Genome Biology and Evolution

SMBE

Gene Conversion Shapes Linear Mitochondrial Genome Architecture

David Roy Smith* and Patrick J. Keeling

Canadian Institute for Advanced Research, Department of Botany, University of British Columbia, Vancouver, Canada

*Corresponding author: E-mail: smithdr@dal.ca.

Accepted: April 3, 2013

Abstract

Recently, it was shown that gene conversion between the ends of linear mitochondrial chromosomes can cause telomere expansion and the duplication of subtelomeric loci. However, it is not yet known how widespread this phenomenon is and how significantly it has impacted organelle genome architecture. Using linear mitochondrial DNAs and mitochondrial plasmids from diverse eukaryotes, we argue that telomeric recombination has played a major role in fashioning linear organelle chromosomes. We find that mitochondrial telomeres frequently expand into subtelomeric regions, resulting in gene duplications, homogenizations, and/or fragmentations. We suggest that these features are a product of subtelomeric gene conversion, provide a hypothetical model for this process, and employ genetic diversity data to support the idea that the greater the effective population size the greater the potential for gene conversion between subtelomeric loci.

Key words: gene duplication, inverted repeat, mitochondrial DNA, nucleotide diversity, plasmid, telomere.

Mitochondrial genomes are often thought of as circular molecules, but linear forms exist throughout the eukaryotic domain (Nosek et al. 2004) and may represent the norm rather than the exception (Bendich 2007). In many cases, linear mitochondrial DNAs (mtDNAs) have evolved telomeres, which are believed to help maintain the ends of the chromosome, independent of telomerase (Nosek et al. 2006), but see Santos et al. (2004). Mitochondrial telomeres are typically a few hundred to a few thousand nucleotides long, almost always arranged in an inverted orientation, where the sequence of the left telomere forms a palindromic repeat with that of the right (Nosek et al. 2004), and are known to be recombinogenic (Morin and Cech 1988). Palindromic telomeres (often called terminal inverted repeats) are also found on other types of linear organelle chromosomes, including mitochondrial plasmids (Handa 2008) and the nucleomorph genomes of chlorarachniophyte and cryptophyte algae (Douglas et al. 2001; Gilson et al. 2006).

Recent analysis of the winged box jellyfish mtDNA (Smith et al. 2012) uncovered a bizarre genome organization. The genome has fragmented into eight linear chromosomes with identical palindromic telomeres, which have expanded through gene conversion into coding regions within the subtelomeres, causing gene duplication, homogenization, and fragmentation events. Some of these events have affected coding regions. For example, a 150-nt segment of a rRNA-coding gene was shown to be identical to the 5′-end of *nad2*, and a *trnM* gene along with a segment of *cox3* have spread to the subtelomeres of three different chromosomes. The predicted large effective population size and high mtDNA mutation rate of winged box jellyfish were hypothesized to have been catalysts for these subtelomeric conversion events (Smith et al. 2012).

It is largely unknown whether subtelomeric gene conversion has impacted the layout of linear organelle chromosomes within other eukaryotic lineages. This is because until recently there were limited data on linear organelle DNAs, and those that were available often lacked sequence information from the telomeres and subtelomeres. Moreover, for some genomes, the evidence for subtelomeric gene conversion may have been overlooked as it often involves short tracts that are difficult to detect (Voigt et al. 2008). Here, using newly available linear mtDNA and mitochondrial plasmid sequences from across the eukaryotic tree of life, we argue that subtelomeric gene conversion is a commonly occurring phenomenon in mitochondrial systems, and one that has significantly impacted organelle chromosome architecture.

Available Linear Mitochondrial Chromosomes and Their Architecture

We searched the literature for mtDNAs and mitochondrial plasmids that 1) assemble as linear chromosomes, 2) contain defined palindromic telomeres, 3) migrate as linear molecules in gel-electrophoresis analyses (or map as such in restriction digest experiments), and 4) are completely sequenced. We identified and downloaded from GenBank 34 mtDNAs and 18 mitochondrial plasmids that meet these criteria (table 1; supplementary table S1, Supplementary Material online). These genomes come from an assortment of species, such as apicomplexan parasites, ciliates, fungi, green algae, land plants, and cnidarians, and include model organisms, such as *Chlamydomonas reinhardtii*, *Hydra oligactis*, *Physarum polycephalum,* and *Tetrahymena thermophila*. These linear DNAs also run the gamut of size (5–70 kb), nucleotide composition (16–57% GC), coding content (2 to >50 protein-coding genes), compactness (55 to >95% coding), and chromosome number (1–8), and some have fragmented and scrambled ribosomal RNA-coding regions (Boer and Gray 1988; Kairo et al. 1994) or horizontally acquired genes (Kayal et al. 2012). All 52 mitochondrial genomes and plasmids have palindromic telomeres, ranging in length from 50 to >15,000 nt, and some of the telomeres are known to terminate with elaborate conformations, such as 3′- or 5′-overhangs, single-stranded closed loops, or covalently bound proteins (Vahrenholz et al. 1993; Nosek et al. 1995; Fricova et al. 2010; Smith et al. 2010).

Duplicate and Homogenized Genes in the Telomeres of Linear Organelle DNAs

We explored the telomeres and subtelomeres of each linear organelle DNA in the data set and uncovered two reoccurring features: subtelomeric gene duplication (when a coding region is completely or partially incorporated into the palindromic telomeres and therefore repeated at each end of the chromosome) and subtelomeric gene homogenization (when segments of two unique subtelomeric loci, each located at different chromosome ends, are identical to one another) (fig. 1). Indeed, 35 (67%) of the 52 linear organelle DNAs show signs of subtelomeric duplication and/or homogenization, of which 22 are mtDNAs and 13 mitochondrial plasmids (table 1). These genomes span all the lineages in our data set, with the exception of green algae.

The types and number of genes involved in these events and the length of the duplication/homogenization tracts vary extensively among the organelle DNAs, ranging from a single protein-, tRNA-, or rRNA-coding region to more than 30 different loci (supplementary table S1, Supplementary Material online). For example, the mitochondrial telomeres of the stramenopile alga *Ochromonas danica* contain duplicates of four tRNAs and a 160-nt fragment of *nad11*, those of the apicomplexan *Theileria equi* harbor replicates of two rRNA-coding regions and a complete copy of *cox3* (Hikosaka et al. 2010), and those of the fungus *Candida viswanathii* have a *trnM* gene and a 240-nt segment of *atp6* (Valach et al. 2011). Similarly, the telomeres of the mitochondrial plasmids from the slime mold *P. polycephalum* and the fungi *Fusarium proliferatum* and *Neurospora intermedia* all contain a 50–700-nt piece of a DNA polymerase gene (*DNApol*) (Chan et al. 1991; Takano et al. 1994; Láday et al. 2008). In the plasmid subtelomeres of the plant *Brassica napus* and fungi *Pichia kluyveri* and *Podospora anserina,* a *DNApol* and an RNA polymerase gene (*RNApol*) have homogenized with one another, so that the first 32–72 residues of their deduced N-terminal amino acid sequences are identical (fig. 1) (Hermanns and Osiewacz 1992; Blaisonneau et al. 1999; Handa et al. 2002). Subtelomeric gene homogenizations also abound in the mitochondrial genomes of cnidarians, such as the moon jelly *Aurelia aurita* and the Atlantic sea nettle *Chrysaora quinquecirrha* (Shao et al. 2006; Park et al. 2012), where the 3′-ends of *cob* and *polB* are identical (but in different reading frames), as well as in the multipartite mtDNAs of *H. magnipapillata* and the box jellyfish *Alatina moseri* (Voigt et al. 2008; Smith et al. 2012), where protein-coding and structural RNA genes have spread their sequences to varying degrees to other subtelomeric loci (fig. 1). Overall, these data suggest that the ends of linear organelle DNAs are prone to major mutational events.

Telomere Expansion through Gene Conversion

Given their abundance and widespread occurrence, the subtelomeric duplication and homogenization events among the linear organelle DNAs that we investigated are likely the result of a common mechanism. One mechanism that is known to shape telomeres in nuclear genomes (McEachern and Iyer 2001; Linardopoulou et al. 2005) and cause gene homogenization, duplication, and inverted repeat expansion in organelle genomes (Goulding 1996; Hao et al. 2010) is gene conversion—a type of genetic recombination where a segment of DNA on one chromosome is converted to that of another through mismatch repair (see Chen et al. [2007] and Maréchal and Brisson [2010] for detailed depictions of the mechanism). Gene conversion has proven to be a significant evolutionary force in organelle systems (Khakhlova and Bock 2006; Hao et al. 2010) and is believed to play a part in organelle DNA maintenance, including the repair of double-stranded breaks (Maréchal and Brisson 2010). Also, a leading model for organelle DNA replication, called recombination-dependent replication, involves gene conversion between long palindromic repeats of complex branched linear DNA molecules (Bendich 2004; Maréchal and Brisson 2010).

Because of their palindromic repeat structure and the large number of organelle chromosomes per organelle, one would expect the telomeres of linear mtDNAs and plasmids to be magnets for gene conversion. Moreover, the ends of linear chromosomes tend to be more recombinogenic and experience more double-stranded breaks than other chromosomal regions (Linardopoulou et al. 2005; Barton et al. 2008), which

Table 1

Architecture of Linear Organelle Genomes with Palindromic Telomeres

Species	Genome Size (kb)	Proteins[a]	Telomere Size (kb)[b]	Subtelomeric Gene Duplication and/or Homogenization[c]	GenBank Accession
Mitochondrial DNA					
Apicomplexans					
Babesia bigemina	5.9	3	0.13	Yes	AB499085
Babesia bovis	6	3	0.16	No	NC_009902
Babesia caballi	5.8	3	0.06	No	AB499086
Babesia gibsoni	5.9	3	0.08	Yes	AB499087
Theileria annulata	5.9	3	0.08	No	NT_167255
Theileria equi	8.2	3	1.6	Yes	AB499091
Theileria orientalis	6.0	3	0.05	No	AB499090
Theileria parva	5.9	3	0.1	Yes	AB499089
Ciliates					
Ichthyophthirius multifiliis	51.7	>30	6.3	Yes	NC_015981
Oxytricha trifallax	69.8	>30	1.5–2.5	Yes	JN383843
Tetrahymena malaccensis	47.7	>30	2.9	Yes	NC_008337
Tetrahymena paravorax	47.5	>30	3.3	Yes	NC_008338
Tetrahymena pigmentosa	47	>30	2.9	Yes	NC_008339
Tetrahymena pyriformis	47.3	>30	2.9	Yes	NC_000862
Tetrahymena thermophila	47.6	>30	2.8	Yes	AF396436
Cnidarians					
Alatina moseri[d]	29	15	0.8–1.1	Yes	JN642329–44
Aurelia aurita "White sea"	16.9	15	0.42	Yes	NC_008446
Aurelia aurita "Yellow sea"	17	15	0.42	Yes	HQ694729
Chrysaora quinquecirrha	~17	15	>0.3	Yes	HQ694730
Clava multicornis	~17	13	>0.2	Yes	NC_016465
Hydra magnipapillata[d]	15.9	13	0.2–0.4	Yes	NC_011220–1
Hydra oligactis	16.3	13	1.5	Yes	NC_010214
Laomedea flexuosa	~16	13	>0.3	Yes	NC_016463
Fungi					
Candida parapsilosis	32.8	14	1.9	No	DQ376035
Candida subhashii	29.8	15	0.7	No	NC_014337
Candida viswanathii	39.2	14	5	Yes	EF536359
Hyaloraphidium curvatum	29.6	14	1.2	No	AF402142
Pneumocystis carinii	22.9	14	0.3–1.2	No	NC_013660
Green algae					
Chlamydomonas reinhardtii	15.8	8	0.5	No	EU306622
Polytomella capuana	13	7	0.9	No	EF645804
Polytomella parva[d]	16.2	7	1.3	No	AY062933–4
Polytomella piriformis[d]	16.1	7	1.3	No	GU108480–1
Heterokonts					
Ochromonas danica	41	33	2.2	Yes	NC_002571
Proteromonas lacertae	48.7	57	15.6	Yes	NC_014338
Mitochondrial plasmid (plasmid name)					
Amoebozoan					
Physarum polycephalum (mF)	14.5	≥2	1.5–1.8	Yes	D29637
Fungi					
Ascobolus immersus (pAI2)	5.1	1	0.6	No	X15982
Blumeria graminis (pBgh)	8	2	0.7	Yes	NC_004935
Claviceps purpurea (pClK1)	6.8	2	0.3	Yes	X15648
Fusarium proliferatum (pFP1)	10.3	2	0.4	Yes	NC_010425
Gelasinospora sp. G114 (pKal)	8.2	2	1.1	Yes	L40494
Moniliophthora roreri (pMR2)	11.5	2	0.05	No	NC_015334
Morchella conica (pMC3-2)	6	1	0.7	No	X63909

(continued)

Table 1 Continued

Species	Genome Size (kb)	Proteins[a]	Telomere Size (kb)[b]	Subtelomeric Gene Duplication and/or Homogenization[c]	GenBank Accession
Neurospora crassa (pMaranhar)	7.1	2	0.4	Yes	X55361
Neurospora intermedia (pHarbin-3)	7.1	2	0.4	Yes	NC_000843
Neurospora intermedia (pKalilo)	8.6	2	1.4	Yes	X52106
Pichia kluyveri (pPK2)	7.2	2	0.5	Yes	Y11606
Pleurotus ostreatus (mlp1)	9.9	2	0.4	No	AF126285
Podospora anserina (pAL2-1)	8.4	2	1	Yes	X60707
Land plants					
Beta vulgaris (p10.4)	10.4	≥2	0.4	No	Y10854
Brassica napus (p11.6)	11.6	≥2	0.3	Yes	AB073400
Zea mays (pS-1)	6.4	≥1	0.2–1.5	Yes	X02451
Zea mays (pS-2)	5.5	≥1	0.2–1.5	Yes	J01426

[a] Number of protein-coding genes based on GenBank accession. Number may change as annotations improve. We tried to ignore intronic and hypothetical ORFs, but for some species, particularly ciliates, it was difficult to distinguish them from standard ORFs. Duplicate genes were counted only once.

[b] Telomere size includes genes and gene fragments that have been incorporated into the palindromic repeat. Values should be considered approximations because in many cases the extreme ends of the genome are not yet sequenced.

[c] Subtelomeric gene duplication is when a protein-, rRNA-, or tRNA-coding sequence is completely or partially incorporated into the terminal palindromic repeats (telomeres) and thus present twice (or more) within the genome: once at each end of a chromosome or chromosomes. Subtelomeric gene homogenization is when a coding region located within, or close to, the terminal palindromic repeats spreads its sequence to other coding or noncoding regions adjacent to the telomeres. See figure 1 for examples. Homogenization events ≥ 5 nt were considered. The genes involved in subtelomeric duplication and/or homogenization and the length of the homogenization tracts are listed in supplementary table S1, Supplementary Material online.

[d] The mitochondrial genomes from these species are fragmented into more than one chromosome: *Alatina moseri* (8), *H. magnipapillata* (2), *P. parva* (2), and *P. piriformis* (2). Genome size is based on the concatenation of all chromosomes.

might further contribute to high rates of telomeric gene conversion in linear organelle DNAs.

If organelle telomeres are gene conversion hotspots, then it could explain the propensity of gene duplications and homogenizations observed within these regions. In figure 2, we outline a model for how gene conversion among the ends of linear organelle chromosomes can lead to telomere expansion and the duplication, fragmentation, and homogenization of subtelomeric loci. In this model, recombination between terminal palindromic repeats results in the branch migration of a Holliday junction across the telomere/subtelomere border and the formation of heteroduplex DNA. Gene conversion then occurs through mismatch repair of one of the two strands in the heteroduplex, causing the expansion of the telomeres into subtelomeric regions. Recurrent telomeric recombination, heteroduplex formation, and mismatch repair eventually cause subtelomeric gene duplication and homogenization (fig. 2*A*–*D*).

Mitochondrial Genetic Diversity and the Frequency of Gene Conversion

If gene conversion is fashioning the ends of linear mitochondrial chromosomes then why did we not observe subtelomeric gene duplication/homogenization in all the mtDNAs and plasmids in our data set? Not one of the green algal mtDNAs and very few of the fungal mtDNAs showed traces of subtelomeric gene conversion (table 1). One reason could be that the frequency of organelle gene conversion differs greatly among the lineages and species we explored.

The rate of gene conversion for a genome is intricately tied to various population genetic forces, including the effective population size—discussed hereafter in terms of the effective number of gene copies per locus in a population (N_g) (Lynch et al. 2006). All else being equal, genomes and genetic loci with a large N_g are expected to have higher rates of gene conversion than those with a small N_g—the reason being that the more copies there are of a given locus, the greater the chance for that locus to recombine with itself. It is difficult to measure N_g, but insights into this fundamental parameter can be gained by measuring within-species nucleotide diversity at noncoding and synonymous sites (π_{silent}). At mutation-drift equilibrium, π_{silent} should approximate $2N_g\mu$: twice the effective number of genes in the population times the mutation rate per nucleotide site per generation (Lynch et al. 2006).

By scanning the literature and mining public databases, we obtained π_{silent} statistics for 10 of the mtDNAs in our data set, five of which show evidence for subtelomeric gene conversion and five that do not (table 2). These diversity data include apicomplexans, ciliates, cnidarians, fungi, and green algae and in most cases were calculated using complete mitochondrial genome sequences from multiple members of a population (see Materials and Methods for details). The π_{silent} estimates differ by more than an order of magnitude

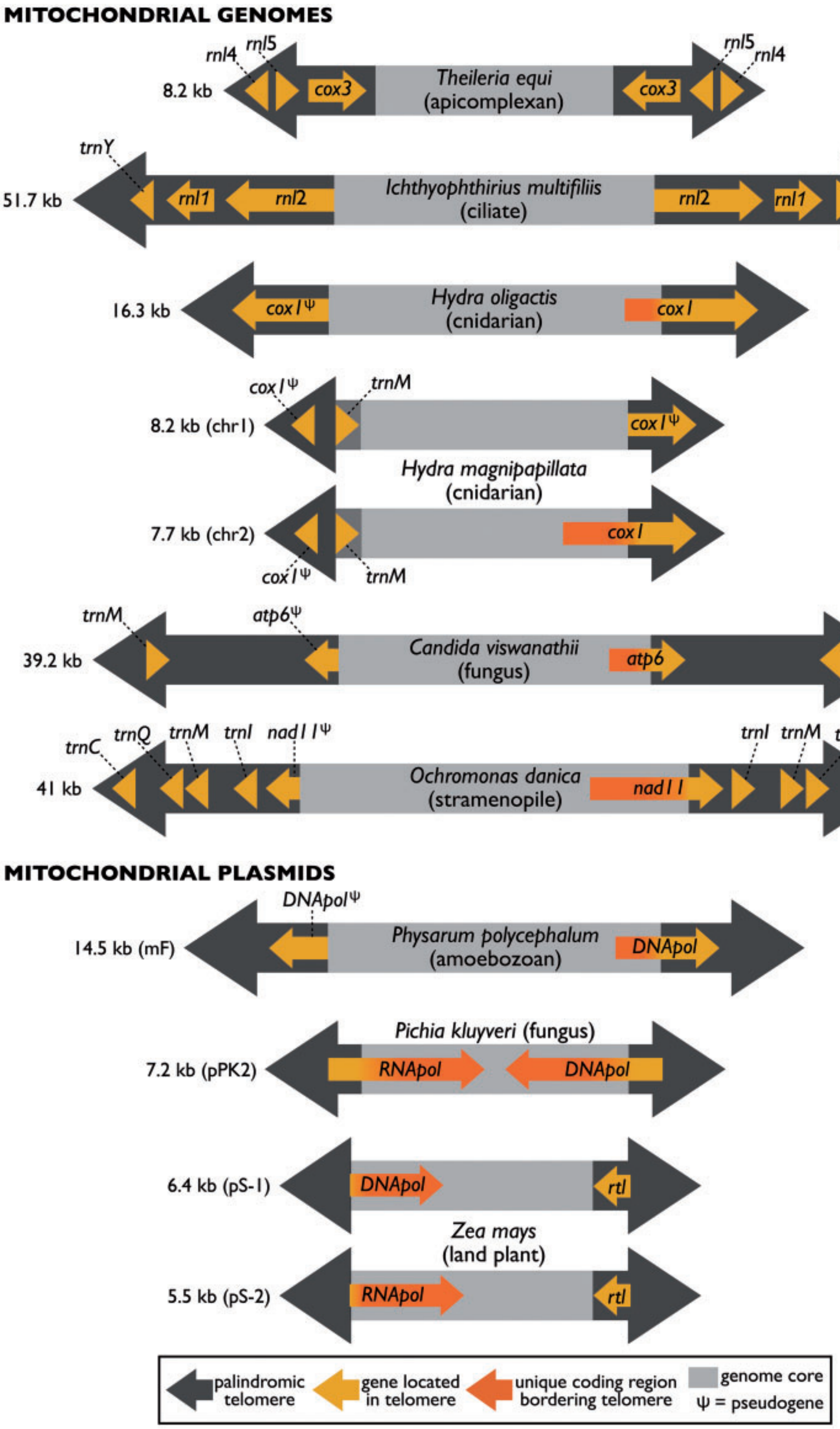

FIG. 1.—Examples of subtelomeric gene duplication and homogenization. Linear mitochondrial genomes and plasmids typically have palindromic telomeres (charcoal), which often contain coding regions (gold). In certain cases, a gene will overlap the telomeric and nontelomeric regions (e.g., *cox1* from the *Hydra oligactis* mtDNA), causing one section of the gene (ψ) to be duplicated at each end of the chromosome (gold) and another section to be present only once (dark orange). Occasionally, segments of two different genes when they each border a telomere can become homogenized, as observed for the N-termini of the *DNApol* and *RNApol* genes from the *Pichia kluyveri* mitochondrial plasmid. See table 1 for more examples of subtelomeric gene duplication and homogenization within linear organelle DNAs. Note: genomes and telomeres are not to scale. Genes within the genomic core (light gray) are not shown. Chromosome lengths (in kilobases) and plasmid names are labeled on the left of the chromosomes. Gene arrows show the transcriptional polarity.

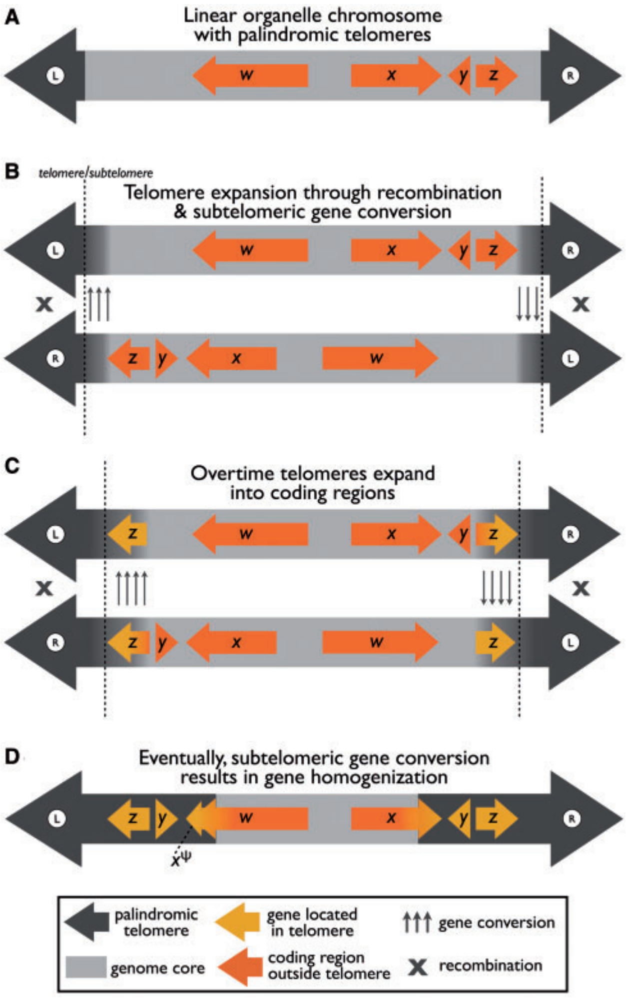

Fig. 2.—Hypothetical model for telomere expansion via gene conversion. (*A*) Linear organelle chromosome with identical palindromic telomeres (charcoal) on the left (L) and right (R) ends and four genes (*w*, *x*, *y*, and *z*) (orange) in the genome core (light gray)—Gene arrows show the transcriptional polarity. (*B*) Two copies of the same linear organelle chromosome aligned in opposite orientations. Recombination (X) between the left and right telomeres of the different chromosomes causes the branch migration of a Holliday junction across the telomere/subtelomere border (dotted line), resulting in the formation of heteroduplex DNA. Heteroduplex is resolved by sequence correction against either strand (in this case the sequence from the right subtelomere), resulting in expansion of the palindromic repeat into noncoding regions of the genome core. (*C*) Overtime, the same processes described in (*B*) result in the expansion of the telomeres into coding regions within the genome core and the almost-complete integration of gene *z* into the chromosome ends. (*D*) Eventually, telomeric recombination and gene conversion cause multiple genes to be integrated into the telomeres and gene homogenization between the C-termini of the *x* and *w* genes.

Table 2

Within-Species Mitochondrial DNA Genetic Diversity

Species	Lineage	Genetic Diversity		Subtelomeric Gene Duplication and/or Homogenization
		Silent Sites	Telomeres	
Alatina moseri	Cnidarian	0.079	0.044	Yes
Hydra magnipapillata	Cnidarian	0.024	0.033	Yes
Babesia bigemina	Apicomplexan	0.021	0.040	Yes
Ichthyophthirius multifiliis	Ciliate	0.013	NA	Yes
Tetrahymena pyriformis	Ciliate	0.012	NA	Yes
Chlamydomonas reinhardtii	Green algae	0.009	0.008	No
Babesia bovis	Apicomplexan	0.005	<0.001	No
Candida parapsilosis	Fungi	<0.001	<0.001	No
Pneumocystis carinii	Fungi	<0.001	<0.001	No
Polytomella parva	Green algae	<0.001	0.018	No

Note.—NA, not available. Genetic diversity is the average pairwise number of nucleotide differences per site. Silent sites include synonymous, noncoding, and/or intronic positions. Values for *A. moseri*, *B. bovis*, *C. parapsilosis*, *C. reinhardtii*, and *P. parva* come from the literature (Smith and Lee 2008, 2011; Smith and Keeling 2012; Smith et al. 2012; Valach et al. 2012). Those for the other species were calculated using available mitochondrial genome data—see Materials and Methods and supplementary table S2, Supplementary Material online, for details.

among the different mitochondrial genomes, ranging from approximately 0.08 in the box jellyfish to less than 0.001 in the fungi *Candida parapsilosis* and *Pneumocystis carinii* and the green alga *Polytomella parva*. The highest π_{silent} values (>0.01) belong to the five species whose mtDNAs harbor subtelomeric duplications and/or homogenizations, such as *H. magnipapillata* and the ciliate *Ichthyophthirius multifiliis*, whereas the lowest values (<0.01) are all found in species whose mitochondrial genomes show no traces of subtelomeric gene conversion, like the green alga *C. reinhardtii* and the fungus *P. carinii* (table 2). These findings are consistent with the hypothesis that a high $2N_g\mu$ can lead to elevated levels of subtelomeric gene conversion and help explain why we did not observe subtelomeric duplication/homogenization events in all the genomes. Nucleotide diversity data from the telomere regions further support these conclusions (table 2).

A multipartite linear genomic architecture might also contribute to elevated levels of subtelomeric gene conversion, provided the different chromosomes harbor similar telomeric repeats. For instance, each copy of the winged box jellyfish mtDNA contains 16 identical palindromic telomeres (eight chromosomes, each with two telomeres). This high telomere copy number per genome likely increases the potential for subtelomeric gene conversion. Some of the highest levels of subtelomeric gene homogenization that we observed were found in multipartite mitochondrial genomes, including that of the winged box jellyfish (table 1).

Conclusion

Linear mitochondrial genomes and plasmids are highly susceptible to subtelomeric gene duplication and homogenization. These mutational events are likely the product of recurrent gene conversion between the ends of organelle chromosomes and appear to occur more frequently in organelle DNAs that harbor large amounts of silent-site diversity. The findings presented here parallel those from studies on gene conversion in the subtelomeres of various nuclear genomes and the inverted repeats of chloroplast DNAs (Goulding 1996), and ultimately highlight the power of gene conversion in shaping organelle genome architecture.

Materials and Methods

The linear mitochondrial genomes and plasmids were downloaded from GenBank on 1 November 2012. We did not include in our data set linear organelle DNAs that lack telomeres, such as the mtDNAs of certain apicomplexan parasites (Hikosaka et al. 2011) or those for which the ends do not form palindromic repeats. We also omitted linear mitochondrial plasmids without annotated protein-coding genes. For each organelle chromosome, we aligned pairwise the left telomere and subtelomere to those of the right with MAFFT (Katoh et al. 2005), implemented through Geneious v6.0.3 (Biomatters Ltd, Auckland, New Zealand), using default parameters and the "auto-fit" algorithm option. For organelle genomes containing more than one chromosome, such as the bipartite mtDNAs of *H. magnipapillata* and *P. parva*, we also performed multiple alignments of the telomeres and subtelomeres from the different chromosomes. We used Basic Local Alignment Search Tool (BLAST)N (executed within Geneious) to search all the organelle genomes (and their different chromosomes) against themselves to uncover potential gene duplication, homogenization, or fragmentation events not identified in the MAFFT alignments.

Mitochondrial genetic diversity data (table 2) came directly from the literature for *A. moseri* (Smith et al. 2012), *Babesia bovis* (Smith and Keeling 2012), *C. parapsilosis* (Valach et al. 2012), *C. reinhardtii* (Smith and Lee 2008), and *P. parva* (Smith and Lee 2011). For all other species (table 2), diversity was calculated with DnaSP v5 (Librado and Rozas 2009), using the Jukes and Cantor correction, and the following strains and loci: *B. bigemina*, complete mtDNAs from five isolates (Kochinda, Bond, BbiS3P, PR, and JG-29); *H. magnipapillata*, complete mtDNAs from two isolates (one from Wuhu, China, and one from Mishima, Japan); *I. multifiliis*: *cox1* sequences from seven isolates (G2 tomont and G2–G7 theront); *Pneumocystis*: complete mtDNAs from two isolates (*carinii* and *jirovecii*); and *T. pyriformis*: *cox1* sequences from six isolates (TRO4 and ATCC strains 30005, 30039, 30202, 30327, and 30331). In certain cases, we measured genetic diversity by assembling complete mtDNAs from next-generation sequencing data in GenBank's Sequence Read Archive. The sources, accession numbers, and assembly details for all the sequences used to calculate genetic diversity are listed in supplementary table S2, Supplementary Material online.

Supplementary Material

Supplementary tables S1 and S2 are available at *Genome Biology and Evolution* online (http://www.gbe.oxfordjournals.org/).

Acknowledgments

This work was supported by a postdoctoral fellowship from the Izaak Walton Killam Memorial Trusts to D.R.S. and by a grant from the Natural Sciences and Engineering Research Council of Canada (227301) to P.J.K. P.J.K. is a Fellow of the Canadian Institute for Advanced Research.

Literature Cited

Barton AB, Pekosz MR, Kurvathi RS, Kaback DB. 2008. Meiotic recombination at the ends of chromosomes in *Saccharomyces cerevisiae*. Genetics 179:1221–1235.

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

Bendich AJ. 2007. The size and form of chromosomes are constant in the nucleus, but highly variable in bacteria, mitochondria and chloroplasts. BioEssays 29:474–483.

Blaisonneau J, Nosek J, Fukuhara H. 1999. Linear DNA plasmid pPK2 of *Pichia kluyveri*: distinction between cytoplasmic and mitochondrial linear plasmids in yeasts. Yeast 15:781–791.

Boer PH, Gray MW. 1988. Scrambled ribosomal RNA gene pieces in *Chlamydomonas reinhardtii* mitochondrial DNA. Cell 55:399–411.

Chan BS, Court DA, Vierula PJ, Bertrand H. 1991. The *kalilo* linear senescence-inducing plasmid of *Neurospora* is an invertron and encodes DNA and RNA polymerases. Curr Genet. 20:225–237.

Chen JM, Cooper DN, Chuzhanova N, Férec C, Patrinos GP. 2007. Gene conversion: mechanisms, evolution and human disease. Nat Rev Genet. 8:762–775.

Douglas S, et al. 2001. The highly reduced genome of an enslaved algal nucleus. Nature 410:1091–1096.

Fricova D, et al. 2010. The mitochondrial genome of the pathogenic yeast *Candida subhashii*: GC-rich linear DNA with a protein covalently attached to the 5′ termini. Microbiology 156:2153–2163.

Gilson PR, et al. 2006. Complete nucleotide sequence of the chlorarachniophyte nucleomorph: nature's smallest nucleus. Proc Natl Acad Sci U S A. 103:9566–9571.

Goulding SE, Olmstead RG, Morden CW, Wolfe KH. 1996. Ebb and flow of the chloroplast inverted repeat. Mol Gen Genet. 252:195–206.

Handa H. 2008. Linear plasmids in plant mitochondria: peaceful coexistences or malicious invasions? Mitochondrion 8:15–25.

Handa H, Itani K, Sato H. 2002. Structural features and expression analysis of a linear mitochondrial plasmid in rapeseed (*Brassica napus* L.). Mol Genet Genomics. 267:797–805.

Hao W, Richardson AO, Zheng Y, Palmer JD. 2010. Gorgeous mosaic of mitochondrial genes created by horizontal transfer and gene conversion. Proc Natl Acad Sci U S A. 107:21576–21581.

Hermanns JJ, Osiewacz HDH. 1992. The linear mitochondrial plasmid pAL2-1 of a long-lived *Podospora anserina* mutant is an invertron encoding a DNA and RNA polymerase. Curr Genet. 22: 491–500.

Hikosaka K, et al. 2010. Divergence of the mitochondrial genome structure in the apicomplexan parasites, *Babesia* and *Theileria*. Mol Biol Evol. 27:1107–1116.

Hikosaka K, et al. 2011. Concatenated mitochondrial DNA of the coccidian parasite *Eimeria tenella*. Mitochondrion 11:273–278.

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

Kairo A, Fairlamb AH, Gobright E, Nene V. 1994. A 7.1 kb linear DNA molecule of *Theileria parva* has scrambled rDNA sequences and open reading frames for mitochondrially encoded proteins. EMBO J. 13: 898–905.

Katoh K, Kuma KI, Toh H, Miyata T. 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res. 33: 511–518.

Kayal E, et al. 2012. Evolution of linear mitochondrial genomes in medusozoan cnidarians. Genome Biol Evol. 4:1–12.

Láday M, Stubnya V, Hamari Z, Hornok L. 2008. Characterization of a new mitochondrial plasmid from *Fusarium proliferatum*. Plasmid 59: 127–133.

Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451–1452.

Linardopoulou EV, et al. 2005. Human subtelomeres are hot spots of interchromosomal recombination and segmental duplication. Nature 437:94–100.

Lynch M, Koskella B, Schaack S. 2006. Mutation pressure and the evolution of organelle genomic architecture. Science 311:1727–1730.

Maréchal A, Brisson N. 2010. Recombination and the maintenance of plant organelle genome stability. New Phytol. 186:299–317.

McEachern MJ, Iyer S. 2001. Short telomeres in yeast are highly recombinogenic. Mol Cell. 7:695–704.

Morin GB, Cech TR. 1988. Mitochondrial telomeres: surprising diversity of repeated telomeric DNA sequences among six species of *Tetrahymena*. Cell 52:367–374.

Nosek J, Dinouel N, Kovac L, Fukuhara H. 1995. Linear mitochondrial DNAs from yeasts: telomeres with large tandem repetitions. Mol Gen Genet. 247:61–72.

Nosek J, Kosa P, Tomáska L. 2006. On the origin of telomeres: a glimpse at the pre-telomerase world. BioEssays 28:182–190.

Nosek J, Tomáska L, Kucejová B. 2004. The chromosome end replication: lessons from mitochondrial genetics. J Appl Biomed. 2:71–79.

Park E, et al. 2012. Estimation of divergence times in cnidarian evolution based on mitochondrial protein-coding genes and the fossil record. Mol Phylogenet Evol. 62:329–345.

Santos JH, Meyer JN, Skorvaga M, Annab LA, Van Houten B. 2004. Mitochondrial hTERT exacerbates free-radical-mediated mtDNA damage. Aging Cell. 3:399–411.

Shao Z, Graf S, Chaga OY, Lavrov DV. 2006. Mitochondrial genome of the moon jelly *Aurelia aurita* (Cnidaria, Scyphozoa): a linear DNA molecule encoding a putative DNA-dependent DNA polymerase. Gene 381: 92–101.

Smith DR, et al. 2012. First complete mitochondrial genome sequence from a box jellyfish reveals a highly fragmented linear architecture and insights into telomere evolution. Genome Biol Evol. 4:52–58.

Smith DR, Hua J, Lee RW. 2010. Evolution of linear mitochondrial DNA in three known lineages of *Polytomella*. Curr Genet. 56:427–438.

Smith DR, Keeling PJ. 2012. Twenty-fold difference in evolutionary rates between the mitochondrial and plastid genomes of species with secondary red plastids. J Eukaryot Microbiol. 59:181–184.

Smith DR, Lee RW. 2008. Nucleotide diversity in the mitochondrial and nuclear compartments of *Chlamydomonas reinhardtii*: investigating the origins of genome architecture. BMC Evol Biol. 8:156.

Smith DR, Lee RW. 2011. Nucleotide diversity of the colorless green alga *Polytomella parva* (Chlorophyceae, Chlorophyta): high for the mitochondrial telomeres, surprisingly low everywhere else. J Eukaryot Microbiol. 58:471–473.

Takano H, Kawano S, Kuroiwa T. 1994. Genetic organization of a linear mitochondrial plasmid (mF) that promotes mitochondrial fusion in *Physarum polycephalum*. Curr Genet. 26:506–511.

Vahrenholz C, Riemen G, Pratje E, Dujon B, Michaelis G. 1993. Mitochondrial DNA of *Chlamydomonas reinhardtii*: the structure of the ends of the linear 15.8-kb genome suggests mechanisms for DNA replication. Curr Genet. 24:241–247.

Valach M, et al. 2011. Evolution of linear chromosomes and multipartite genomes in yeast mitochondria. Nucleic Acids Res. 39:4202–4219.

Valach M, et al. 2012. Mitochondrial genome variability within the *Candida parapsilosis* species complex. Mitochondrion 12:514–519.

Voigt O, Erpenbeck D, Wörheide G. 2008. A fragmented metazoan organellar genome: the two mitochondrial chromosomes of *Hydra magnipapillata*. BMC Genomics 9:350.

Associate editor: Gertraud Burger