

Organelle Genomes and Endosymbionts

2.1 Mitochondria – 22

- 2.1.1 Origin and Evolution of Mitochondria – 22
- 2.1.2 Animal Mitochondrial Genomes – 25
- 2.1.3 Mitochondrial Genomes of Plants and Algae – 26
- 2.1.4 Mitochondrial Genomes of «Other» Eukaryotes – 28

2.2 Plastids – 29

- 2.2.1 Origin and Evolution of Plastids – 29
- 2.2.2 Plastid Genomes – 31
- 2.2.3 Plastids in the Amoeba *Paulinella chromatophora* – 32

2.3 Heritable Bacterial Endosymbionts – 33

- 2.3.1 Primary Endosymbionts – 33
- 2.3.2 Secondary Endosymbionts – 35

2.4 DNA Barcoding – 35

References – 37

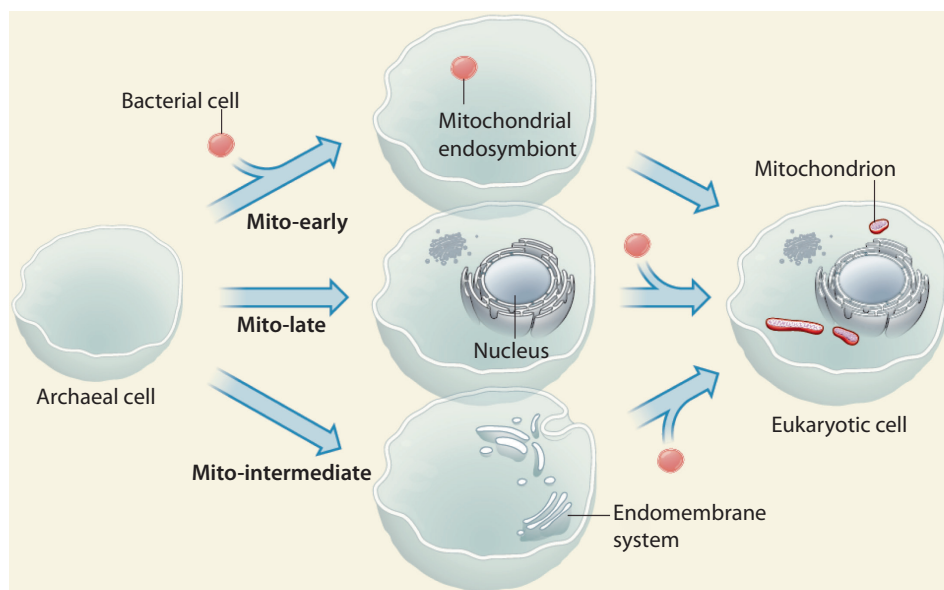
- Mitochondria and chloroplasts are cell organelles of endosymbiotic origin which carry their own genomes.
- Chloroplasts were acquired by their hosts either by primary endosymbiosis (uptake of a cyanobacterium by the last common ancestor of Archaeplastida) or secondary endosymbiotic events (symbiosis of a eukaryote host with chloroplast-bearing algae).
- Organelle genomes differ in size, structure and gene content across taxa.
- Mitochondrial and chloroplast markers are used as standard for DNA barcoding of animals and plants.
- Heritable bacterial endosymbionts are broadly classified as either primary symbionts or secondary symbiont and are commonly found in eukaryotes, especially insects.
- Endosymbiont genomes are highly streamlined and included the smallest reported genomes of all living organisms.

2.1 Mitochondria

2.1.1 Origin and Evolution of Mitochondria

Mitochondria are double-membrane-bound cell organelles which contain their own genome and carry out the replication, transcription and translation of DNA. With the publication of the outstanding book *Origin of Eukaryotic Cells* by Lynn Margulis (1970), the old idea that mitochondria evolved from free-living bacteria via symbiosis got broad attention in the scientific community. The development of cloning and sequencing techniques in the 1970s allowed sequencing of mitochondrial genes and basically confirmed this hypothesis (Gray 2012). Using phylogenomic analyses of different sets of orthologous genes, mitochondria are now firmly placed within Alphaproteobacteria as part of the Rickettsiales (Wang and Wu 2015). Interestingly, this taxon comprises a large variety of bacterial endosymbionts which, similar to mitochondria, also harbour strongly reduced genomes (e.g. the genera *Rickettsia* and *Wolbachia*; ► see Sect. 2.3). Whereas the phylogenetic placement of mitochondria seems well settled, the circumstances under which this symbiosis evolved remain under debate (Gray and Archibald 2012). Several eukaryote taxa lack mitochondria, including Microsporidia (fungi), Trichomonadida (Excavata, Fornicata), Diplomonadida (Excavata, Fornicata) and Archamoebae (Amoebozoa) (Cavalier-Smith 1987; Keeling 1998). It has been suggested that these taxa primarily lack mitochondria uniting them as an early branching clade of eukaryotes called Archezoa (Cavalier-Smith 1983). This would mean that the acquisition of mitochondria took place during eukaryote evolution. Interestingly, molecular phylogenetic analyses including the first available ribosomal sequence data initially supported the early branching of amitochondriate taxa, thereby supporting the Archezoa hypothesis (Sogin 1989). However, this idea was later rejected based on several findings. First, all recent amitochondriate taxa have been demonstrated to bear double-membrane-bounded organelles which are referred to as mitosomes or hydrogenosomes (Hjort et al. 2010) and interpreted to be derived from mitochondria. Hydrogenosomes synthesize ATP under anaerobic condition and thereby produce hydrogen, whereas in the case of mitosomes, the function remains unclear. Hydrogenosomes are also found in taxa which never have

been speculated to be primarily amitochondriate, e.g. in some Loricifera (Metazoa, Ecdysozoa) (Danovaro et al. 2010). It is obvious that these types of reduced organelles evolved several times convergently across eukaryotes. With the oxymonad *Monocercomonoides* sp., only a single case of a eukaryotic taxon lacking any form of mitochondrion has been described (Karnkowska et al. 2016). However, also in this case, a secondary loss of mitochondria is clearly supported by a phylogenetic analysis. Second, phylogenetic analyses using other or more genes clearly proved that the basal branching placement of amitochondriate taxa might be due to systematic errors. Instead, these analyses firmly placed former archezoan taxa as derived eukaryotes. *Microsporidia* are supported as part of the fungi and archaeamobans group deeply within Amoebozoa, and Diplomonadida and Triplomonadida are part of the Excavata (Embley and Martin 2006; Katz and Grant 2014). Finally, PCR-based and genomic analyses found genes which have been transferred from the mitochondrion to the nucleus in amitochondriate taxa (Embley and Martin 2006). Whereas the Archezoa hypothesis has been firmly put to rest and no recent primarily amitochondriate eukaryotes are known, it is still discussed in which order the events leading to the eukaryotic cell have evolved (■ Fig. 2.1). In a mitochondrial-early scenario, it is assumed that the acquisition of mitochondria basically defines the evolution of eukaryotes, suggesting that the last eukaryote common ancestor already had a mitochondrion. In contrast, mitochondrial-late scenarios assume that early eukaryotes already show some cellular complexity, thereby distinguishing them from archaeans, and that mitochondria were acquired via endosymbiosis later in evolution (Ettema 2016). Genes of eukaryotes are of different ancestry, and the origin of many



■ **Fig. 2.1** Origin of eukaryotic cells and their mitochondria. Three different scenarios for the acquisition of mitochondria exist. In the mito-early scenario, the acquisition of the mitochondrion via endosymbiosis already took place in the last eukaryote common ancestor. Mito-late and mito-intermediate scenarios assume already complexly organized eukaryotic ancestors which later in evolution acquired the mitochondrion via endosymbiosis (Reprinted by permission from Macmillan Publishers Ltd: Nature (Ettema 2016), Copyright 2016)

can be traced back to archaean or bacterial clades. As mitochondria are likely of alphaproteobacterial origin, genes related to them, either located in the organelle or transferred to the nucleus, can be traced back to these bacteria in a phylogenetic analysis (McInerney et al. 2014). By using comparative studies of the mitochondrial proteome, a conserved core of proteins descended from the ancestral mitochondrion has been identified (Gray 2015). Similarly, core eukaryote nuclear genes of different functional classes can be identified, whose origin also can be traced back in a phylogenomic analysis, often favouring an archaeal origin. With the help of phylogenetic gene family analyses, the relative age of a given group of genes can be estimated. In case of the mitochondria-early hypothesis, it has to be assumed that there are no differences in the age of genes of archaeal and alphaproteobacterial origin. However, a study testing this hypothesis clearly found support that mitochondria-related genes of an alphaproteobacterial origin are significantly younger than eukaryotic genes of other origin (Pittis and Gabaldón 2016). This would support a mitochondria-late scenario, where an already cellularly complex organized eukaryotic host would have acquired the mitochondrion. Whereas most of the genes without mitochondrial origin can be traced back to an archaeal origin, several other genes are of bacterial origin from different clades, underlining the chimeric nature of early eukaryotes. It remains difficult to distinguish if the acquisition of these diverse sets of genes stems from several events of horizontal gene transfer or maybe previous endosymbiotic associations with other bacteria (Pittis and Gabaldón 2016).

Several adaptive hypotheses exist to explain the ancestral function of mitochondria (Lynch 2007). The primary function of recent mitochondrial organelles is the acquisition of energy, and this might also reflect their ancestral role. Other hypotheses describe ancestral functions like oxygen scavenging, photosynthate acquisition or hydrogen acquisition. Under the latter hypothesis, mitochondria are postulated to originate in a hydrogen-dependent autotrophic archaeal host that lived in a fully anaerobic environment. In this relationship the ancestral role of mitochondria was to provide the host with hydrogen produced by fermentation of organic substrates (Martin and Muller 1998).

Given that mitochondria are coevolving with their hosts for more than 1.5 billion years, it comes without surprise that recent mitochondria differ strikingly across taxa. Notably, mitochondrial genomes of extant eukaryotes differ strongly in size, structure and gene content (Burger et al. 2003b; Nosek and Tomáška 2003). Usually the mitochondrial genome is organized as a single circular molecule, as typical for most prokaryote genomes. However, many deviations from this circular organization have been described. Linearly organized mitochondrial genomes are not rare, as, for example, found in ciliates or medusozoan cnidarians (Kayal et al. 2012; Burger et al. 2000). Several different solutions of maintaining the telomeres of linearly organized mitochondrial DNA have been reported, including hairpin structures, inverted and non-inverted repeat sequences or terminal proteins (Nosek and Tomáška 2003). Mitochondrial genomes are not always encoded on a single molecule, but can also be organized on two or more circular (e.g. in several insects and nematode species or in the flowering plant *Amborella trichopoda*) or linear molecules (e.g. the cnidarian *Hydra magnipapillata* with two linear fragments or the opisthokont *Amoebidium parasiticum* bearing many linear fragments) (Burger et al. 2003a; Gibson et al. 2007; Cameron et al. 2011; Voigt et al. 2008; Rice et al. 2013). Mitochondrial genomes show major differences in the size, with the smallest genomes in a range of around 6–7 Kb (e.g. in several apicomplexans) to the biggest known genomes in the size of several Mb

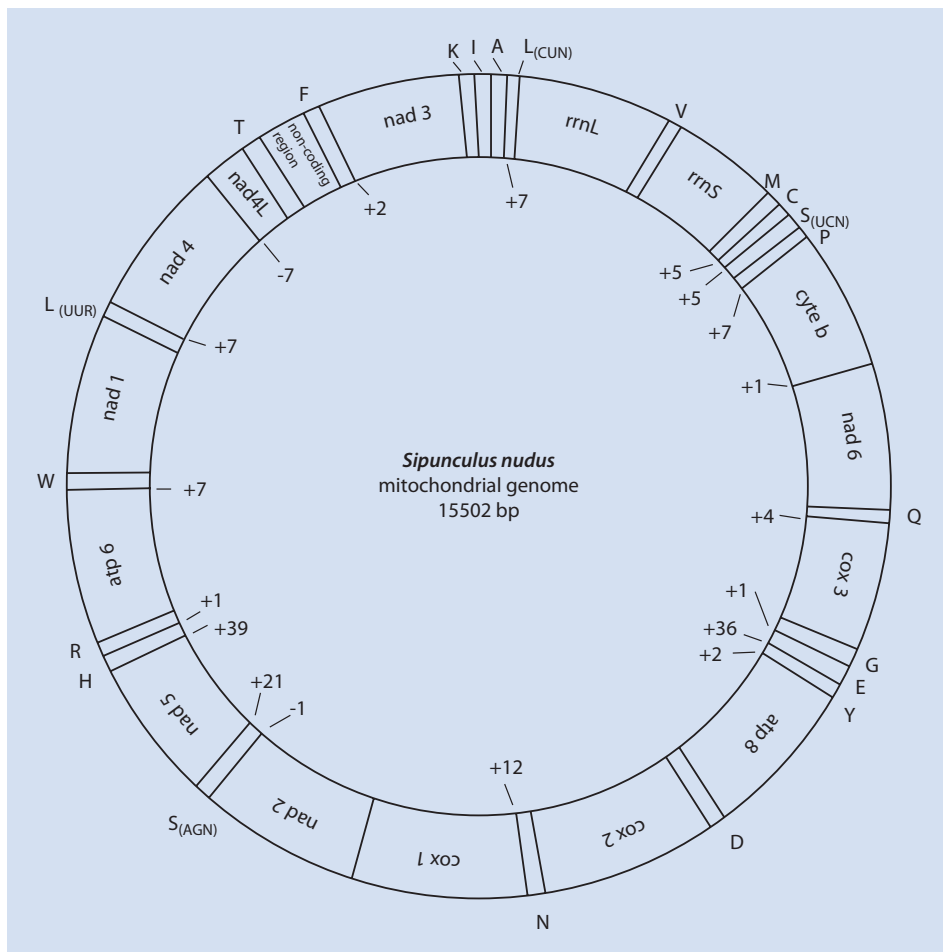
(e.g. 11.3 Mb in the flowering plant *Silene conica*), thereby exceeding the size of some nuclear eukaryote genomes (Hikosaka et al. 2010; Sloan et al. 2012).

The gene content of mitochondrial genomes is greatly reduced compared with Alphaproteobacteria. Large-scale phylogenetic comparisons differ in the number of proteins from 394 to 842 that were (minimally) likely part of the ancestral mitochondrial proteome (Gabaldón and Huynen 2007; Wang and Wu 2014). In contrast, proteins encoded in recent mitochondria range from 3 (as in the apicomplexan *Plasmodium*) to 66 (jakobid Excavata) (Gray 2015). Proteins encoded on mitochondrial genomes are usually involved in the respiratory chain or its corresponding translation system (Lithgow and Schneider 2010). Mitochondrial translation alone requires more than 100 proteins and many other essential housekeeping genes, most of which are encoded in the nucleus and are imported by mitochondria (Dolezal et al. 2006). Altogether, for some species, it is reported that up to 1000 genes are encoded in the nucleus, synthesized in the cytosol and imported to the mitochondria (Lithgow and Schneider 2010). The horizontal transfer of genes from the mitochondrion to the nucleus is further complicated by the presence of differences in the genetic code. The genetic code of mitochondria varies among organisms, and at least 16 deviations from the standard code are reported across eukaryotes, with animals showing the highest diversity (Knight et al. 2001).

2.1.2 Animal Mitochondrial Genomes

Animal mitochondrial genomes usually range in a size between 11 and 20 Kb (Gissi et al. 2008), even though some examples exist with genome sizes of up to 43 Kb as in Placozoa (Dellaporta et al. 2006). Typically, these densely packed genomes encode for 13 protein-coding genes, 22 tRNAs and 2 ribosomal RNAs (■ Fig. 2.2), which all can be located on either strand (Bernt et al. 2013). One protein-coding gene (*atp8*) has been lost convergently in (among others) many Platyhelminthes, Acoelomorpha and Nematoda. It is not unusual that genes begin with alternative starting codons and stop codons are often incomplete. Moreover, tRNAs are often truncated and undergo RNA editing (Börner et al. 1997) or are missing completely (Gissi et al. 2008). Only few examples of the existence of introns are reported, which are in most cases self-splicing group II introns (Huchon et al. 2015). Animal mitochondria are transmitted maternally, even though some examples of doubly uniparental inheritance (DUI) exist. Transmission via DUI is known for several bivalve molluscs and is characterized by the presence of two distinct gender-associated mitochondrial DNAs, where one is transmitted via eggs (F, female) and the other one transmitted through sperm (M, male) (Passamonti et al. 2011). In this case, females are homoplasmic as they only receive mitochondria from their mother, whereas males are heteroplasmic receiving the organelles from both parents. F and M mitochondrial genomes can differ up to 50% in their nucleotide sequence (Doucet-Beaupré et al. 2010).

Substitution rates of mitochondrial genes are several times faster than those of single-copy nuclear genes in most bilaterian animals (Brown et al. 1979). This made mitochondrial genes ideal markers for population genetic, phylogeographic, phylogenetic and barcoding studies (Avise 2004). However, in non-bilaterian animals like Cnidaria and Porifera, mitochondrial substitutions rates are much lower, making these genes less suitable for barcoding or population genetics (Shearer et al. 2002; Huang et al. 2008).



■ **Fig. 2.2** The circular mitochondrial genome of the annelid *Sipunculus nudus* is typical for animal genomes. The densely packed genome encodes for 13 protein-coding genes (*atp1-atp8*, *cox1-cox3*, *cytb*, *nad1-nad6*), 2 ribosomal RNAs (small (*rrnS*) and large (*rrnL*) subunit) and 22 tRNAs (specified in one-letter code) (Reprinted from Mwinyi et al. (2009))

2.1.3 Mitochondrial Genomes of Plants and Algae

Whereas animal mitochondrial genomes are rather uniformly organized, plant mitochondrial genomes exhibit a great diversity in size, structure and gene content (Mower et al. 2012; Liu et al. 2012). Mitochondrial genomes have been sequenced for all major clades of plants (rhodophytes, chlorophytes, charophytes, hornworts, liverworts, mosses, ferns, lycophytes, gymnosperms, angiosperms), showing a variety in genome size of a thousand-fold ranging from 13 Kb in chlorophytes up to 11.3 Mb in angiosperms (Mower et al. 2012; Sloan et al. 2012). Remarkable are the mitochondrial genomes of land plants (embryophytes) which are prone to recombination, RNA editing, trans-splicing, insertion of DNA from the chloroplast and nuclear genomes as well as from distant taxa and ongoing gene transfer into the nucleus (Knoop 2012). Land plant mitochondrial genomes are

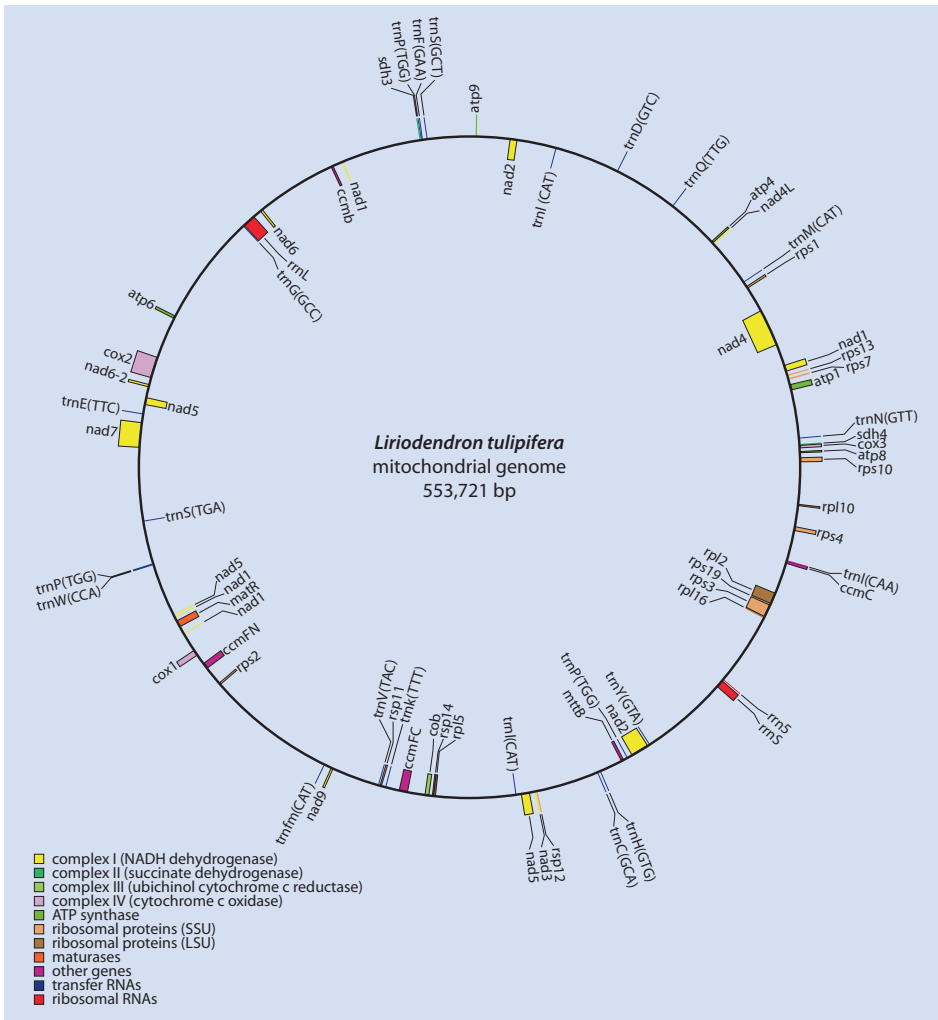


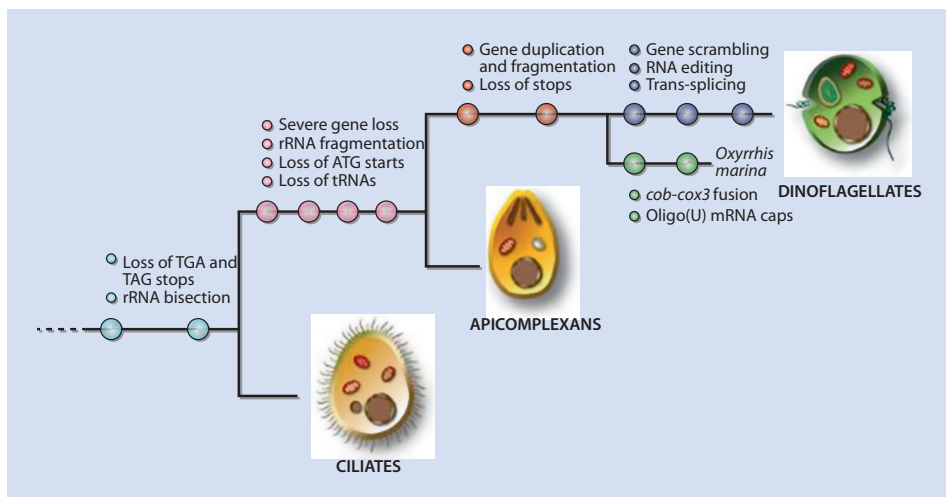
Fig. 2.3 The circular mitochondrial genome of the angiosperm land plant *Liriodendron tulipifera*. As typical for land plants, the genome is expanded in size, harbouring large intergenic regions. Genes inside and outside the circle are transcribed from different strands (The genome has been published by Richardson et al. (2013) and was redrawn using OGDRAW (Lohse et al. 2013))

highly variable in their size, content and structure. The genome size of most land plants exceeds 200 Kb, even though usually less than 20% are encoding for proteins or RNAs (Fig. 2.3). The remaining genome content is dominated by the presence of group I and II introns and large intergenic regions. The number of introns in land plants ranges from 19 to 37, and their position is relatively conserved within clades (Mower et al. 2012). Intergenic regions include repetitive elements, as well as integrations from the nucleus and chloroplast of their own and foreign genomes (Chaw et al. 2008; Rice et al. 2013). However, the origin of most of the excessive intergenic regions of the hugely expanded land plant genomes remains unclear (Sloan et al. 2012).

Plants have a much larger number of genes encoded in their mitochondrial genomes than animals. The number of identified genes in most plants ranges from 42 to 69; however, some chlorophyte green algae only have around 10 genes (Mower et al. 2012; Fan and Lee 2002). As such genes for ribosomal RNAs, tRNAs, ribosomal proteins, a twin-arginine translocase subunit (*tatC*) and *nad* (NADH dehydrogenase), *sdh* (succinate dehydrogenase), *cob* (cytochrome b), *cox* (cytochrome c oxidase), *ccm* (cytochrome c maturation) and *atp* (ATP synthase) subunits are found (Knoop 2012). Interestingly, even though plant mitochondrial genomes show high structural variability, the substitution rates of their encoded genes are rather low in most species (Christensen 2013), making them less suitable as molecular markers in population level or barcoding studies (► see Sect. 2.4).

2.1.4 Mitochondrial Genomes of «Other» Eukaryotes

Of outstanding interest from an evolutionary point of view are the mitochondrial genomes of jakobid flagellates, a group of unicellular eukaryotes which are part of the Excavata (Katz and Grant 2014). Jakobid mitochondrial genomes range in their size from 65 to 100 Kb, while showing also a compact organization with a high coding density ranging from 80% to 93% (Burger et al. 2013). These mitochondrial genomes are the most gene rich among eukaryotes, with nearly 100 genes. Unique among eukaryotes is the presence of genes involved in transcription and quality control of translation. Moreover, some highly conserved gene clusters are found which are interpreted as remnants of an operon structure inherited from the bacterial ancestor of mitochondria. In summary, the genome organization of jakobid mitochondria is suggested to most closely resemble the ancestral pattern of all eukaryotes (Lang et al. 1997). In contrast, the most reduced mitochondrial genomes are found among apicomplexans and dinoflagellates (■ Fig. 2.4), which are sister



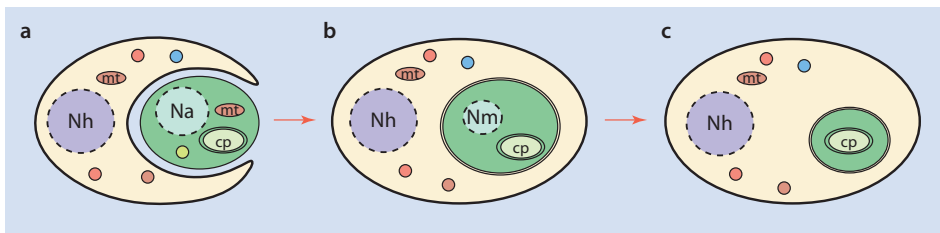
■ **Fig. 2.4** Mitochondrial genome evolution in alveolates. Dinoflagellata and Apicomplexa represent sister taxa, whose current ancestor already possessed a strongly reduced mitochondrial genome, harbouring only three protein-coding genes and fragments of the ribosomal RNAs. Further modifications of the mitochondrial genomes evolved in these lineages (Reprinted by permission from John Wiley and Sons (Waller and Jackson 2009), Copyright 2009)

groups within the taxon Alveolata (SAR clade) (Katz and Grant 2014). The mitochondrial genome of the apicomplexan malaria parasite *Plasmodium falciparum* is encoded on tandemly repeated linear copies of 6 Kb which include only three protein-coding genes (*cox1*, *cox3*, *cob*) and fragments of the small and large subunit of the ribosomal RNA (Feagin et al. 1997). Similarly organized mitochondrial genomes have been revealed for other Apicomplexa, even though the number of ribosomal fragments, the order of genes and the number of copies of tandem repeats (monomeric vs. multiple copies) can vary (Hikosaka et al. 2013). The genome content of dinoflagellates is similar to Apicomplexa (Waller and Jackson 2009). However, some substantial modifications can be found in these genomes, including massive amplification and recombination of the genome. Moreover, trans-splicing is required for generating *cox3* transcripts, and RNA editing of most genes is ubiquitous (Jackson et al. 2012). Dinoflagellates can have surprisingly large genomes given the reduced genome content. The genome of *Symbiodinium minutum* is around 326 Kb, of which 99% are non-coding, even though transcribed (Shoguchi et al. 2015).

2.2 Plastids

2.2.1 Origin and Evolution of Plastids

It is well accepted that plastids originated later than mitochondria in eukaryote evolution, and some eukaryotes bear plastids whereas others not. The first plastids stem from the uptake of a cyanobacterium by the ancestor of Archaeplastida, a clade uniting glaucocystophytes, Rhodophyta and Viridiplantae (green algae and land plants) (Gray 1999). Oxygenic photosynthesis, the conversion of H_2O and CO_2 into energy-rich sugars and O_2 , evolved in the lineage of *Cyanobacteria* more than 3.5 billion years ago, which enduringly transformed life on earth due to oxygen enrichment in the atmosphere (Gould et al. 2008; Hohmann-Marriott and Blankenship 2011). Eukaryotes were able to co-opt photosynthesis by integrating their cyanobacterial endosymbiont. Interestingly, whereas it seems that plastids are all closely related, the organisms that contain them are from diverse eukaryotic clades. Besides Archaeplastida, plastids are further found in euglenids, apicomplexans, haptophytes, cryptomonads, heterokonts and dinoflagellates (Keeling 2010). This can be explained by multiple layers of endosymbiotic events, so-called secondary endosymbiosis (■ Fig. 2.5). In this case a



■ **Fig. 2.5** Evolution of secondary endosymbiosis. **a** A red or green alga with a chloroplast surrounded by a double membrane is engulfed by a eukaryotic host. **b** Eukaryotic host carries a chloroplast with four membranes and the vestigial endosymbiont nucleus (nucleomorph). **c** Eukaryotic host with chloroplast surrounded by four membranes, and the algal nucleus has been completely reduced. *Abbreviations*: cp chloroplast, mt mitochondrium, Na nucleus of the alga, Nh nucleus of the host, Nm nucleomorph

eukaryote carrying a chloroplast has been phagocytized by another eukaryote leading to a subsequent integration of the new endosymbiont. Such events can be distinguished from primary endosymbiosis by the morphology of the plastids, as they still carry additional cell membranes stemming from the phagocytosis. Whereas primary endosymbionts bear two plastid membranes, secondary endosymbionts have four such membranes, which are sometimes reduced to three, as in euglenids and dinoflagellates (Keeling 2013). Moreover, a few cases of tertiary endosymbiosis events have been documented for some dinoflagellate lineages. Species of the genera *Karenia* and *Karlodinium* lost their plastids and gained new ones from a haptophyte species (Tengs et al. 2000). In the case of secondary endosymbiosis, the algal nucleus is usually completely reduced, while only the chloroplast remains (■ Fig. 2.5c). However, in some cases a vestigial nucleus of the algal symbiont is retained (■ Fig. 2.5b). These relicts are called nucleomorphs and can be found in cryptomonads and chlorarachniophytes (Keeling 2010). The actual number of events of secondary and tertiary symbiosis remains still debated, but at least eight distinct evolutionary events are suggested (Cavalier-Smith 2003; Keeling 2013, 2010).

The primary role of the plastid is to conduct photosynthesis, in which case they are called chloroplasts. However, some plastids seem to have lost their photosynthetic ability, e.g. in Apicomplexa (Köhler et al. 1997). As typical for obligate endosymbionts, many unnecessary genes got lost, whereas several other genes were transferred to the host nucleus. Nevertheless, plastids retain a small part of their ancestral genome, which might be due to the fact that hydrophobic proteins are difficult to transport to the organelle or that organelles are needed to be in control of expression for genes which are part of the electron transport chain as a redox regulation (Allen 2015; Timmis et al. 2004). Additionally, some cases of transfer of nuclear genes into the plastid genomes are documented (Keeling 2009). Whereas primary plastids are located in the cytoplasm, secondary plastids are found within the endomembrane system. All genes necessary for plastid function which are encoded in the nucleus have a targeting system to arrive at the plastid and to cross its inner and outer envelopes (Strittmatter et al. 2010). Around 40% of the plastid proteome consists of proteins which seem to be derived from the host nuclear genome or various bacterial lineages outside *Cyanobacteria* (Suzuki and Miyagishima 2010).

Similar to mitochondria, plastid genomes are usually organized as circular molecules, with one genome per circle. Additionally, long, polyploid linear molecules and branched molecules undergoing replication seem to be also abundant (Bendich 2004). Plastid genomes of Archaeplastida are usually around 100–200 Kb in size, and the molecule shows a quadripartite structure due to the presence of two large inverted repeats, which divide the molecule into a large and a small single-copy region (■ Fig. 2.6). Usually around 60–250 genes are encoded on chloroplasts, and they are normally organized as operons. The inverted repeats include the ribosomal RNA genes (16S, 23S and 5S rRNA), as well as some other genes. The number of tRNAs varies between 27 and 31, and a variable number of ribosomal protein genes is usually present. Protein-coding genes are part of photosystems I and II, the cytochrome b6 complex as well as ATP synthase (Green 2011).

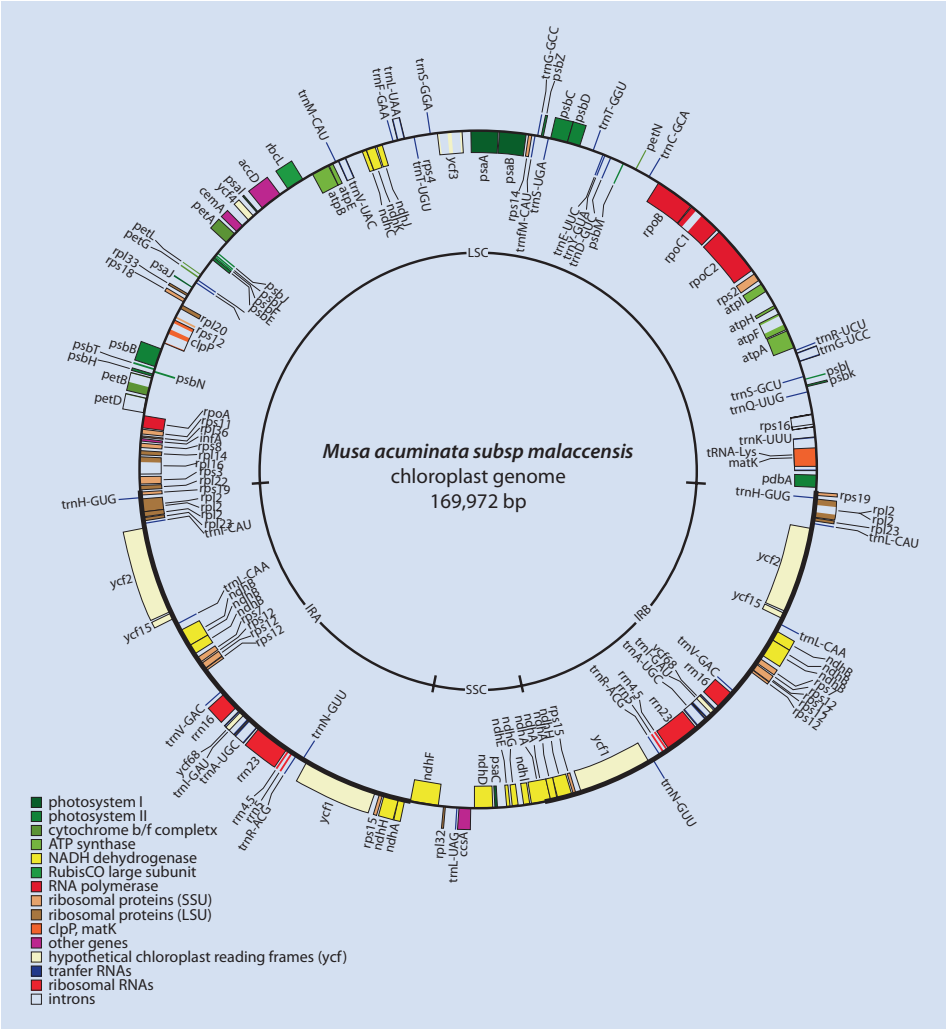


Fig. 2.6 Chloroplast genome of the wild Malaysian banana *Musa acuminata*. Different groups of genes indicated by colours, with genes inside and outside the outer circle transcribed from different strands. Borders of the large single-copy region (LSC) and small (SSC) single-copy region, as well of inverted repeat regions (IRA, IRB) indicated on the inner circle (Chloroplast genome published by Martin et al. (2013) and redrawn with OGDRAW (Lohse et al. 2013))

2.2.2 Plastid Genomes

The composition and size of plastids can vary dramatically across taxa. The largest chloroplast genomes are found in green algae, with the ~500 Kb genome of *Floydiella terrestris* as the actual record holder (Brouard et al. 2010). The large size in these taxa is mainly due

to repetitive regions and not due to the number of retained genes, which is with ~100 similar to that of most land plants. In contrast, the chloroplast genomes of red algae (Rhodophyta) bear with 220–250, the largest number of genes of any sequenced chloroplasts (Janouškovec et al. 2013). It seems that red algae show a much slower rate of plastid-to-nucleus gene transfer. Moreover, red algae plastid genomes also lack inverted repeats. The other extreme in terms of gene number comes from dinoflagellates. The chloroplasts of most photosynthetic dinoflagellates contain the light-harvesting pigment peridinin and are surrounded by three membranes, suggesting a secondary symbiosis, putatively stemming from a red algal host. The organization of these chloroplasts is highly unusual, and only a small number of genes (17) is retained, which are located on different minicircles. Moreover, other chloroplast types stemming from different symbiosis events are also described in this taxon, which accordingly differ in size and organization (Dorrell and Howe 2015). Half of the described dinoflagellate lineages lost their plastids (Green 2011). Apicomplexa, the sister taxon of dinoflagellates, are parasitic eukaryotes (including important pathogens of humans or livestock as *Plasmodium*, *Toxoplasma*, *Eimeria*) which bear a name-giving organelle called apicoplast, which is also derived from a secondary uptake of a chloroplast of putatively red algal origin. However, unlike chloroplasts these organelles do not show photosynthetic activity. Nevertheless, apicoplasts retain their own genome and expression machinery and are involved in the synthesis of fatty acids, isoprenoids, iron sulphur clusters and haem (Lim and McFadden 2010).

Transcription in chloroplasts can be mediated by two different types of RNA polymerase or by a combination of both. According to the location where the corresponding polymerase is encoded, these types are abbreviated as PEP (plastid-encoded plastid RNA polymerase) or NEP (nuclear-encoded plastid RNA polymerase) (Yagi and Shiina 2014). Chloroplast genes are categorized into three classes according to the promoter they bear for their transcription. Class I genes are photosynthesis-related genes and are mainly transcribed by PEP, class II genes comprise mainly of housekeeping genes transcribed by both PEP and NEP, and class III genes (e.g. the gene *accD* and the *rpoB* operon) are transcribed by NEP (Hajdukiewicz et al. 1997). Chloroplasts represent a highly oxidative environment leading to an increased mutation rate. Post-transcriptional repair by RNA editing is widely used to restore affected genes by insertion, deletion or modification of specific nucleotides. Using this mechanism, mainly C to U, but also some U to C, conversions are conducted (Kotera et al. 2005).

Chloroplast genes have been extensively used for phylogenetic and phylogeographic questions. A combination of the genes *rbcL* and *matK* has been proposed as barcode for the identification of plant species (Hollingsworth et al. 2009). Conserved primers for easy amplification of the *rbcL* gene are available, and it is by far the most widely used gene in plant systematics, with over 50,000 published sequences in NCBI GenBank (Li et al. 2015). Several other plastid genome regions or fragments such as *atpF-H*, *matK*, *psbK-I*, *rbcL*, *ropC1*, *rpoB*, *trnH-psbA* and *trnL-F* have been also widely used in plant molecular systematic studies. Especially the advent of next-generation sequencing techniques enabled an increase in sequencing of complete chloroplast genomes for phylogenomic studies of land plants and green algae (Ruhfel et al. 2014; Lemieux et al. 2014).

2.2.3 Plastids in the Amoeba *Paulinella chromatophora*

The amoeba *Paulinella chromatophora* (■ Fig. 2.7) also contains two plastids with photosynthetic activities which have been demonstrated to likely originate from an

■ **Fig. 2.7** The amoeba *Paulinella chromatophora* contains plastids with photosynthetic activity originating from an independent endosymbiotic uptake (Picture provided by Eva C.M. Nowack)

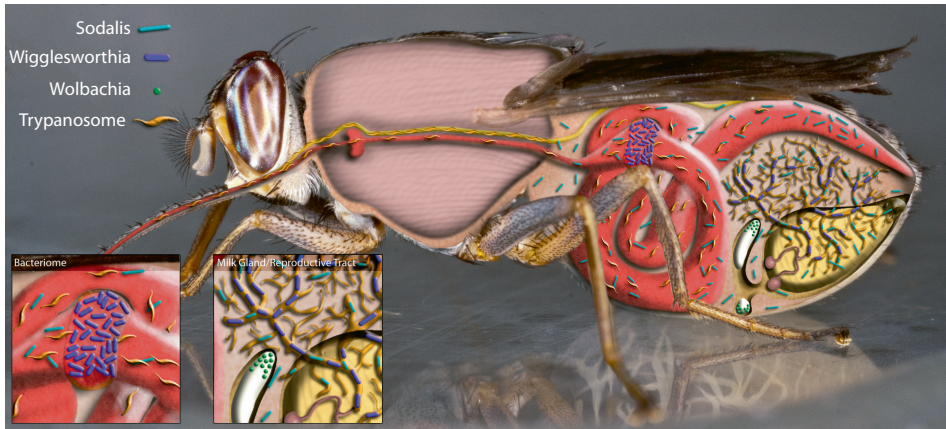


independent endosymbiotic uptake from a cyanobacterium (Marin et al. 2005; Nowack 2014). Interestingly, the endosymbiotic origin of these organelles is with an assumed age of 60–200 mya much younger than the primary endosymbiotic uptake of plastids by Archaeplastida. Morphological studies and sequencing of rRNA genes suggested that these plastids stem from a member of the cyanobacterial *Synechococcus* clade. The plastids bear two envelope membranes surrounding a thick peptidoglycan wall. It could be shown that the endosymbiont lost 75% of its ancestral genome (Bodyl et al. 2012; Nowack et al. 2008). Reduction of the chromatophore genome led to the loss of many important biosynthetic pathways. For compensation, numerous (229) nuclear genes were acquired by horizontal gene transfer of which around 25% came from the endosymbiont (Nowack et al. 2016). The evolution and establishment of such a protein import mechanism qualify the chromatophores of *Paulinella* as cell organelles in a strict sense (Bodyl et al. 2012; Keeling and Archibald 2008; Nowack and Grossman 2012). Conversely, some genes seem also to be imported from the host genome into the plastid genome (Mackiewicz and Bodyl 2010).

2.3 Heritable Bacterial Endosymbionts

2.3.1 Primary Endosymbionts

Mitochondrial and chloroplast organelles evolved from an ancient symbiosis of bacteria with its archaeal or eukaryote host. Heritable bacterial endosymbionts are widespread across eukaryotic taxa, and complex relationships between host and symbionts have been described. Whereas most described endosymbionts belong to Bacteria, some examples from Archaea are also known (van Hoek et al. 2000). The best investigated examples stem from insects, where bacterial endosymbionts are often inherited maternally, as also typical for mitochondria (Ferrari and Vavre 2011). Endosymbionts can be loosely classified into either primary symbionts (P-symbionts) or secondary symbionts (S-symbionts) (Moran



■ **Fig. 2.8** Tsetse flies (*Glossina* spp.) always carry the P-symbiont *Wigglesworthia*, which resides in a specialized organ called bacteriome. Furthermore, often the presence of two S-symbionts can be observed: *Sodalis* and *Wolbachia* (Reprinted by permission from Elsevier Ltd: Trends in Parasitology (Weiss and Aksoy 2011), Copyright 2011)

et al. 2008). P-symbionts are obligate mutualists which are required for the survival or reproduction of the host. Typically, such endosymbionts reside in specialized organs called bacteriomes. Well-studied examples are endosymbionts of insects with a specialized diet. E.g. tsetse flies (*Glossina* sp.) feed exclusively on blood and rely on microbial symbionts to supply amino acids and vitamins the host is not able to synthesize (Aksoy 2000). These flies harbour the gammaproteobacterium *Wigglesworthia glossinidia* as P-symbiont, which resides in specific epithelial cells (bacteriocytes) forming the bacteriome (Balmand et al. 2013) (■ Fig. 2.8). *Wigglesworthia* provides its host with vitamins and supports the digestion of the blood meal. Moreover, female tsetse flies cured from its endosymbiont are infertile (Pais et al. 2008). As typical for P-symbionts, *Wigglesworthia* has a streamlined and highly reduced genome of only 700 Kb (Akman et al. 2002).

The smallest reported genomes are found in P-symbionts of sap-feeding insects, often retaining only a minimal gene set (McCutcheon and Moran 2012). With 139 Kb, the smallest genome is reported for the mealybug P-symbiont *Candidatus Tremblaya princeps* (López-Madriral et al. 2011). The genome of this betaproteobacterium contains only 120 protein-coding genes and misses several essential genes. Interestingly, in the cytoplasm of *Tremblaya* is another endosymbiont resident, the gammaproteobacterium *Candidatus Moranella endobia*, which supports essential functions of its bacterial host (von Dohlen et al. 2001; Husnik et al. 2013). This highly degenerated endosymbiont genomes led to a complete dependency of their hosts, often blurring the distinction between organelles and endosymbionts (McCutcheon and Keeling 2014). For example, a protein of a gene, which has been horizontally transferred to its host nucleus, has been demonstrated to be transported back to its obligate endosymbiont in aphids (Nakabachi et al. 2014). The evolution of protein targeting systems to redirect the products of horizontally transferred genes back to the symbiont is regarded as one of the major transitions in organelle evolution (Cavalier-Smith and Lee 1985). Aphids usually possess intracellular gammaproteobacteria of the genus *Buchnera*, which are transmitted vertically (to the offspring) via the ovary. This symbiotic relationship is obligate for both partners, as aphids without symbionts have a low fitness or are infertile, and *Buchnera* are unknown outside their aphid hosts (Douglas

1998). The symbiosis between aphids and *Buchnera* is very old, suggested to be established ~200 mya. Consequently, co-diversification between aphid hosts and their *Buchnera* symbionts can be found (Baumann 2005), and phylogenetic analyses of symbiont genes were even helpful to resolve aphid relationships (Novakova et al. 2013).

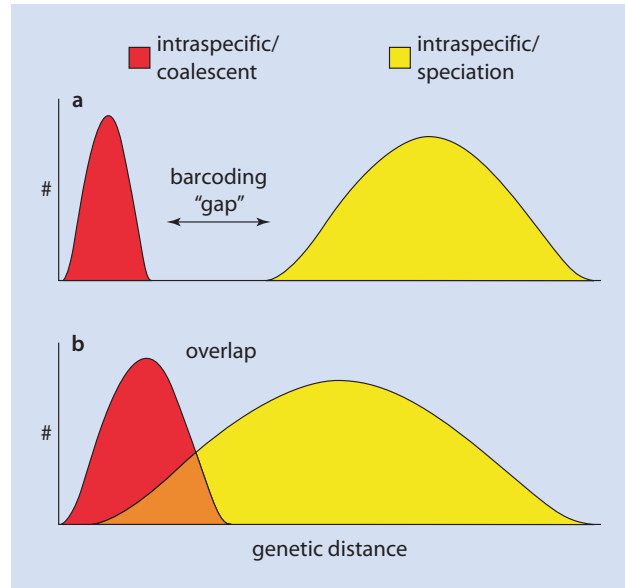
2.3.2 Secondary Endosymbionts

S-symbionts are bacterial symbionts which can be facultative mutualists and reproductive manipulators or have completely unknown effects on their host (Moran et al. 2008). Usually these symbionts reside in different cell types of their hosts and often invade reproductive organs, but can also be found in fluids of the body cavity. Unlike P-symbionts, there is usually no co-diversification found between S-symbionts and their hosts. Prevalence of S-symbionts in host populations can range from infecting some few up to all individuals, e.g. tsetse flies often carry two S-symbionts: the facultative mutualist *Sodalis glossinidius* (Enterobacteriaceae) and the alphaproteobacterium *Wolbachia pipientis* (■ Fig. 2.8). Whereas the P-symbiont is found in all tsetse flies, the infection prevalence of the S-symbionts can strongly vary across species and populations ranging from 1.4% up to 93.7% in the case of *Sodalis* (Dennis et al. 2014). *Wolbachia* is found in arthropods and filarial nematodes, and it is estimated that infections occur in ~40% of all terrestrial arthropod species (Zug and Hammerstein 2012). This ubiquity explains why *Wolbachia* is one of the best investigated endosymbionts and relationships with their hosts are investigated in many cases. Within *Wolbachia*, many different supergroups are described, which differ in their host range and their symbiotic relationships (Gerth et al. 2014). By far the most of the known *Wolbachia* strains belong to either supergroups A or B, mostly infecting insects, but also other arthropod species. Interestingly, it has been shown that the origin of these supergroups coincides with the diversification of hyperdiverse insect lineages ~200 mya (Gerth and Bleidorn 2016). *Wolbachia* are well known as reproductive manipulators of their hosts. As vertical transmission is exclusively maternal, several mechanisms to enhance the spread across the host population are described. These include distortion of the host population sex ratio via parthenogenesis, male killing or feminization, as well as induction of cytoplasmic incompatibility (CI). In the case of CI, eggs of uninfected females are incompatible with the sperm of infected males, thereby preventing successful mating (Werren et al. 2008). As typical for S-symbionts, there is usually no co-diversification pattern between *Wolbachia* and their arthropod hosts, suggesting repeated instances of horizontal transmission between unrelated species (Werren et al. 1995).

2.4 DNA Barcoding

DNA-based species identification by a universal DNA barcode of few standard DNA regions became firstly established for animals (Hebert et al. 2003a) and later also standard in plants, fungi and other eukaryotes (Hollingsworth et al. 2009; Schoch et al. 2012; Saunders and McDevit 2012). The idea to use molecular markers for species identification and delimitation was already in use for decades in prokaryotes (Tindall et al. 2010), most commonly utilizing the 16S rRNA gene. For eukaryotes, DNA barcoding has been most successfully developed for animals, where a 658 bp region of the mitochondrial cytochrome oxidase 1 gene (*cox1*) is used as standard marker. The choice of this mitochondrial

Fig. 2.9 Schematic representation of the DNA barcoding gap. **a** In the ideal case, there is no overlap between intraspecific and interspecific genetic variability, thereby creating a barcoding gap. **b** In many «real-world» examples, an area of overlap of the genetic variability between interspecific and intraspecific comparisons exists (Reprinted from Meyer and Paulay (2005))



marker has several advantages: (I) nearly universal primer pairs for the *cox1* fragment are available (Folmer et al. 1994); (II) mitochondrial DNA is available in a much higher copy number per cell than nuclear DNA, thereby alleviating DNA extraction and amplification; and (III) the existence of a «barcoding gap» is proposed (Fig. 2.9), where interspecific genetic variation clearly exceeds intraspecific variation (Hebert et al. 2003b). Mitochondrial DNA of plants evolves much slower than its animal counterpart, and consequently with the genes *rbcl* and *matk*, two chloroplast markers are currently in use for DNA barcoding of plants (Hollingsworth et al. 2009). In fungi, barcoding relies on the nuclear ITS regions, so-called internal transcribed spacers separating the tandemly repeated ribosomal RNA genes (Schoch et al. 2012). As reference for this approach serves the Barcode of Life Data Systems (BOLD) (Ratnasingham and Hebert 2007), which links specimen information, metadata and genetic sequence data. The Consortium for the Barcode of Life (CBOL) coordinates and promotes the standardization of DNA barcoding. Many country- and taxon-specific initiatives contribute to the growth of the database. In February 2016, around 2.5 million *cox1* sequences from more than 175,000 animal species were accessible in BOLD. DNA barcoding offers several practical applications including protection of endangered species, product authentication, control of invasive and pest species, biodiversity monitoring, diet analyses, linking larval of developmental with adult stages and the discovery of new species.

The reliance on one or few markers also promoted several critiques of the barcoding approach. Especially, focussing solely on organelle markers may be misleading due to reduced effective population size, introgression, maternal inheritance, inconsistent mutation rate, pseudogenization or heteroplasmy (Galtier et al. 2009). The presence of endosymbionts manipulating the host reproduction and thereby altering inheritance patterns of maternally transmitted genes may imply further complications (Gerth et al. 2011). Moreover, the existence of a «barcoding gap» might be an artefact generated through an insufficient sampling across taxa and populations (Wiemers and Fiedler 2007). Nevertheless, DNA barcoding became popular, and especially the advent of

next-generation sequencing techniques allowed metabarcoding studies estimating the diversity of communities previously difficult to handle as, e.g. from soil, permafrost or the deep sea (Valentini et al. 2009). Metabarcoding describes the simultaneous amplification of DNA barcodes from mass collections of organisms or environmental DNA (Yu et al. 2012). Such studies usually discover a huge amount of DNA sequences which do not match with any entry for BOLD, and species-delimitation methods are needed for classification. The most popular methods are based on the generalized mixed Yule coalescent (GMYC) model (Pons et al. 2006) or Poisson tree processes (PTP) (Zhang et al. 2013).

References

- Akman L, Yamashita A, Watanabe H, Oshima K, Shiba T, Hattori M, Aksoy S (2002) Genome sequence of the endocellular obligate symbiont of tsetse flies, *Wigglesworthia glossinidia*. *Nat Genet* 32:402–407
- Aksoy S (2000) Tsetse – a haven for microorganisms. *Parasitol Today* 16:114–118
- Allen J (2015) Why chloroplasts and mitochondria retain their own genomes and genetic systems: colocalization for redox regulation of gene expression. *Proc Natl Acad Sci U S A* 112:10231–10238
- Awise JC (2004) Molecular markers, natural history, and evolution. Sinauer Associates, Inc, Sunderland
- Balmand S, Lohs C, Aksoy S, Heddi A (2013) Tissue distribution and transmission routes for the tsetse fly endosymbionts. *J Invertebr Pathol* 112(Suppl 1):S116–S122
- Baumann P (2005) Biology of bacteriocyte-associated endosymbionts of plant sap-sucking insects. *Annu Rev Microbiol* 59:155–189
- Bendich AJ (2004) Circular chloroplast chromosomes: the grand illusion. *Plant Cell* 16:1661–1666
- Bernt M, Braband A, Schierwater B, Stadler PF (2013) Genetic aspects of mitochondrial genome evolution. *Mol Phylogenet Evol* 69:328–338
- Bodyl A, Mackiewicz P, Gagat P (2012) Organelle evolution: *Paulinella* breaks a paradigm. *Curr Biol* 22:R304–R306
- Börner GV, Yokobori S-I, Mörl M, Dörner M, Pääbo S (1997) RNA editing in metazoan mitochondria: staying fit without sex. *FEBS Lett* 409:320–324
- Brouard J-S, Otis C, Lemieux C, Turmel M (2010) The exceptionally large chloroplast genome of the green alga *Floydiella terrestris* illuminates the evolutionary history of the Chlorophyceae. *Genome Biol Evol* 2:240–256
- Brown WM, George M, Wilson AC (1979) Rapid evolution of animal mitochondrial DNA. *Proc Natl Acad Sci U S A* 76:1967–1971
- Burger G, Forget L, Zhu Y, Gray MW, Lang BF (2003a) Unique mitochondrial genome architecture in unicellular relatives of animals. *Proc Natl Acad Sci U S A* 100:892–897
- Burger G, Gray MW, Forget L, Lang BF (2013) Strikingly bacteria-like and gene-rich mitochondrial genomes throughout jakobid protists. *Genome Biol Evol* 5:418–438
- Burger G, Gray MW, Franz Lang B (2003b) Mitochondrial genomes: anything goes. *Trends Genet* 19:709–716
- Burger G, Zhu Y, Littlejohn TG, Greenwood SJ, Schnare MN, Lang BF, Gray MW (2000) Complete sequence of the mitochondrial genome of *Tetrahymena pyriformis* and comparison with *Paramecium aurelia* mitochondrial DNA. *J Mol Biol* 297:365–380
- Cameron SL, Yoshizawa K, Mizukoshi A, Whiting MF, Johnson KP (2011) Mitochondrial genome deletions and minicircles are common in lice (Insecta: Phthiraptera). *BMC Genomics* 12:394
- Cavalier-Smith T (1983) A 6-kingdom classification and a unified phylogeny. In: Schenk HEA, Schwemmler WS (eds) *Endocytobiology. II. Intracellular Space as Oligogenetic Ecosystem*. Walter de Gruyter, Berlin, pp 1027–1034
- Cavalier-Smith T (1987) Eukaryotes with no mitochondria. *Nature* 326:332–333
- Cavalier-Smith T (2003) Genomic reduction and evolution of novel genetic membranes and protein-targeting machinery in eukaryote-eukaryote chimaeras (meta-algae). *Philos Trans R Soc Lond Ser B Biol Sci* 358:109–134
- Cavalier-Smith T, Lee JJ (1985) Protozoa as hosts for endosymbioses and the conversion of symbionts into organelles. *J Protozool* 32:376–379

- Chaw S-M, Chun-Chieh Shih A, Wang D, Wu Y-W, Liu S-M, Chou T-Y (2008) The mitochondrial genome of the gymnosperm *Cycas taitungensis* contains a novel family of short interspersed elements, Bpu sequences, and abundant RNA editing sites. *Mol Biol Evol* 25:603–615
- Christensen AC (2013) Plant mitochondrial genome evolution can be explained by DNA repair mechanisms. *Genome Biol Evol* 5:1079–1086
- Danovaro R, Dell'Anno A, Pusceddu A, Gambi C, Heiner I, Møbjerg Kristensen R (2010) The first metazoa living in permanently anoxic conditions. *BMC Biol* 8:1–10
- Dellaporta SL, Xu A, Sagasser S, Jakob W, Moreno MA, Buss LW, Schierwater B (2006) Mitochondrial genome of *Trichoplax adhaerens* supports Placozoa as the basal lower metazoan phylum. *Proc Natl Acad Sci U S A* 103:8751–8756
- Dennis JW, Durkin SM, Horsley Downie JE, Hamill LC, Anderson NE, MacLeod ET (2014) *Sodalis glossinidius* prevalence and trypanosome presence in tsetse from Luambe National Park, Zambia. *Parasit Vectors* 7:1–11
- Dolezal P, Likic V, Tachezy J, Lithgow T (2006) Evolution of the molecular machines for protein import into mitochondria. *Science* 313:314–318
- Dorrell RG, Howe CJ (2015) Integration of plastids with their hosts: lessons learned from dinoflagellates. *Proc Natl Acad Sci USA* 112(33):10247–10254
- Doucet-Beaupré H, Breton S, Chapman EG, Blier PU, Bogan AE, Stewart DT, Hoeh WR (2010) Mitochondrial phylogenomics of the Bivalvia (Mollusca): searching for the origin and mitogenomic correlates of doubly uniparental inheritance of mtDNA. *BMC Evol Biol* 10:50
- Douglas AE (1998) Nutritional interactions in insect-microbial symbioses: aphids and their symbiotic bacteria *Buchnera*. *Annu Rev Entomol* 43:17–37
- Embley TM, Martin W (2006) Eukaryotic evolution, changes and challenges. *Nature* 440:623–630
- Ettema TJG (2016) Evolution: mitochondria in the second act. *Nature* 531:39–40
- Fan J, Lee RW (2002) Mitochondrial genome of the colorless green alga *Polytomella parva*: two linear DNA Molecules with homologous inverted repeat termini. *Mol Biol Evol* 19:999–1007
- Feagin JE, Mericle BL, Werner E, Morris M (1997) Identification of additional rRNA fragments encoded by the *Plasmodium falciparum* 6 kb element. *Nucleic Acids Res* 25:438–446
- Ferrari J, Vavre F (2011) Bacterial symbionts in insects or the story of communities affecting communities. *Philos Trans R Soc Lond Ser B Biol Sci* 366:1389–1400
- Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol Mar Biol Biotechnol* 3:294–299
- Gabaladón T, Huynen MA (2007) From endosymbiont to host-controlled organelle: the hijacking of mitochondrial protein synthesis and metabolism. *PLoS Comput Biol* 3:e219
- Galtier N, Nabholz B, Glemin S, Hurst GDD (2009) Mitochondrial DNA as a marker of molecular diversity: a reappraisal. *Mol Ecol* 18:4541–4550
- Gerth M, Bleidorn C (2016) Comparative genomics provides a timeframe for *Wolbachia* evolution and exposes a recent biotin synthesis operon transfer. *Nat Microbiol* 2:16241
- Gerth M, Gansauge MT, Weigert A, Bleidorn C (2014) Phylogenomic analyses uncover origin and spread of the *Wolbachia* pandemic. *Nat Commun* 5:5117
- Gerth M, Geißler A, Bleidorn C (2011) *Wolbachia* infections in bees (Anthophila) and possible implications for DNA barcoding. *Syst Biodivers* 9:319–327
- Gibson T, Blok VC, Phillips MS, Hong G, Kumarasinghe D, Riley IT, Dowton M (2007) The mitochondrial subgenomes of the nematode *Globodera pallida* are mosaics: evidence of recombination in an animal mitochondrial genome. *J Mol Evol* 64:463–471
- Gissi C, Iannelli F, Pesole G (2008) Evolution of the mitochondrial genome of Metazoa as exemplified by comparison of congeneric species. *Heredity* 101:301–320
- Gould SB, Waller RF, McFadden GI (2008) Plastid evolution. *Annu Rev Plant Biol* 59:491–517
- Gray MW (1999) Evolution of organellar genomes. *Curr Opin Genet Dev* 9:678–687
- Gray MW (2012) Mitochondrial evolution. *Cold Spring Harb Perspect Biol* 4:a011403
- Gray MW (2015) Mosaic nature of the mitochondrial proteome: implications for the origin and evolution of mitochondria. *Proc Natl Acad Sci U S A* 112:10133–10138
- Gray MW, Archibald JM (2012) Origins of mitochondria and plastids. In: Bock R, Knoop V (eds) *Genomics of chloroplasts and mitochondria, Advances in photosynthesis and respiration*, vol 35. Springer Science + Business Media B.V, Dordrecht, pp 1–30
- Green BR (2011) Chloroplast genomes of photosynthetic eukaryotes. *Plant J* 66:34–44

References

- Hajdukiewicz PTJ, Allison LA, Maliga P (1997) The two RNA polymerases encoded by the nuclear and the plastid compartments transcribe distinct groups of genes in tobacco plastids. *EMBO J* 16:4041–4048
- Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003a) Biological identifications through DNA barcodes. *Proc R Soc Lond B Biol Sci* 270:313–321
- Hebert PDN, Ratnasingham S, de Waard JR (2003b) Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proc R Soc Lond B Biol Sci* 270:S96–S99
- Hikosaka K, Kita K, Tanabe K (2013) Diversity of mitochondrial genome structure in the phylum Apicomplexa. *Mol Biochem Parasitol* 188:26–33
- Hikosaka K, Watanabe Y-I, Tsuji N, Kita K, Kishine H, Arisue N, Palacpac NMQ, Kawazu S-I, Sawai H, Horii T, Igarashi I, Tanabe K (2010) Divergence of the mitochondrial genome structure in the apicomplexan parasites, *Babesia* and *Theileria*. *Mol Biol Evol* 27:1107–1116
- Hjort K, Goldberg AV, Tsousis AD, Hirt RP, Embley TM (2010) Diversity and reductive evolution of mitochondria among microbial eukaryotes. *Philos Trans R Soc Lond Ser B Biol Sci* 365:713–727
- Hohmann-Marriott MF, Blankenship RE (2011) Evolution of photosynthesis. *Annu Rev Plant Biol* 62:515–548
- Hollingsworth PM, Forrest LL, Spouge JL, Hajibabaei M, Ratnasingham S, van der Bank M, Chase MW, Cowan RS, Erickson DL, Fazekas AJ, Graham SW, James KE, Kim K-J, Kress WJ, Schneider H, van AlphenStahl J, Barrett SCH, van den Berg C, Bogarin D, Burgess KS, Cameron KM, Carine M, Chacón J, Clark A, Clarkson JJ, Conrad F, Devey DS, Ford CS, Hedderson TAJ, Hollingsworth ML, Husband BC, Kelly LJ, Kesanakurti PR, Kim JS, Kim Y-D, Lahaye R, Lee H-L, Long DG, Madriñán S, Maurin O, Meusnier I, Newmaster SG, Park C-W, Percy DM, Petersen G, Richardson JE, Salazar GA, Savolainen V, Seberg O, Wilkinson MJ, Yi D-K, Little DP (2009) A DNA barcode for land plants. *Proc Natl Acad Sci U S A* 106:12794–12797
- Huang D, Meier R, Todd PA, Chou LM (2008) Slow mitochondrial COI sequence evolution at the base of the metazoan tree and its implications for DNA barcoding. *J Mol Evol* 66:167–174
- Huchon D, Szitenberg A, Shefer S, Ilan M, Feldstein T (2015) Mitochondrial group I and group II introns in the sponge orders Agelasida and Axinellida. *BMC Evol Biol* 15:278
- Husnik F, Nikoh N, Koga R, Ross L, Duncan Rebecca P, Fujie M, Tanaka M, Satoh N, Bachtrog D, Wilson Alex CC, von Dohlen CD, Fukatsu T, McCutcheon John P (2013) Horizontal gene transfer from diverse bacteria to an insect genome enables a tripartite nested mealybug symbiosis. *Cell* 153:1567–1578
- Jackson CJ, Gornik SG, Waller RF (2012) The mitochondrial genome and transcriptome of the basal dinoflagellate *Hematodinium* sp.: character evolution within the highly derived mitochondrial genomes of dinoflagellates. *Genome Biol Evol* 4:59–72
- Janouškovec J, Liu S-L, Martone PT, Carré W, Leblanc C, Collén J, Keeling PJ (2013) Evolution of red algal plastid genomes: ancient architectures, introns, horizontal gene transfer, and taxonomic utility of plastid markers. *PLoS One* 8:e59001
- Karnkowska A, Vacek V, Zubáčová Z, Treitli Sebastian C, Petrželková R, Eme L, Novák L, Žárský V, Barlow Lael D, Herman Emily K, Soukal P, Hroudová M, Doležal P, Stairs Courtney W, Roger Andrew J, Eliáš M, Dacks Joel B, Vlček Č, Hampl V (2016) A eukaryote without a mitochondrial organelle. *Curr Biol* 26:1274–1284
- Katz LA, Grant JR (2014) Taxon-rich phylogenomic analyses resolve the eukaryotic tree of life and reveal the power of subsampling by sites. *Syst Biol* 64:406–415
- Kayal E, Bentlage B, Collins AG, Kayal M, Pirro S, Lavrov DV (2012) Evolution of linear mitochondrial genomes in medusozoan cnidarians. *Genome Biol Evol* 4:1–12
- Keeling PJ (1998) A kingdom's progress: Archezoa and the origin of eukaryotes. *BioEssays* 20:87–95
- Keeling PJ (2009) Role of horizontal gene transfer in the evolution of photosynthetic eukaryotes and their plastids. In: Gogarten MB, Gogarten JP, Olendzenski LC (eds) *Horizontal gene transfer, Methods in molecular biology*, vol 532. Humana Press, New York, pp 501–515
- Keeling PJ (2010) The endosymbiotic origin, diversification and fate of plastids. *Philos Trans R Soc Lond Ser B Biol Sci* 365:729–748
- Keeling PJ (2013) The number, speed, and impact of plastid endosymbioses in eukaryotic evolution. *Annu Rev Plant Biol* 64:583–607
- Keeling PJ, Archibald JM (2008) Organelle evolution: what's in a name? *Curr Biol* 18:R345–R347
- Knight RD, Freeland SJ, Landweber LF (2001) Rewiring the keyboard: evolvability of the genetic code. *Nat Rev Genet* 2:49–58
- Knoop V (2012) Seed plant mitochondrial genomes: complexity evolving. In: Bock R, Knoop V (eds) *Genomics of chloroplasts and mitochondria, Advances in photosynthesis and respiration*, vol 35. Springer Science + Business Media B.V, Dordrecht, pp 175–200

- Köhler S, Delwiche CF, Denny PW, Tilney LG, Webster P, Wilson RJM, Palmer JD, Roos DS (1997) A plastid of probable green algal origin in apicomplexan parasites. *Science* 275:1485–1489
- Kotera E, Tasaka M, Shikanai T (2005) A pentatricopeptide repeat protein is essential for RNA editing in chloroplasts. *Nature* 433:326–330
- Lang BF, Burger G, O’Kelly CJ, Cedergren R, Golding GB, Lemieux C, Sankoff D, Turmel M, Gray MW (1997) An ancestral mitochondrial DNA resembling a eubacterial genome in miniature. *Nature* 387:493–497
- Lemieux C, Otis C, Turmel M (2014) Chloroplast phylogenomic analysis resolves deep-level relationships within the green algal class Trebouxiophyceae. *BMC Evol Biol* 14:211
- Li X, Yang Y, Henry RJ, Rossetto M, Wang Y, Chen S (2015) Plant DNA barcoding: from gene to genome. *Biol Rev* 90:157–166
- Lim L, McFadden GI (2010) The evolution, metabolism and functions of the apicoplast. *Philos Trans R Soc Lond Ser B Biol Sci* 365:749–763
- Lithgow T, Schneider A (2010) Evolution of macromolecular import pathways in mitochondria, hydro-genosomes and mitosomes. *Philos Trans R Soc Lond Ser B Biol Sci* 365:799–817
- Liu Y, Wang B, Cui P, Li L, Xue J-Y, Yu J, Qiu Y-L (2012) The mitochondrial genome of the lycophyte *Huperzia squarrosa*: the most archaic form in vascular plants. *PLoS One* 7:e35168
- Lohse M, Drechsel O, Kahlau S, Bock R (2013) OrganellarGenomeDRAW – a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. *Nucleic Acids Res* 41:W575–W581
- López-Madrigal S, Latorre A, Porcar M, Moya A, Gil R (2011) Complete genome sequence of «*Candidatus Tremblaya princeps*» strain PCVAL, an intriguing translational machine below the living-cell status. *J Bacteriol* 193:5587–5588
- Lynch M (2007) The origins of genome architecture. Sinauer Assoc, Sunderland
- Mackiewicz P, Bodyl A (2010) A hypothesis for import of the nuclear encoded *PsaE* Protein of *Paulinella chromatophora* (Cercozoa, Rhizaria) into its cyanobacterial endosymbionts/plastids via the endo-membrane system. *J Phycol* 46:847–859
- Margulis L (1970) Origin of eukaryotic cells. Yale University Press, New Haven
- Marin B, Nowack EC, Melkonian M (2005) A plastid in the making: evidence for a second primary endosymbiosis. *Protist* 156:425–432
- Martin G, Baurens F-C, Cardi C, Aury J-M, D’Hont A (2013) The complete chloroplast genome of banana (*Musa acuminata*, Zingiberales): insight into plastid monocotyledon evolution. *PLoS One* 8:e67350
- Martin W, Muller M (1998) The hydrogen hypothesis for the first eukaryote. *Nature* 392:37–41
- McCutcheon JP, Keeling PJ (2014) Endosymbiosis: protein targeting further erodes the organelle/symbiont distinction. *Curr Biol* 24:R654–R655
- McCutcheon JP, Moran NA (2012) Extreme genome reduction in symbiotic bacteria. *Nat Rev Microbiol* 10:13–26
- McInerney JO, O’Connell MJ, Pisani D (2014) The hybrid nature of the Eukaryota and a consilient view of life on earth. *Nat Rev Microbiol* 12:449–455
- Meyer CP, Paulay G (2005) DNA barcoding: error rates based on comprehensive sampling. *PLoS Biol* 3:e422
- Moran NA, McCutcheon JP, Nakabachi A (2008) Genomics and evolution of heritable bacterial symbionts. *Annu Rev Genet* 42:165–190
- Mower JP, Sloan DB, Alverson AJ (2012) Plant mitochondrial genome diversity: the genomics revolution. In: Wendel FJ, Greilhuber J, Dolezel J, Leitch JI (eds) Plant genome diversity, Plant genomes, their residents, and their evolutionary dynamics, vol 1. Springer, Vienna, pp 123–144
- Mwinyi A, Meyer A, Bleidorn C, Lieb B, Bartolomaeus T, Podsiadlowski L (2009) Mitochondrial genome sequence and gene order of *Sipunculus nudus* give additional support for an inclusion of Sipuncula into Annelida. *BMC Genomics* 10:27
- Nakabachi A, Ishida K, Hongoh Y, Ohkuma M, Miyagishima S-Y (2014) Aphid gene of bacterial origin encodes a protein transported to an obligate endosymbiont. *Curr Biol* 24:R640–R641
- Nosek J, Tomáška Ľ (2003) Mitochondrial genome diversity: evolution of the molecular architecture and replication strategy. *Curr Genet* 44:73–84
- Novakova E, Hypša V, Klein J, Foottit R, Von Dohlen CD, Moran NA (2013) Reconstructing the phylogeny of aphids (Hemiptera: Aphididae) using DNA of the obligate symbiont *Buchnera aphidicola*. *Mol Phylogenet Evol* 68:42–54
- Nowack ECM (2014) *Paulinella chromatophora* – rethinking the transition from endosymbiont to organelle. *Acta Soc Bot Pol* 83:387–397

References

- Nowack ECM, Grossman AR (2012) Trafficking of protein into the recently established photosynthetic organelles of *Paulinella chromatophora*. *Proc Natl Acad Sci U S A* 109:5340–5345
- Nowack ECM, Melkonian M, Glöckner G (2008) Chromatophore genome sequence of *Paulinella* sheds light on acquisition of photosynthesis by eukaryotes. *Curr Biol* 18:410–418
- Nowack ECM, Price DC, Bhattacharya D, Singer A, Melkonian M, Grossman AR (2016) Gene transfers from diverse bacteria compensate for reductive genome evolution in the chromatophore of *Paulinella chromatophora*. *Proc Natl Acad Sci U S A* 113:12214–12219
- Pais R, Lohs C, Wu Y, Wang J, Aksoy S (2008) The obligate mutualist *Wigglesworthia glossinidia* Influences reproduction, digestion, and immunity processes of Its host, the Tsetse fly. *Appl Environ Microbiol* 74:5965–5974
- Passamonti M, Ricci A, Milani L, Ghiselli F (2011) Mitochondrial genomes and doubly uniparental inheritance: new insights from *Musculista senhousia* sex-linked mitochondrial DNAs (Bivalvia Mytilidae). *BMC Genomics* 12:442
- Pittis AA, Gabaldón T (2016) Late acquisition of mitochondria by a host with chimaeric prokaryotic ancestry. *Nature* 531:101–104
- Pons J, Barraclough TG, Gomez-Zurita J, Cardoso A, Duran DP, Hazell S, Kamoun S, Sumlin WD, Vogler AP (2006) Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Syst Biol* 55:595–609
- Ratnasingham S, Hebert PDN (2007) BOLD: the barcode of life data system (<http://www.barcodinglife.org>). *Mol Ecol Notes* 7:355–364
- Rice DW, Alverson AJ, Richardson AO, Young GJ, Sanchez-Puerta MV, Munzinger J, Barry K, Boore JL, Zhang Y, dePamphilis CW, Knox EB, Palmer JD (2013) Horizontal transfer of entire genomes via mitochondrial fusion in the angiosperm *Amborella*. *Science* 342:1468–1473
- Richardson AO, Rice DW, Young GJ, Alverson AJ, Palmer JD (2013) The «fossilized» mitochondrial genome of *Liriodendron tulipifera*: ancestral gene content and order, ancestral editing sites, and extraordinarily low mutation rate. *BMC Biol* 11:29
- Ruhfel B, Gitzendanner M, Soltis P, Soltis D, Burleigh J (2014) From algae to angiosperms-inferring the phylogeny of green plants (Viridiplantae) from 360 plastid genomes. *BMC Evol Biol* 14:23
- Saunders GW, McDevitt DC (2012) Methods for DNA barcoding photosynthetic protists emphasizing the macroalgae and diatoms. In: Kress JW, Erickson LD (eds) *DNA barcodes: methods and protocols*. Humana Press, Totowa, pp 207–222
- Schoch CL, Seifert KA, Huhndorf S, Robert V, Spouge JL, Levesque CA, Consortium FB (2012) Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proc Natl Acad Sci U S A* 109:6241–6246
- Shearer TL, van Oppen MJH, Romano SL, Wörheide G (2002) Slow mitochondrial DNA sequence evolution in the Anthozoa (Cnidaria). *Mol Ecol* 11:2475–2487
- Shoguchi E, Shinzato C, Hisata K, Satoh N, Mungpakdee S (2015) The large mitochondrial genome of *Symbiodinium minutum* reveals conserved noncoding sequences between dinoflagellates and api-complexans. *Genome Biol Evol* 7:2237–2244
- Sloan DB, Alverson AJ, Chuckalovcak JP, Wu M, McCauley DE, Palmer JD, Taylor DR (2012) Rapid evolution of enormous, multichromosomal genomes in flowering plant mitochondria with exceptionally high mutation rates. *PLoS Biol* 10:e1001241
- Sogin ML (1989) Evolution of eukaryotic microorganisms and their small subunit ribosomal RNAs. *Am Zool* 29:487–499
- Strittmatter P, Soll J, Bölter B (2010) The chloroplast protein import machinery: a review. In: Economou A (ed) *Protein secretion, Methods in molecular biology*, vol 619. Humana Press, New York, pp 307–321
- Suzuki K, Miyagishima S-Y (2010) Eukaryotic and eubacterial contributions to the establishment of plastid proteome estimated by large-scale phylogenetic analyses. *Mol Biol Evol* 27:581–590
- Tengs T, Dahlberg OJ, Shalchian-Tabrizi K, Klaveness D, Rudi K, Delwiche CF, Jakobsen KS (2000) Phylogenetic analyses indicate that the 19'Hexanoyloxy-fucoxanthin-containing dinoflagellates have tertiary plastids of haptophyte origin. *Mol Biol Evol* 17:718–729
- Timmis JN, Ayliffe MA, Huang CY, Martin W (2004) Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nat Rev Genet* 5:123–135
- Tindall BJ, Rosselló-Móra R, Busse H-J, Ludwig W, Kämpfer P (2010) Notes on the characterization of prokaryote strains for taxonomic purposes. *Int J Syst Evol Microbiol* 60:249–266
- Valentini A, Pompanon F, Taberlet P (2009) DNA barcoding for ecologists. *Trends Ecol Evol* 24:110–117

- van Hoek AHAM, van Alen TA, Sprakel VSI, Leunissen JAM, Brigge T, Vogels GD, Hackstein JHP (2000) Multiple acquisition of methanogenic archaeal symbionts by anaerobic ciliates. *Mol Biol Evol* 17:251–258
- Voigt O, Erpenbeck D, Wörheide G (2008) A fragmented metazoan organellar genome: the two mitochondrial chromosomes of *Hydra magnipapillata*. *BMC Genomics* 9:350
- von Dohlen CD, Kohler S, Alsop ST, McManus WR (2001) Mealybug beta-proteobacterial endosymbionts contain gamma-proteobacterial symbionts. *Nature* 412:433–436
- Waller RF, Jackson CJ (2009) Dinoflagellate mitochondrial genomes: stretching the rules of molecular biology. *BioEssays* 31:237–245
- Wang Z, Wu M (2014) Phylogenomic reconstruction indicates mitochondrial ancestor was an energy parasite. *PLoS One* 9:e110685
- Wang Z, Wu M (2015) An integrated phylogenomic approach toward pinpointing the origin of mitochondria. *Sci Rep* 5:7949
- Weiss B, Aksoy S (2011) Microbiome influences on insect host vector competence. *Trends Parasitol* 27:514–522
- Werren JH, Baldo L, Clark ME (2008) *Wolbachia*: master manipulators of invertebrate biology. *Nat Rev Microbiol* 6:741–751
- Werren JH, Zhang W, Guo LR (1995) Evolution and phylogeny of *Wolbachia*: reproductive parasites of arthropods. *Proc R Soc Lond B Biol Sci* 261:55–63
- Wiemers M, Fiedler K (2007) Does the DNA barcoding gap exist? – a case study in blue butterflies (Lepidoptera: Lycaenidae). *Front Zool* 4:8
- Yagi Y, Shiina T (2014) Recent advances in the study of chloroplast gene expression and its evolution. *Front Plant Sci* 5:61
- Yu DW, Ji Y, Emerson BC, Wang X, Ye C, Yang C, Ding Z (2012) Biodiversity soup: metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. *Methods Ecol Evol* 3:613–623
- Zhang J, Kapli P, Pavlidis P, Stamatakis A (2013) A general species delimitation method with applications to phylogenetic placements. *Bioinformatics* 29:2869–2876
- Zug R, Hammerstein P (2012) Still a host of hosts for *Wolbachia*: analysis of recent data suggests that 40% of terrestrial arthropod species are infected. *PLoS One* 7:e38544

Phylogenomics

An Introduction

Bleidorn, C.

2017, XIII, 222 p. 89 illus., 87 illus. in color., Hardcover

ISBN: 978-3-319-54062-7