paraphyletic (*dotted line*). Losses of functional genes from the respective mtDNAs are indicated by arrows on the branches of the model phylogeny. Losses of ribosomal proteins (*rpl*, *rps*), subunits of succinate dehydrogenase (*sdh*), and cytochrome c biogenesis (*ccb*) occur independently multiple times. Conversely, independent losses of core respiratory subunits are rare (*atp*8 and *nad*7, *underlined*). Sequential losses of *rpl*14, *rps*8, and *rpl*6 from the mtDNA deep along the tree's backbone can be taken as support for this model embryophyte phylogeny. RNA editing emerges with embryophytes and is secondarily lost in marchantiid liverworts (*open triangle*). The bryophyte-tracheophyte transition zone (*gray shading*) is characterized by multiple parallel losses of mt genes and an increase of reverse (U to C) RNA editing. A strong increase in DNA recombinational activity in the plant mitochondrial genomes and the integration of promiscuous DNA originating from the chloroplast and nucleus apparently emerges with the tracheophyte lineage (*black dot*)

These lineages, the "pteridophytes" *sensu lato* (ferns and "fern allies"), are now unequivocally distinguished into lycophytes and monilophytes (Pryer et al. 2001). Only the latter is the sister clade to seed plants and comprises the true ferns, the horsetails (*Equisetum*) and the whisk ferns (*Psilotum, Tmesipteris*). Reports on the mtDNAs of ferns and "fern allies" have been very rare (Palmer et al. 1992). As a first mtDNA of this group the one of the quillwort *Isoetes engelmannii* has been reported very recently (Grewe et al. 2009).

A monophyletic group of land plants and related green algal lineages has been recognized as the clade of "streptophytes" (see Fig. 1.1). The available mtDNAs of several streptophyte algae are a very important addition to the collection of embryophyte mtDNAs for insights on phylogeny and molecular evolution in the green lineage: *Chaetosphaeridium globosum* (Turmel et al. 2002a), *Mesostigma viride* (Turmel et al. 2002b), *Chara vulgaris* (Turmel et al. 2003), and *Chlorokybus atmophyticus* (Turmel et al. 2007). Whereas the former three algal mtDNAs are of moderate sizes between 42 and 68 kbp, the mitochondrial DNA of *Chlorokybus* at 202 kbp even exceeds the *Marchantia* mtDNA of 186 kbp in size. As novelties among streptophyte mtDNAs the *Chlorokybus* mtDNA features a *nad*10 and a *trn*L(gag) gene and 13 introns at novel insertion sites with a surprising number of 10 group II introns in tRNA-encoding genes. Somewhat puzzling is the mtDNA of the tiny unicellular chlorophyte green alga *Ostreococcus tauri*: On the one hand the 44 kbp mtDNA is very gene dense, on the other hand it consists of a large segmental duplication of 44% of the chondrome (Robbens et al. 2007).

1.4 Ongoing Gene Transfer to the Nucleus

The functional transfer of genes from an endosymbiont genome to the nuclear genome of the host is frequently referred to as endosymbiotic gene transfer (EGT, see Timmis et al. 2004). An apparent stasis of EGT has been reached in the animal lineage, where largely the same (small) gene complement is conserved across phylogenetic distances dating back some 500 million years. Animal mtDNAs usually encode 22 tRNAs, 2 rRNAs, and 13 genes encoding proteins of the respiratory chain complexes I, III, IV, and V (the NADH dehydrogenase, apocytochrome B, the cytochrome c oxidase, and the ATPase): *nad*1-6 and *nad*4L, *cob*, *cox*1-3, and *atp*6 and 8. Typically, the gene complement of plant mtDNAs is extended by genes for additional subunits of these respiratory chain complexes (*atp*1, 4, and 9, *nad*7 and 9) and for two subunits of complex II, the succinate dehydrogenase (*sdh*3 and 4). Moreover, *ccb* genes encoding proteins of cytochrome c biogenesis, *rps* and *rpl* genes encoding proteins of the small and large ribosomal subunit and *tat*C, a transport protein subunit (twin arginine translocase), are encoded in plant mtDNAs. Plant mitochondrial gene complement varies widely among angiosperms in particular for *rps, rpl* and *sdh* genes (see Fig. 1.1) showing that endosymbiotic gene transfer is an ongoing process also in

recent times of plant evolution (Liu et al. 2009). The identification of genes in the mitochondrial DNA of one but not in another plant lineage has generally been taken as a first hint for such recent EGT. The *rps*10 gene was a typical early example along such lines (Knoop et al. 1995), and then found to be transferred to the nucleus independently in many plant lineages (Adams et al. 2000). The same point was made for the *sdh* genes (Adams et al. 2001b) and the *rpl* and *rps* genes similarly turned out to be transferred to the nucleus via EGT many times independently in a systematic survey of angiosperms (Palmer et al. 2000; Adams et al. 2002b; Adams and Palmer 2003).

In contrast, only three examples have as yet been identified for functional EGT of any of the core components of respiratory chain complexes I–V during land plant evolution. The example of *cox*2 transfer among legumes (Fabales) is particularly intriguing given that the necessary steps of nuclear copy establishment, activation of the nuclear copy, and subsequent defunctionalization and disintegration of the mitochondrial copy (Brennicke et al. 1993) can be traced among closely related taxa of this clade (Nugent and Palmer 1991; Covello and Gray 1992; Adams et al. 1999). Secondly, functional *atp*8 genes are apparently lost from *Allium* (Adams et al. 2002b) and are also missing from the two completed hornwort mtDNAs (Li et al. 2009; Xue et al. 2010). It will be interesting to see whether EGT of *atp*8 may turn out as an apomorphy of the entire hornwort clade.

Finally, a third example of gene transfer concerns the interesting issue of the *nad*7 gene in liverworts (Kobayashi et al. 1997), which appears to be a very ancient event in this plant clade. Once a gene is transferred via EGT it normally disintegrates quickly and the identification of a recognizable pseudogene copy in the mitochondrial genome can be taken as evidence for evolutionary recent gene transfer. Strikingly, however, *nad*7 remains present in the liverwort mtDNAs as a pseudogene after functional establishment of the nuclear *nad*7 copy instead of quick pseudogene disintegration as generally observed in the angiosperms (Groth-Malonek et al. 2007b). Whether this long-term pseudogene survival simply reflects particular slow structural mtDNA evolution due to lack of mtDNA recombination in the liverworts (see below) or a remaining functional necessity of the pseudogene in liverworts remains to be seen. The former possibility is supported by the extreme degrees of gene synteny between the two fully sequenced liverwort mtDNAs, in fact also in comparison to moss and hornwort mtDNAs (Wang et al. 2009; Xue et al. 2010). The latter idea of remaining functionality in a pseudogene retained in mitochondria is supported through the second known example for exceptional long-term retention of a pseudogene after EGT, in this case among angiosperms: an *rps*14 pseudogene is conserved for some 80 million years in the mtDNAs of grasses (Ong and Palmer 2006).

Similar to *atp*8, independent EGT of *nad*7 has obviously occurred in the hornworts (see Fig. 1.1). In fact, the recently determined complete mtDNA sequences of two hornworts (Li et al. 2009; Xue et al. 2010) show many pseudogenes or recognizable pseudogene fragments of mitochondrial genes such as the *ccb* genes and most of the *rpl* and *rps* genes, which suggests an increase of EGT activity in

the hornwort lineage. Surprisingly, the *ccb* genes and many *rpl* and *rps* genes are also lacking in the lycophyte *Isoetes* mtDNA as well (Grewe et al. 2009). Given our current understanding of embryophyte phylogeny (see Fig. 1.1) a rise in EGT activity may have taken place at the time of bryophyte-tracheophyte transition in land plant evolution. Nevertheless, these gene losses need to be explained as independent events in the hornwort and tracheophyte lineages given that the affected genes are present in seed plants.

Obviously the genes known to be subject to frequent, recent, and independent EGT in embryophytes (*ccb, rpl, rps, sdh*) are the ones that are absent from the modern metazoan mtDNAs altogether, suggesting that similar functional selective pressures exist for retention of genes in the organelles. It is striking to see that the *tat*C gene so far not found subject to recent EGT in the embryophyte lineage has recently also been identified in a very basal metazoan lineage, the sponge *Oscarella* (Wang and Lavrov 2007).

There is good evidence from the study of EGT in plants that functional gene transfer is mediated by RNA (e.g., Covello and Gray 1992; Nugent and Palmer 1991; Wischmann and Schuster 1995). Functional nuclear genes are inserted in the RNA-edited version and certainly lack the organellar introns. The RNA-mediated EGT serving to functionally migrate a gene from mtDNA to the nucleus is, however, accompanied by DNA-based transfer of mitochondrial sequences during evolution, mostly without any discernible function. Small fragments of mitochondrial and plastid DNA (called "numts" and "nupts") have been identified in many nuclear genomes of diverse eukaryotes. In fact, a full mtDNA sequence copy is even present on chromosome 2 in the nucleus of *Arabidopsis* (Lin et al. 1999; Stupar et al. 2001). Exceptionally such insertions of organelle DNA fragments may possibly contribute new functionality in their new nuclear location such as the insertion of a group II intron fragment in one of two lectin genes in *Dolichos biflorus* (Knoop and Brennicke 1991).

Several examples of EGT show that mitochondrial genes must not necessarily be transferred to the nucleus in a 1:1 fashion. In some cases genes are functionally split upon EGT: the transfer of *rpl*2 in flowering plants (Adams et al. 2001a) or of *cox*2 in chlamydomonad algae (Perez-Martinez et al. 2001) are examples. Other particular cases of gene fissions upon EGT are observed in protist groups. The *sdh*2 gene is split in two parts in Euglenozoa upon relocation to the nucleus with the protein subunits independently reimported into the mitochondria (Gawryluk and Gray 2009). In the case of *cox*1 only the C-terminal portion is nuclear encoded in diverse protist groups, whereas the truncated gene remained in the mitochondria (Gawryluk and Gray 2010).

Secondly, a gene may be functionally substituted such as *rps*13 apparently replaced upon EGT by an aminoterminal extension of *rps*19 in *Arabidopsis* (Sanchez et al. 1996). Thirdly, a mitochondrial gene may have been replaced by a homologous gene originating from chloroplast or nuclear DNA as was shown for *rps*13 and *rps*8, respectively (Adams et al. 2002a). Obviously, this substitution process can also work in the opposite direction with mitochondrial genes replacing chloroplast counterparts as was shown for *rps*16 (Ueda et al. 2008).

1.5 Plant Mitochondrial Genomes: Structures

The notion that "plant mitochondrial DNA evolves rapidly in structure, but slowly in sequence" (Palmer and Herbon 1988) still holds generally true. Accelerated primary sequence drift has, however, been observed in angiosperm genera such as *Silene* (Sloan et al. 2009), *Pelargonium,* and *Plantago* (Palmer et al. 2000; Cho et al. 2004; Parkinson et al. 2005), or the fern genus *Lygodium* (Vangerow et al. 1999). Efficient DNA mutation repair mechanisms that are currently being elucidated (Boesch et al. 2009) probably contribute to the generally slow primary sequence drift in plant mitochondrial sequences.

Like in the liverwort *Marchantia*, simple circular mtDNA molecules were deduced as mitochondrial genomes also for the moss, for the hornworts and the streptophyte algae. It has been pointed out though that many simple, circular mtDNAs may actually just be an illusion of mapping studies and the corresponding genomes could in fact rather be populations of coexisting, overlapping, linear, or branched molecules (Oldenburg and Bendich 2001; Bendich 2007).

Recombinationally active mtDNAs were found in angiosperms, resulting in dynamic, multipartite and complex structures of the flowering plant mitochondrial genomes (e.g., Ogihara et al. 2005; Palmer and Shields 1984; Sugiyama et al. 2005). Variations of mitochondrial genome structure may even be observed within different ecotypes, isolates, or strains of single species such as *Arabidopsis* or maize (Ullrich et al. 1997; Allen et al. 2007). Active DNA recombination in flowering plant mitochondria leads to populations of coexisting genomic rearrangements of shifting stoichiometry and possibly also to heteroplasmy (see Woloszynska 2010 for a recent review).

Repeated DNA sequences obviously play a major role in generating plant mitochondrial genome variabilities (Andre et al. 1992; Lilly and Havey 2001; Kubo and Mikami 2007), and it may be assumed that homologous sequences need to have a certain length to serve as substrates for homologous recombination. Small repeated sequences not participating in recombination are particularly striking in the *Cycas taitungensis* chondrome. Some 500 sequence elements of 36 bp in the *Cycas* mtDNA are characterized by the terminal direct repeats of the AAGG tetranucleotide and the internal recognition sites Bpu10I which suggested the label "Bpu sequences" (Chaw et al. 2008). On the other hand, much longer sequence repeats are not per se targets for active recombination. An example is a large part of a *cob* gene intron (cobi783g2)[1] copied into an intergenic region in the mtDNA of *Marchantia* and other liverworts (Groth-Malonek et al. 2007a), without any indication for actively ongoing recombination producing alternative mtDNA gene arrangements (Oda et al. 1992a).

The mitochondrial DNA of the lycophyte *Isoetes engelmannii* in contrast features at least two dozen DNA recombination sites resulting in alternative, coexisting genomic arrangements (Grewe et al. 2009). Similarly, we find strong evidence

[1]Intron nomenclature consists of gene name followed by the letter i, the number of the nucleotide in the mature reading frame preceding the insertion site and a qualifier designating groupI/II introns as g1/2. See Box 1.1 for details.

for mtDNA recombination in the lycophyte sister lineage *Selaginella* (J.H., F.G., and V.K., unpublished observations). Hence, in the light of conserved circular chondromes in bryophytes, mtDNA recombination is obviously a gain in the earliest tracheophyte lineages (see Fig. 1.1). Interestingly, though, in contrast to the dynamically rearranging chondromes of the Brassicaceae (Palmer and Shields 1984; Unseld et al. 1997; Handa 2003), the counter-example of a simple nonrecombining circular mtDNA structure in *Brassica hirta* also exists, even in the same plant family (Palmer and Herbon 1987), suggesting that the recombinational activity may be secondarily lost in evolution.

Nuclear factors involved in plant mtDNA recombination are increasingly being identified (e.g., Manchekar et al. 2006; Shedge et al. 2007; Odahara et al. 2009). Another chapter in this volume will be entirely devoted to the functional analysis of DNA recombinational dynamics of plant mitochondrial genomes (see Chap. 2).

1.6 The Introns in Embryophyte Mitochondrial DNAs

In spite of their many idiosyncrasies, plant mitochondrial genome data have proved useful to help in the phylogenetic analysis of embryophytes. Not only coding sequences but also introns, non-coding intergenic regions, gene rearrangements and a pseudogene have been exploited as phylogenetically informative regions (see Knoop 2010).

The occurrence of mitochondrial group I and group II introns (see Box 1.1 and Chap. 6) differ significantly between land plant clades. So far, some 100 different intron insertion sites have been identified in embryophyte mitochondrial genes (Fig. 1.2). Their appearances have been taken as evidence for a sister group relationship of liverworts to all other land plants (Qiu et al. 1998), as well as for a sister group relationship of hornworts and tracheophytes (Groth-Malonek et al. 2005, see Fig. 1.1). The majority of group I and group II introns appear to be rather stable in presence within a given group, in fact can be typical signatures of a given plant clade (Vangerow et al. 1999; Pruchner et al. 2001; Pruchner et al. 2002; Dombrovska and Qiu 2004). Consequently, such introns can be useful loci for phylogenetic analyses in a given clade and several of them have been used in that way (e.g., Beckert et al. 1999; Beckert et al. 2001; Wikström and Pryer 2005; Volkmar and Knoop 2010; Wahrmund et al. 2010). Rare variability of mitochondrial intron presence among angiosperms (see Fig. 1.2) may in fact suggest that independent intron losses occur even less frequently than total gene losses after EGT, but final conclusions need more systematic and extensive taxon sampling. Most likely, the overall loss of intron-encoded ORFs involved in intron mobility (see Box 1.1) in seed plant mtDNAs plays a major role in determining intron position stabilities. Interestingly, of 20 mitochondrial introns among angiosperms, 19 are also conserved in the mtDNA of the gymnosperm *Cycas taitungensis* (Chaw et al. 2008). Recent studies of the *rps*3 gene in a wide sampling of gymnosperms (featuring intron rps3i257g2[1] absent in angiosperms) suggest that independent

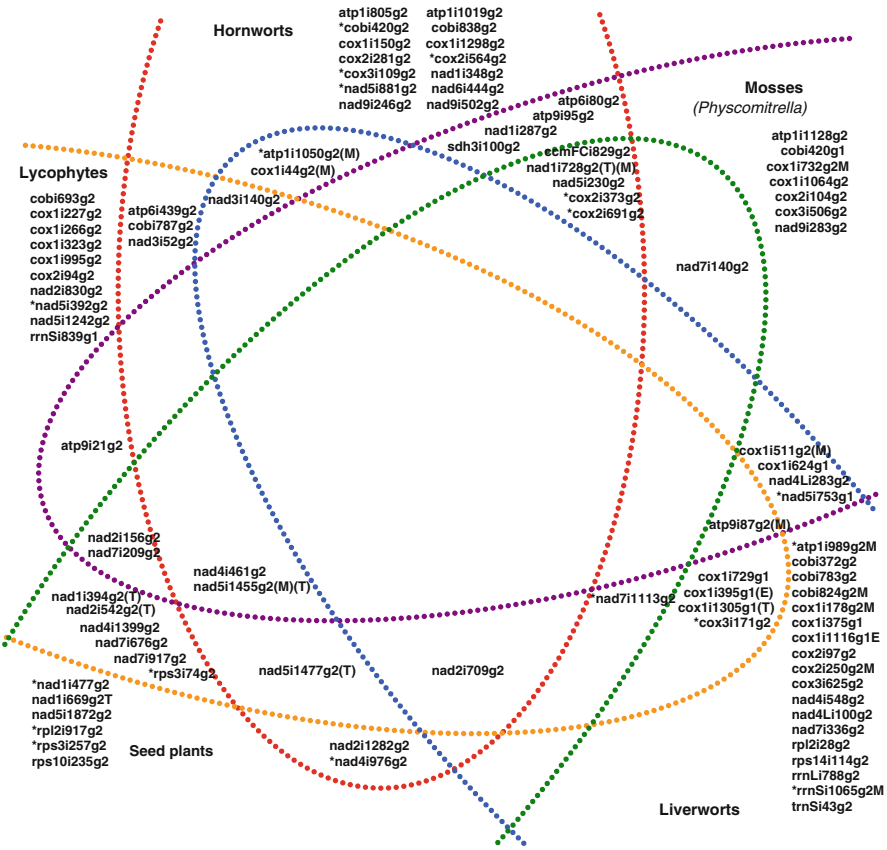

Fig. 1.2 Venn diagram showing the occurrences of 97 mitochondrial introns in the major plant clades, for which complete mitochondrial genome sequences are currently available. The intron nomenclature designates the respective gene, the coding nucleotide number preceding the intron insertion site and the type of intron (g1/g2) as previously proposed (Dombrovska and Qiu 2004; Knoop 2004). The nomenclature is here suggested to be extended by qualifiers "T" for *trans*-splicing and "M" or "E" for intron-borne maturase or endonuclease ORFs, respectively. Known alternative states of *cis*- vs. *trans*-splicing or presence of introns-borne ORFs are indicated with brackets. Variability of intron presence mostly indicating occasional secondary losses within a given group is indicated with an asterisk, e.g., nad1i477g2 absent in Geraniaceae (Bakker et al. 2000), atp1i989g2, and atp1i1050g2 absent in *Treubia lacunosa* (database accession DQ646161), nad5i753g1 absent in *Takakia ceratophylla* (database accession DQ268963), and rps3i257g2 hitherto only identified in some gymnosperms (Ran et al. 2010). Other cases rather suggest rare individual gains: nad5i392g2 in *Huperzia* (Vangerow et al. 1999), nad5i881g2 in *Anthoceros* (Beckert et al. 1999), or cox3i171g2 in *Lycopodium* (Hiesel et al. 1994). Not included in the diagram are introns that are part of promiscuous DNA originating from the chloroplast such as trnVi39g2 in *Cycas* and introns atp1i361g2, cox1i747g2, and nad2i258g2 so far identified only in individual monilophyte taxa (Dombrovska and Qiu 2004; Wikström and Pryer 2005; Bégu and Araya 2009), for which so far no complete mtDNA sequence has been determined

mitochondrial intron losses can occur through retro-processing (Ran et al. 2010) but may be similarly rare as among angiosperms.

As a so far unique exception for the generally conserved presence of plant mitochondrial introns within a given clade, one particular group I intron in the *cox*1 gene, cox1i729g1, originally identified in the angiosperm *Peperomia polybotrya* (Vaughn et al. 1995), has subsequently also been found with a patchy distribution among distant angiosperms. Intron cox1i729g1 has been suggested to originate from a fungal donor source and to invade angiosperm mitochondrial genomes independently (Adams et al. 1998; Cho et al. 1998). A recent reinvestigation with large angiosperm taxon sampling and careful phylogenetic analyses has questioned this idea and instead found that rare gains, if not a single unique gain followed by numerous independent losses are a more satisfying explanation (Cusimano et al. 2008). Yet further taxon sampling, however, reinforced the idea of multiple independent gains of cox1i729g1, now assuming numerous horizontal gene transfers across plant species borders (see below) instead of independent acquisitions from a fungal donor (Sanchez-Puerta et al. 2008).

The mRNAs of three genes in angiosperm mtDNAs – *nad*1, *nad*2, and *nad*5 – are disrupted and need to be assembled via *trans*-splicing group II introns (Chapdelaine and Bonen 1991; Knoop et al. 1991; Pereira de Souza et al. 1991; Wissinger et al. 1991; Binder et al. 1992). Five ancient *trans*-splicing introns in these genes (nad1i394g2, nad1i669g2, nad2i542g2, nad5i1455g2 and nad5i1477g2) are conserved in the gymnosperm *Cycas taitungensis* (Chaw et al. 2008) and their evolutionary histories had earlier been traced back to *cis*-splicing ancestor present in pteridophytes, hornworts, and mosses (Malek et al. 1997; Malek and Knoop 1998; Groth-Malonek et al. 2005). While the disintegrations of these five introns may have been very rare, possibly even singular events, the disruption of intron nad1i728g2 into *trans*-arrangements occurred later in angiosperm evolution multiple times independently (Qiu and Palmer 2004). Obviously the gain of recombinational activity in tracheophyte evolution has played a major role in evolving *trans*-splicing introns. Taken together with the fact that novel group II introns frequently show up during plant evolution (see Fig. 1.2), whereas group I intron history is largely dominated by losses in plant mitochondrial DNA, this may explain why *trans*-splicing group II introns but so far no *trans*-splicing group I introns were discovered. Finally, a first *trans*-splicing group I intron has recently been identified in *Isoetes engelmannii* (Grewe et al. 2009). Preliminary data indicate that yet more *trans*-splicing introns will be discovered in the mtDNA of the sister lycophyte *Selaginella* (J.H, F.G., and V.K., unpublished observations).

1.7 RNA Editing

In plant mitochondria, RNA editing (Box 1.2 and see Chap. 7) comes in the form of pyrimidine nucleotide conversions from cytidine to uridine and in the reverse direction. The main task of RNA editing in plants largely appears to be the restoration of evolutionary conserved codon identities to create functional proteins, which may

Box 1.2 RNA editing

RNA editing defines any changes in the nucleotide sequence of a transcript that results in a nucleotide sequence which could principally be directly encoded in the DNA template. As such it encompasses the insertions, deletions and exchanges of any of the four standard nucleotides of the RNA alphabet (A: adenosine, C: cytidine, G: guanidine, and U: uridine) in a RNA molecule. The definition also includes the deamination of adenosine (A) to inosine (I) nucleotides, which are read as guanidine (G) nucleotides upon translation but excludes chemical modifications leading to nonstandard bases such as for example dihydro-uridine or pseudo-uridine in tRNAs. The term RNA editing was originally coined for the uridine insertion type of RNA editing discovered in trypanosome mitochondria (Benne et al. 1986). Many other types of mitochondrial, cytosolic and chloroplast RNA editing have been discovered subsequently (Knoop and Brennicke 1999). A nomenclature has recently been suggested to label RNA editing sites (Rüdinger et al. 2009) and the software tool PREPACT has been developed to analyze and catalogue the abundant RNA editing sites in plant organelles (Lenz et al. 2009). For more details see Chap. 7.

include the creation of start and stop codons of the respective reading frames. Somewhat more rarely, RNA editing in structural RNAs such as introns or tRNAs may help to stabilize or reestablish base-pairings. Deamination and/or transamination of the pyrimidine bases are most likely the underlying biochemical mechanisms for base conversion. Given that one chapter of this book (see Chap. 7) is exclusively devoted to plant mitochondrial RNA editing, we will here mainly focus on the phylogenetic perspective of RNA editing evolution among embryophytes.

RNA editing in flowering plant chloroplasts (Hoch et al. 1991) was discovered only briefly after its identification in mitochondria (Covello and Gray 1989; Gualberto et al. 1989; Hiesel et al. 1989). Phylogenetic studies subsequently showed that the occurrence of RNA editing in mitochondria and chloroplasts is surprisingly congruent: RNA editing appears to be absent in algae but has been identified in chloroplast and mitochondria of all land plant clades with the unique exception of the marchantiid liverworts (Malek et al. 1996; Freyer et al. 1997; Steinhauser et al. 1999). Similarly the C-to-U type of editing is dominating in both organelles among seed plants, mosses and (nonmarchantiid) liverworts, whereas frequent "reverse" U-to-C editing is only seen in hornworts, lycophytes, and ferns (see Fig. 1.1) in chloroplasts and mitochondria at the same time (Steinhauser et al. 1999; Vangerow et al. 1999; Kugita et al. 2003; Wolf et al. 2004; Grewe et al. 2009). Mitochondrial RNA editing frequencies vary widely across land plant phylogeny, ranging from zero sites in marchantiid liverworts over very few, e.g., only 11 in the moss *Physcomitrella* (Rüdinger et al. 2009) to predictions of over 1,000 RNA editing sites in the gymnosperm *Cycas* or the lycophyte *Isoetes* (Chaw et al. 2008; Grewe et al. 2009).

Biochemical and genetic functional studies suggest very similar mechanisms for recognition of editing sites in transcripts of mitochondria and chloroplasts. In particular, specific members of the vastly extended pentatricopeptide repeat protein (PPR) gene families of plants (Lurin et al. 2004) have been shown to be involved in RNA editing site recognition both in chloroplasts (Kotera et al. 2005) and in mitochondria (Zehrmann et al. 2009). A particular sub-group of PPR proteins with a carboxy-terminal extension ending in the highly conserved DYW tripeptide had previously been connected to RNA editing in plants. The "DYW domain" shows similarity to cytidine deaminases potentially indicating direct involvement in the biochemical process of base conversion through deamination of cytidine to uridine and DYW domain PPR proteins are consistently present in taxa showing RNA editing but could not be identified in those lacking RNA editing (Salone et al. 2007; Rüdinger et al. 2008). Indeed, there seems to be a quantitative correlation between the numbers of organellar editing sites and nuclear DYW-type PPR genes (Rüdinger et al. 2008; Rüdinger et al. 2009). Nevertheless, the existence of plant organellar RNA editing as such remains a mystery given that no convincing evidence for functional gains such as physiological regulation of protein activity or the creation of protein diversity through partial editing (e.g., Mower and Palmer 2006) has been found. RNA editing may largely be a correction mechanism, possibly compensating on RNA level for mutations occurring on DNA level associated with the establishment of embryophytes. The wide variability of editing frequencies including the complete absence in marchantiid liverworts and the rise in reverse U-to-C editing activity in hornworts, lycophytes and monilophytes (see Fig. 1.1) remain, however, mysterious at present. Given that establishment of plant RNA editing obviously comes at the cost of creation and maintenance of large nuclear gene families such as the PPR gene family the phenomenon remains all the more puzzling.

1.8 Gene Transfer Deviations: Promiscuous DNA

Besides EGT other forms of interorganellar gene transfer have been identified as well in plant cells. Flowering plant mitochondrial genomes have been shown to have a surprising disposition to integrate and perpetuate foreign DNA from the chloroplast or nuclear genomes. The initial discovery of a 12 kbp fragment of cpDNA in the maize mitochondrial genome (Stern and Lonsdale 1982) led J. Ellis to suggest the term "promiscuous DNA" at that time (Ellis 1982). Not only chloroplast DNA, but also promiscuous DNA fragments of nuclear origin were subsequently observed in many angiosperm mtDNAs (Schuster and Brennicke 1987), e.g., all different types of retrotransposon fragments in *Arabidopsis* (Knoop et al. 1996). Similarly, sequence inserts originating from the chloroplast and nuclear genome have also been found in the mtDNA of the gymnosperm *Cycas* (Wang et al. 2007). Finally, chloroplast and nuclear promiscuous DNA fragments have been identified in the mtDNA of the lycophyte *Isoetes* (Grewe et al. 2009), showing that the propensity for integrating foreign DNA arose with early tracheophyte evolution.

Notably, an active mechanism of DNA import in potato mitochondria has been described (Koulintchenko et al. 2003), the physiologic relevance of which is unclear at present. In contrast to tracheophytes, no promiscuous DNA has as yet been identified in the mtDNAs of bryophytes. The ability to integrate promiscuous sequences may well depend on the gain of recombinational activity in the tracheophyte lineage (see Fig. 1.1).

1.9 Horizontal Gene Transfer

A series of publications has shown that plant mitochondrial genomes may be the prime examples of donors and acceptors of DNA via frequent horizontal gene transfer (HGT) among eukaryotes (Box 1.3). Evidence for DNA transfer across species borders in plant mitochondria initially came from the surprising identification of genes in certain species that had apparently been regained in the mitochondria after they had previously been transferred to the nucleus via EGT in the respective plant lineage (Bergthorsson et al. 2003). Host–parasite interactions seem to play a major role for mitochondrial HGT in plants (Davis and Wurdack 2004; Mower et al. 2004). Not only angiosperms, but also the gymnosperm *Gnetum* (Won and Renner 2003) and the fern *Botrychium* (Davis et al. 2005) have been identified as acceptor species for mitochondrial DNA transferred from other taxa via HGT. A particularly striking example seems to be the case of the mtDNA in the basal angiosperm *Amborella trichopoda,* which contains many gene sequences of foreign, including bryophyte, origins (Bergthorsson et al. 2004; J. Palmer, personal communication). Intimate physical plant-plant contacts that allow the exchange of DNA into cells

Box 1.3 Horizontal gene transfer

Horizontal gene transfer (HGT) is defined as the migration of DNA sequences across species borders or alternatively as the nonsexual movement of genetic information between two organisms (Keeling and Palmer 2008; Bock 2010). Hence, HGT is an exception to (or extension of) the common vertical transmission of genetic material via cell division or the mating of sexually compatible species. In most cases, HGT is identified as a strong discrepancy between a gene phylogeny to a reliably supported organismal phylogeny. HGT is a common phenomenon in bacterial evolution (Lawrence and Hendrickson 2003), but examples of HGT are increasingly also being identified in eukaryotes, mostly in protists (Andersson 2005). In plants, the majority of identified cases of HGT concern the movement of DNA between mitochondrial genomes in particular (Richardson and Palmer 2006), whereas HGT in plastids seems to occur much more rarely (Rice and Palmer 2006). The term *lateral gene transfer* (LGT) is often used synonymously for HGT.

that develop flowering meristems are an obvious prerequisite. Besides host–parasite interactions epiphytism, illegitimate pollination, or natural grafting may be envisaged for such plant–plant interactions. Once such heterologous cell-to-cell contacts have been established, the obvious question certainly remains why the mitochondrial genome in particular appears to be much more prone to HGT than the chloroplast or nuclear genomes. Most likely the processes of highly active DNA recombination (see Chap. 3) and the readiness for fusion and fission of plant mitochondria, which may rather be seen as a "discontinuous whole" than as continuously separate organelles (Logan 2010; and see Chap. 2), very likely play a major role. A strong gain of DNA recombinational activity during early tracheophyte evolution as reflected with the recently sequenced *Isoetes* mtDNA may at the same time explain restructured genomes as well as the readiness to integrate promiscuous DNA originating from the other two genomic compartments or mtDNA from other plant taxa. If true, one may predict that bryophytes are not prone for acceptance of mtDNA via HGT but only act as donors as was observed for the *Amborella* example (Bergthorsson et al. 2004). Indeed, phylogenetic analyses of several mitochondrial loci at dense sampling of mosses have not given a hint for HGT so far (Beckert et al. 1999; Pruchner et al. 2002; Wahrmund et al. 2009; Volkmar and Knoop 2010; Wahrmund et al. 2010).

1.10 An Extended Perspective: What Else?

What other idiosyncrasies may be identified further down the road in plant mitochondrial genome research? The plethora of peculiarities in plant mitochondrial genomes, including DNA recombination leading to multipartite genome structures, intron gains and losses mainly in early land plant evolution, disrupted *trans*-splicing group II and group I introns, frequent RNA editing, acceptance of promiscuous DNA from the chloroplast and nuclear genomes, and an ongoing endosymbiotic and horizontal DNA transfer is rivaled only by the idiosyncrasies observed in certain protist mtDNAs (Gray et al. 2004). Among the most peculiar mitochondrial genome organizations is the one of the diplonemid protist *Diplonema papillatum,* which consists of more than 100 circular minichromosomes carrying gene modules which give rise to transcripts that have to be matured by an as yet uncharacterized mode of ligation or splicing (Marande and Burger 2007). Similarly idiosyncratic is the organization of the mtDNA in the ichthyosporean *Amoebidium parasiticum,* which features a large population of coexisting linear instead of circular minichromosomes with peculiar terminal sequence motifs (Burger et al. 2003). Extreme degrees of mtDNA recombination as recently observed for the lycophyte *Isoetes* may suggest that yet more complexity of mitochondrial genomes could be expected in other isolated plant lineages (Grewe et al. 2009). The compact and circular mitochondrial genome of the jakobiid protist *Reclinomonas americana* (Lang et al. 1997) reminds of animal mtDNAs in structure but is yet more gene-rich than the ones of plants. Surprisingly though, *rpl*10, a gene that was supposed to be absent