International Journal of Biological Macromolecules 242 (2023) 124568

Contents lists available at ScienceDirect

International Journal of Biological Macromolecules

journal homepage: www.elsevier.com/locate/ijbiomac

Rapid evolution, rearrangements and whole mitogenome duplication in the Australian stingless bees *Tetragonula* (Hymenoptera: Apidae): A steppingstone towards understanding mitochondrial function and evolution

Elaine Françoso [a,b,*], Alexandre Rizzo Zuntini [c], Paulo Cseri Ricardo [b], Priscila Karla Ferreira Santos [d], Natalia de Souza Araujo [e], João Paulo Naldi Silva [b], Leonardo Tresoldi Gonçalves [f], Rute Brito [g], Rosalyn Gloag [h], Benjamin A. Taylor [i], Brock A. Harpur [i], Benjamin P. Oldroyd [h], Mark J.F. Brown [a], Maria Cristina Arias [b]

[a] *Centre for Ecology, Evolution and Behaviour, Department of Biological Sciences, School of Life Sciences and the Environment, Royal Holloway University of London, Egham TW20 0EX, UK*
[b] *Departamento de Genética e Biologia Evolutiva, Instituto de Biociências, Universidade de São Paulo, São Paulo, SP 05508-090, Brazil*
[c] *Royal Botanic Gardens, Kew TW9 3AE, UK*
[d] *Department of Biology, Utah State University, Logan, UT, USA*
[e] *Unit of Evolutionary Biology & Ecology, Université libre de Bruxelles (ULB), Brussels, Belgium*
[f] *Departamento de Genética, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil*
[g] *Instituto de Biotecnologia, Universidade Federal de Uberlândia, Uberlândia, MG, Brazil*
[h] *School of Life and Environmental Sciences, The University of Sydney, NSW, 2006, Australia*
[i] *Department of Entomology, Purdue University, West Lafayette, Indiana, USA*

ARTICLE INFO

Keywords:
Rapid evolution
Mitogenome duplication
Mitogenome rearrangements
Bees
Tetragonula carbonaria
T. hockingsi

ABSTRACT

The extreme conservation of mitochondrial genomes in metazoans poses a significant challenge to understanding mitogenome evolution. However, the presence of variation in gene order or genome structure, found in a small number of taxa, can provide unique insights into this evolution. Previous work on two stingless bees in the genus *Tetragonula* (*T. carbonaria* and *T. hockingsi*) revealed highly divergent *CO1* regions between them and when compared to the bees from the same tribe (Meliponini), indicating rapid evolution. Using mtDNA isolation and Illumina sequencing, we elucidated the mitogenomes of both species. In both species, there has been a duplication of the whole mitogenome to give a total genome size of 30,666 bp in *T. carbonaria*; and 30,662 bp in *T. hockingsi*. These duplicated genomes present a circular structure with two identical and mirrored copies of all 13 protein coding genes and 22 tRNAs, with the exception of a few tRNAs that are present as single copies. In addition, the mitogenomes are characterized by rearrangements of two block of genes. We believe that rapid evolution is present in the whole Indo-Malay/Australasian group of Meliponini but is extraordinarily elevated in *T. carbonaria* and *T. hockingsi*, probably due to founder effect, low effective population size and the mitogenome duplication. All these features - rapid evolution, rearrangements, and duplication - deviate significantly from the vast majority of the mitogenomes described so far, making the mitogenomes of *Tetragonula* unique opportunities to address fundamental questions of mitogenome function and evolution.

1. Introduction

The eukaryotic cell originated at least 1.75 billion years ago, when two prokaryotic lineages (archaeon and bacterium) fused, establishing an intimate symbiotic relationship [1,2]. Most of the genes in the original bacterial genome were transferred to the archaeon genome. In animals, the number of genes in the former archaeon genome increased from 3000 to 20,000; while the number of genes of the ancestral bacterium reduced from 3000 to 37 genes, resulting in the nuclear genome and the mitochondrial genome as we currently know them, respectively

* Corresponding author at: Centre for Ecology, Evolution and Behaviour, Department of Biological Sciences, School of Life Sciences and the Environment, Royal Holloway University of London, Egham TW20 0EX, UK.
E-mail address: francoso.e@gmail.com (E. Françoso).

https://doi.org/10.1016/j.ijbiomac.2023.124568
Received 8 December 2022; Received in revised form 16 March 2023; Accepted 10 April 2023
Available online 24 April 2023

[3]. While the nuclear genome of eukaryotes has been expanded in size, via an increase in the number of introns, proliferation of transposons, gene duplications, and increase in length with intergenic regions, animal mitochondrial genomes have evolved in the opposite direction, losing genes to the nuclear genome [4]. This way, animal mitochondrial genomes normally contain no introns or large intergenic spacers [5], with a few exceptions [6].

Nuclear and mitochondrial genes have an intimate and synchronized interaction that is fundamental to respiratory function and results in ATP production via oxidative phosphorylation (OXPHOS) - one of the most important biological functions - through the mitochondrial electron transport system [3,7]. Four of the five mitochondrial electron transport system subunits are produced by the interaction of the 13 proteins encoded in the mitochondrial genome, together with around 73 proteins encoded by genes from the nuclear genome [7]. Because of its central importance for cellular respiration, the mitochondrial genome is evolutionarily conserved with respect to both the order and number of genes. It is small and relatively uniform in size (15–20 kb in length) among vertebrate and invertebrate animals [5,8]. However, while a circular molecule encoding 13 protein genes, two rRNA genes, and 22 tRNA genes is considered the typical mitogenome structure, several exceptions have been described [6]. Gene rearrangements have been reported in ascidia (tunicates) [9,10], birds [11], reptiles [12–14], and amphibians [15–17]. Mitochondrial genome expansions are known from seed beetles [18] and in ark shell bivalves [19]; a fragmented mitochondrial genome (in which the mitochondrial genes are split across two or more circular mt-chromosomes) was identified in booklice [20]; and linearization has occurred in some terrestrial isopod species [21,22]. In Hymenoptera, significant rearrangements in the gene order have been described so far in the parasitic wasp *Nasonia* [23], in the Neotropical stingless bee *Frieseomelitta varia* [24], and in two species of *Lepidotrigona* [25], a stingless bee genus from Indo-Malay/Australasian group.

Tetragonula carbonaria and *Tetragonula hockingsi* are highly eusocial stingless bees (Apidae: Tribe Meliponini) distributed across tropical and subtropical regions of Australia [26,27]. They are part of a cryptic species complex of endemic Australian *Tetragonula* that includes at least two other species (*T. davenporti* and *T. mellipes*), known as the Carbonaria complex [26,28,29]. Initial efforts to sequence a fragment of the mitochondrial *CO1* (cytochrome oxidase I) gene in these bees returned sequences with gaps and stop codons, which were recently confirmed to be the misamplification of numts (nuclear DNA from mitochondrial origin) [30]. After isolating mtDNA from nuclear DNA, the numts were confirmed in the nuclear fraction but surprisingly, conservative *CO1*-barcode primers for bees and PCR amplification of the pure mitochondrial fraction failed to recover the true mt-*CO1* sequences. Instead, Illumina sequencing of the whole mitogenomes using the purified mtDNA isolate recovered *CO1* sequences in both species that were extremely divergent relative to bees from the same tribe (Meliponini), with high nucleotide polymorphism between *T. carbonaria* and *T. hockingsi*, indicating rapid evolution [30]. The remaining reads of this sequencing showed that the mitogenomes of these species were non-standard, and a challenge to assemble. Here we assemble the complete mitogenomes of *T. carbonaria* and *T. hockingsi*, and then discuss possible links between their enigmatic structure and the extreme rapid mitogenome evolution in this group.

2. Materials and methods

2.1. Sampling, DNA extraction and sequencing

One thorax each of frozen individuals of *T. carbonaria* (Sydney, NSW, 2009, R.Brito Tc40) and *T. hockingsi* (Cardwell, QLD, 2009, R.Brito 179) were used for mitochondrial isolation, following a protocol for isolating insect mitochondrial genomes [31] that separates mitochondria from nucleus by sequential centrifugation. The pure mtDNA extractions from each species were then sequenced using the Miseq sequencer with the Nextera DNA XT library kit for 150 bp reads (paired reads); 20 million reads were generated per library. Sequencing was performed by the Macrogen Company (South Korea).

2.2. Mitogenome assembly and gene annotation

The first mitogenome assemblies and alignments were made using Geneious Pro 10.2.5 software [32], through the DeNovo alignment method and the reference guided method using the complete mitogenomes of other species as reference. The genomes used were recovered from GenBank: *Apis mellifera syriaca* (KY926882), *Bombus ignitus* (NC10967), *Lepidotrigona flavibasis* (MN747147), and *Melipona bicolor* (AF466146). The final mitogenome structure was assembled manually in Geneious Pro 10.2.5 software. Annotations were made using MITOS2 WebServer [33]. For comparative analyses, we selected 22 bee mitogenomes comprising 13 different genera and six different families. The wasp *Philanthus triangulum*, from the superfamily Apoidea, and the shrimp *Squilla mantis,* which represents the ancestral pancrustacean mitochondrial genome organization [34], were used as out-groups.

2.3. Sequence analyses

To confirm the nucleotide sequence of challenging regions, i.e., regions containing inversions, rearrangements, ambiguous tRNA and AT-rich regions found in the mitogenomes of *T. carbonaria* and *T. hockingsi*, 13 primers were designed for PCR amplification and Sanger sequencing (Table S1). PCR tests using different combinations of these primers were made according to Françoso & Arias [35] with the annealing temperature gradient ranging from 38 to 48 °C. The mitogenomes' structures and the ambiguous tRNA were confirmed using the whole genome complete data (nuclear + mitochondrial; long read sequencing) from a *T. carbonaria* individual larva (Brisbane, QLD) generated independently and shipped to Dovetail Genomics for Hi-C processing. These reads were aligned using the duplicated genome as reference in Geneious Pro 10.2.5 software and the overall alignment coverage and quality were evaluated. To confirm the mitogenome size, 30 ng of *T. hockingsi* mtDNA was loaded in a 0.4 % agarose gel with low voltage (37 V) for a total of 10 h, using the lambda DNA/*HindIII* ladder (higher fragment size of 23,130 bp; Promega).

3. Results

3.1. First assembly - characterization of the whole mitogenome of *T. carbonaria* and *T. hockingsi*: idiosyncrasies and linear structure

The mitogenomes of *T. carbonaria* and *T. hockingsi* were assembled, characterized and compared with complete mitogenomes of other bee species that are available in GenBank (Fig. 1 and Table 2). The first version of the mitogenomes of both species presented a linear structure (i.e., the tail did not overlap with the head) and revealed atypical characteristics shared by both species. These are:

(i) A low AT content of ~72 %: this value is lower than those reported for any other bee species reported to date (which range from 78.6 % in *Andrena comellia* to 86.8 % in *Bombus ignitus* and *Melipona scutellaris*), but is nonetheless higher than that of *Squilla mantis*, representing the ancestral pancrustacean mitogenome (70 %; Table 2);

(ii) Rearrangements: An entire block containing the genes *ND6, CytB, ND1, 16S, 12S* and some tRNA genes is translocated and inverted; and another entire block containing the genes *ND5, ND4, ND4L* and some tRNA genes is translocated, though gene order is retained (Fig. 1). These inversions and rearrangements were confirmed by PCR amplification and Sanger sequencing (Figs. S1 and S2), and by mapping the reads of whole genome Illumina sequencing data of another individual of *T. carbonaria* to our

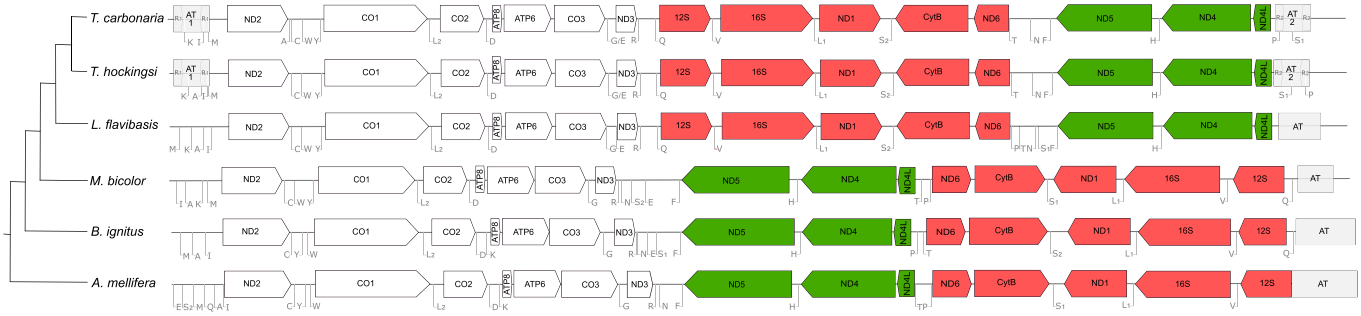

Fig. 1. First assembly (incomplete and linear) of the mitogenomes of *Tetragonula carbonaria* and *Tetragonula hockingsi*, compared to mitogenomes deposited in GenBank of *Lepidotrigona flavibasis* (MN747147), *Melipona bicolor* (AF466146), *Bombus ignitus* (NC10967) and *Apis mellifera syriaca* (KY926882). The relationship among these species is showed before their names, and were based on the published global stingless bee phylogeny (Rasmussen & Cameron 2010). In red are the translocated and inverted genes of *T. carbonaria* and *T. hockingsi,* and in green are the translocated genes of *T. carbonaria* and *T. hockingsi*. Ambiguous tRNA (G/E tRNA, between CO3 and ND3 in *T. carbonaria* and *T. hockingsi*): in which nearly half of the aligned reads generated in the Illumina sequencing presented a cytosine, and the other half present a thymine in the second position of the anticodon, encoding two different amino acids: glutamic acid and glycine, respectively. AT-rich regions (in grey): In *T.carbonaria* and *T. hockingsi,* there is a repeated and mirrored region at the head (R1) of their mitogenomes, with an AT-rich region between them (AT1). At the tail of their mitogenomes, there is another different repeated and mirrored region (R2) with an internal region inside them, being all these regions AT-rich. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

assembly, resulting in consistent and regular coverage across the mitogenome (Fig. S3);

(iii) A repeated region at the head of the mitogenomes (based on a linear representation of the mitogenomes, as shown in Fig. 1): Two duplicated and mirrored regions (R1) were found at the head of each mitogenome, with an internal AT-rich region (AT1; Fig. 1 and Tables 1 and 2). These regions were confirmed by PCR amplification and Sanger sequencing for both species (Figs. S1 and S2);

(iv) A repeated region at the tail of the mitogenomes (as above, as shown in Fig. 1): Two duplicated and mirrored regions were found at the tail of each mitogenome (R2), with an internal region inside them, all these regions being AT rich (AT2; Fig. 1 and Tables 1 and 2). These regions are different from the repeated regions at the head of the mitogenomes and were confirmed by PCR amplification and Sanger sequencing for *T. carbonaria* (Fig. S1);

(v) An ambiguous tRNA: An ambiguous tRNA in which nearly half of the aligned reads generated in the Illumina sequencing of purified mtDNA presented a cytosine (50.6 % and 51.7 % for *T. carbonaria* and *T. hockingsi*, respectively), and the other half present a thymine (49.4 % and 48.3 % for *T. carbonaria* and *T. hockingsi*, respectively) in the second position of the anticodon, encoding two different amino acids, glutamic acid and glycine. This ambiguous tRNA was confirmed by PCR amplification and Sanger sequencing (Figs. S1 and S2), in which double peaks can be observed (representing C and T) at that locus (Figs. S4 and S5). Likewise, the same polymorphism was present in an approximately equal proportion of reads from the additional whole-genome sequencing data of *T. carbonaria*.

3.2. Final assembly - interpreting the first characterization of the whole mitogenome of T. carbonaria and T. hockingsi: a duplicated mitogenome

Although both mitogenomes were easily assembled and all the genes were recovered, the first structure assembled was not circular. Nevertheless, the repeated regions and the two AT-rich regions, together with the ambiguous tRNA found in both species, instead point to a circular structure with the duplication of the entire mitogenome of both species separated from each other by two distinct fragments of the original AT-rich region (Table 1; Figs. 2 and 3).

This duplication was confirmed by PCR and Sanger sequencing of some specific positions (Tables S1 and S2, and Figs. S1 and S2); by agarose gel, that confirmed a genome size >23,130 bp (Fig. S6); and by aligning reads generated by long read sequencing (Hi-C) of another individual of *T. carbonaria* to the duplicated mitogenomes (Fig. S3). The coverage across all genes and other regions was constant and with no gaps, and thus consistent with a full duplication (rather than duplication only of certain genes). Despite the usual level of base variation, expected due to Illumina sequencing error, the reads are all uniform, indicating that the two versions of each gene found in the duplicated mitogenome present the same sequence, except for the ambiguous tRNA and the tRNAs located in the AT regions.

3.3. Differences between the mitogenomes of T. carbonaria and T. hockingsi

The pairwise nucleotide divergence between *T. carbonaria* and *T. hockingsi* across the full mitogenomes was 20.2 %, which is extraordinarily high for what are believed to be recently-diverged congeneric and cryptic species. Among the protein coding genes divergence ranges from 29.2 % (*ATP8*) to 16.4 % (*CO1*). This corroborates previous work showing high pairwise nucleotide divergence at mt-*CO1* between these species and is consistent with a high rate of nucleotide substitution [30]. The main differences in structure between these mitogenomes results from tRNA genes rearrangements. *Tetragonula carbonaria* presents an inversion of $tRNA^{Lys}$ (when comparing with the outgroups) and a translocation of $tRNA^{Ala}$ from cluster 1 to cluster 2 in both sides of each duplicated mitogenome, while *T. hockingsi* presents $tRNA^{Ala}$ as single copy (Fig. 1). Still, the general characterization and structure of their mitogenomes are the same.

4. Discussion

4.1. Duplication of the whole mitogenome of T. carbonaria and T. hockingsi

Here we found un unprecedent duplication of the whole mitogenome of *T. carbonaria* and *T. hockingsi*. Changes in mitogenome architecture are extremely rare due to its importance for cellular respiration [3]. Where they have been reported previously, they have involved a change from a single circular chromosome to either a fragmentation into multiple circular chromosomes, or to a linearized mitogenome [20]. In contrast, *Tetragonula carbonaria* and *T. hockingsi* have a circular mitogenome constituted of two duplicated and mirrored single mitogenomes (Figs. 2 and 3), with two different AT-rich regions between them - a structure known as an amphimer [36]. This atypical structure may be the only form of DNA that can stably retain large duplications with no

Table 1
Characterization of the complete mitochondrial genomes of *Tetragonula carbonaria* and *Tetragonula hockingsi*. The codons of the tRNA genes are shown in brackets. CDS: coding sequence.

Tetragonula carbonaria					*Tetragonula hockingsi*				
Name	Type	Position	Length (bp)	Strand	Name	Type	Position	Length (bp)	Strand
AT1	AT	1–124	124		AT1	AT	1–144	144	
Repeat 1	Repeated region	265–125	141		Repeat 1	Repeated region	145–300	156	
tRNA$^{Lys\ [ttt]}$	tRNA	56–125	70	+	*tRNA$^{Lys\ [ttt]}$*	tRNA	2–71	70	−
tRNA$^{Ile\ [gat]}$	tRNA	169–235	67	−	*tRNA$^{Ala\ [tgc]}$*	tRNA	101–165	65	−
tRNA$^{Met\ [cat]}$	tRNA	262–331	70	+	*tRNA$^{Ile\ [gat]}$*	tRNA	170–237	68	−
ND2	CDS	483–1346	864	+	*tRNA$^{Met\ [cat]}$*	tRNA	265–332	68	+
tRNA$^{Ala\ [tgc]}$	tRNA	1446–1509	64	−	*ND2*	CDS	477–1349	873	+
tRNA$^{Cys\ [gca]}$	tRNA	1448–1513	66	+	*tRNA$^{Cys\ [gca]}$*	tRNA	1460–1526	67	+
tRNA$^{Trp\ [tca]}$	tRNA	1552–1619	68	+	*tRNA$^{Trp\ [tca]}$*	tRNA	1581–1649	69	+
tRNA$^{Tyr\ [gta]}$	tRNA	1643–1708	66	−	*tRNA$^{Tyr\ [gta]}$*	tRNA	1679–1746	68	−
CO1	CDS	1798–3294	1497	+	*CO1*	CDS	1819–3345	1527	+
tRNA$^{Leu2\ [taa]}$	tRNA	3354–3420	67	+	*tRNA$^{Leu2\ [taa]}$*	tRNA	3391–3456	66	+
CO2	CDS	3442–4086	645	+	*CO2*	CDS	3478–4122	645	+
tRNA$^{Asp\ [gtc]}$	tRNA	4112–4180	69	+	*tRNA$^{Asp\ [gtc]}$*	tRNA	4148–4216	69	+
ATP8	CDS	4181–4342	162	+	*ATP8*	CDS	4217–4381	165	+
ATP6	CDS	4348–5019	672	+	*ATP6*	CDS	4456–5058	603	+
CO3	CDS	5028–5795	768	+	*CO3*	CDS	5067–5834	768	+
tRNA$^{Glu\ [ttc]}$	tRNA	5825–5892	68	+	*tRNA$^{Glu\ [ttc]}$*	tRNA	5873–5943	71	+
ND3	CDS	5923–6237	315	+	*ND3*	CDS	5944–6288	345	+
tRNA$^{Arg\ [tcg]}$	tRNA	6246–6311	66	−	*tRNA$^{Arg\ [tcg]}$*	tRNA	6298–6364	67	−
tRNA$^{Gln\ [ttg]}$	tRNA	6449–6515	67	+	*tRNA$^{Gln\ [ttg]}$*	tRNA	6475–6541	67	+
srRNA (12S)	rRNA	6525–7294	770	+	*srRNA (12S)*	rRNA	6550–7322	773	+
tRNA$^{Val\ [tac]}$	tRNA	7294–7360	67	+	*tRNA$^{Val\ [tac]}$*	tRNA	7319–7385	67	+
lrRNA (16S)	rRNA	7349–8733	1385	+	*lrRNA (16S)*	rRNA	7387–8756	1370	+
tRNA$^{Leu1\ [tag]}$	tRNA	8711–8779	69	+	*tRNA$^{Leu1\ [tag]}$*	tRNA	8734–8800	67	+
ND1	CDS	8780–9664	885	+	*ND1*	CDS	8813–9658	846	+
tRNA$^{Ser2\ [tga]}$	tRNA	9705–9771	67	−	*tRNA$^{Ser2\ [tga]}$*	tRNA	9726–9792	67	−
CytB	CDS	9836–10,876	1041	−	*CytB*	CDS	9857–10,930	1074	−
ND6	CDS	10,942–11,430	489	−	*ND6*	CDS	10,963–11,463	501	−
tRNA$^{Thr\ [tga]}$	tRNA	11,505–11,573	69	+	*tRNA$^{Thr\ [tga]}$*	tRNA	11,526–11,593	68	+
tRNA$^{Asn\ [gtt]}$	tRNA	11,676–11,745	70	+	*tRNA$^{Asn\ [gtt]}$*	tRNA	11,710–11,778	69	+
tRNA$^{Phe\ [gaa]}$	tRNA	11,771–11,838	68	−	*tRNA$^{Phe\ [gaa]}$*	tRNA	11,792–11,854	63	−
ND5	CDS	12,088–13,461	1374	−	*ND5*	CDS	12,061–13,482	1422	−
tRNA$^{His\ [gtg]}$	tRNA	13,498–13,564	67	−	*tRNA$^{His\ [gtg]}$*	tRNA	13,519–13,585	67	−
ND4	CDS	13,581–14,858	1278	−	*ND4*	CDS	13,596–14,813	1218	−
ND4L	CDS	14,882–15,133	252	−	*ND4L*	CDS	14,899–15,150	252	−
tRNA$^{Pro\ [tgg]}$	tRNA	15,240–15,305	66	−	*tRNA$^{Pro\ [tgg]}$*	tRNA	15,271–15,336	66	−
tRNA$^{Ser1\ [tct]}$	tRNA	15,365–15,425	61	+	*tRNA$^{Ser1\ [tct]}$*	tRNA	15,374–15,434	61	−
Repeat 2	Repeated region	15,227–15,338	112		Repeat 2	Repeated region	15,241–15,368	128	
AT2	AT	15,339–15,450	112		AT2	AT	15,369–15,438	70	
Repeat 2	Repeated region	15,451–15,562	112		Repeat 2 region	Repeated region	15,566–15,439	128	
tRNA$^{Pro\ [tgg]}$	tRNA	15,486–15,551	66	+	*tRNA$^{Pro\ [tgg]}$*	tRNA	15,471–15,536	66	+
ND4L	CDS	15,658–15,909	252	+	*ND4L*	CDS	15,657–15,908	252	+
ND4	CDS	15,933–17,210	1278	+	*ND4*	CDS	15,994–17,211	1218	+
tRNA$^{His\ [gtg]}$	tRNA	17,227–17,293	67	+	*tRNA$^{His\ [gtg]}$*	tRNA	17,222–17,288	67	+
ND5	CDS	17,330–18,703	1374	+	*ND5*	CDS	17,325–18,746	1422	+
tRNA$^{Phe\ [gaa]}$	tRNA	18,953–19,020	68	+	*tRNA$^{Phe\ [gaa]}$*	tRNA	18,953–19,015	63	+
tRNA$^{Asn\ [gtt]}$	tRNA	19,046–19,115	70	−	*tRNA$^{Asn\ [gtt]}$*	tRNA	19,029–19,097	69	−
tRNA$^{Thr\ [tga]}$	tRNA	19,218–19,286	69	−	*tRNA$^{Thr\ [tga]}$*	tRNA	19,214–19,281	68	−
ND6	CDS	19,361–19,849	489	+	*ND6*	CDS	19,344–19,844	501	+
CytB	CDS	19,915–20,955	1041	+	*CytB*	CDS	19,877–20,950	1074	+
tRNA$^{Ser2\ [tga]}$	tRNA	21,020–21,086	67	+	*tRNA$^{Ser2\ [tga]}$*	tRNA	21,015–21,081	67	+
ND1	CDS	21,127–22,011	885	−	*ND1*	CDS	21,149–21,994	846	−
tRNA$^{Leu1\ [tag]}$	tRNA	22,011–22,079	69	−	*tRNA$^{Leu1\ [tag]}$*	tRNA	22,007–22,073	67	−
lrRNA (16S)	rRNA	22,058–23,442	1385	−	*lrRNA (16S)*	rRNA	22,051–23,420	1370	−
tRNA$^{Val\ [tac]}$	tRNA	23,430–23,496	67	−	*tRNA$^{Val\ [tac]}$*	tRNA	23,488–23,422	0	−
srRNA (12S)	rRNA	23,497–24,266	770	−	*srRNA (12S)*	rRNA	23,485–24,257	773	−
tRNA$^{Gln\ [ttg]}$	tRNA	24,276–24,342	67	−	*tRNA$^{Gln\ [ttg]}$*	tRNA	24,266–24,332	67	−
tRNA$^{Arg\ [tcg]}$	tRNA	24,480–24,545	66	+	*tRNA$^{Arg\ [tcg]}$*	tRNA	24,443–24,509	67	+
ND3	CDS	24,554–24,868	315	−	*ND3*	CDS	24,519–24,863	345	−
tRNA$^{Gly\ [tcc]}$	tRNA	24,899–24,966	68	−	*tRNA$^{Gly\ [tcc]}$*	tRNA	24,864–24,934	71	−
CO3	CDS	24,996–25,763	768	−	*CO3*	CDS	24,973–25,740	768	−
ATP6	CDS	25,772–26,443	672	−	*ATP6*	CDS	25,749–26,351	603	−
ATP8	CDS	26,449–26,610	162	−	*ATP8*	CDS	26,426–26,590	165	−
tRNA$^{Asp\ [gtc]}$	tRNA	26,611–26,679	69	−	*tRNA$^{Asp\ [gtc]}$*	tRNA	26,591–26,659	69	−
CO2	CDS	26,705–27,349	645	−	*CO2*	CDS	26,685–27,329	645	−
tRNA$^{Leu2\ [taa]}$	tRNA	27,371–27,437	67	−	*tRNA$^{Leu2\ [taa]}$*	tRNA	27,351–27,416	66	−
CO1	CDS	27,497–28,993	1497	−	*CO1*	CDS	27,462–28,988	1527	−
tRNA$^{Tyr\ [gta]}$	tRNA	29,083–29,148	66	+	*tRNA$^{Tyr\ [gta]}$*	tRNA	29,061–29,128	68	+
tRNA$^{Trp\ [tca]}$	tRNA	29,172–29,239	68	−	*tRNA$^{Trp\ [tca]}$*	tRNA	29,226–29,158	69	−
tRNA$^{Cys\ [gca]}$	tRNA	29,277–29,342	66	−	*tRNA$^{Cys\ [gca]}$*	tRNA	29,347–29,281	67	−
tRNA$^{Ala\ [tgc]}$	tRNA	29,281–29,344	64	+	*ND2*	CDS	29,458–30,330	873	−

(*continued on next page*)

Table 1 (*continued*)

Tetragonula carbonaria					*Tetragonula hockingsi*				
Name	Type	Position	Length (bp)	Strand	Name	Type	Position	Length (bp)	Strand
ND2	CDS	29,447–30,307	861	–	$tRNA^{Met\ [cat]}$	tRNA	30,475–30,542	68	–
$tRNA^{Met\ [cat]}$	tRNA	30,459–30,528	70	–	$tRNA^{Ile\ [gat]}$	tRNA	30,570–30,637	68	+
$tRNA^{Ile\ [gat]}$	tRNA	30,555–30,621	67	+	Repeat 1	Repeated region	30,507–30,662	156	
Repeat 1	Repeated region	30,525–30,665	141						

Table 2
Characterization of mitochondrial genomes of bees, highlighting the species *Tetragonula carbonaria* and *Tetragonula hockingsi* (in bold), that curiously presented two regions compatible with the AT region in bees. *Philanthus triangulum* (wasp) and *Squilla mantis* (shrimp; ancestral pancrustacean) were used as outgroups. Adapted from Françoso et al. (2019).

Species	Length (bp)	A (%)	T (%)	AT (%)	C (%)	G (%)	CG (%)	AT rich region			GenBank access number
								Length	A (%)	T (%)	
Andrena bicolor	15,422	43.6	35.8	79.4	13.4	7.2	20.6	–	–	–	KT164666
Andrena camellia	15,065	46	32.6	78.6	14.7	6.7	21.4	396	48.2	31.1	KX241615
Apis andreniformis	17,529	41.7	44.2	85.9	7.5	6.6	14.1	1123	47.2	47.3	KF736157
Apis cerana	15,895	42.3	41.6	83.9	9.8	6.3	16.1	562	46.3	49.6	NC014295 (GQ162109)
Apis cerana japonica	15,917	42.3	41.7	84	9.7	6.3	16	568	46.6	50.5	AP017314
Apis florea	17,694	44.7	41.4	86.1	8.8	5.1	13.9	1987	56.6	37.1	JX982136 (NC_021401)
Apis koschevnikovi	16,050	42.5	42.1	84.6	9.3	6.1	15.4	952	47	48.2	KY348372
Apis mellifera capensis	16,470	43.3	41.5	84.8	9.6	5.6	15.2	881	51.4	44.7	KX870183
Bombus consobrinus	17,966	43.7	43	86.7	8.8	4.5	13.3	–	–	–	MF995069
Bombus ignitus	16,434	43.5	43.3	86.8	8.4	4.8	13.2	859	49	47	DQ870926 (NC010967)
Bombus lapidarius	17,817	44.1	41.8	85.9	9.7	4.4	14.1	–	–	–	KT164641
Colletes gigas	15,885	42.1	44.1	86.2	7.4	6.4	13.8	539	41.4	42.1	NC026218 (KM978210)
Frieseomelitta varia	15,144	39.6	48.2	87.8	5.3	6.9	12.2	198	49.5	48	WNWW01002174
Halictus tumulorum	15,268	40.4	41.1	81.5	12.2	6.3	18.5	–	–	–	KT164609
Hylaeus dilatatus	15,475	44.3	41.5	85.8	8.8	5.4	14.2	–	–	–	NC026468 (KP126800)
Lasioglossum lativentre	13,069	40.2	40.3	80.5	12.8	6.6	19.5	–	–	–	KT164682
Lepidotrigona flavibasis	15,408	38.9	39.4	78.3	10.7	11	21.7	344	47	38.8	MN747147
Megachile sculpturalis	16,581	42.8	40.6	83.4	10.7	5.9	16.6	1289	42.7	46	NC028017 (KT223644)
Melipona bicolor	14,422	44	42.7	86.7	8.3	5	13.3	255	42	56.1	NC004529
Melipona scutellaris	14,862	43.9	42.9	86.8	8.2	5	13.2	–	–	–	NC026198 (KP202303)
Nomada goodeniana	15,201	41	45.2	86.2	6.9	6.9	13.8	–	–	–	RSZAXPI001602–16
Rediviva intermixa	16,875	38.5	41.3	79.8	10.6	9.6	20.2	1186	27.8	26	NC030284 (KR864834)
Tetrapedia diversipes	15,358	42.6	41.7	84.3	9.8	5.9	15.7	588	43	43.5	NC060989
Tetragonula carbonaria	**30,666**	**35.9**	**36.5**	**72.3**	**13.9**	**13.8**	**27.7**	**AT1: 124**	**44.4**	**42.7**	OQ918628
								AT2: 112	**46.4**	**39.3**	
Tetragonula hockingsi	**30,662**	**36**	**36.4**	**72.4**	**14**	**13.6**	**27.6**	**AT1: 144**	**40.3**	**46.5**	OQ918629
								AT2: 70	**42.9**	**42.9**	
Philanthus triangulum	16,029	44.4	39.2	83.6	10.3	6.1	16.4	1039	39.5	46.2	NC017007 (JN871914)
Squilla mantis	15,994	35.1	35.1	70.2	16.8	13	29.8	862	42.2	34.5	AY639936

impairments of replication and transcription [36]. Although amphimeric genomes have been reported in some bacterial and yeast plasmids, and in chloroplasts of algae and higher plants, an amphimeric structure in mitochondrial genomes is extremely rare [36]. The duplication here reported could have arisen in the ancestral species of the Carbonaria species complex from the replication of a single mitogenome in which copies broke and fused at the control region, forming a final structure compound of two mirrored identical single mitogenomes (Fig. 4), except the $tRNA^{GLY}$ and the $tRNA^{GLU}$ (the "ambiguous tRNA"), that are present in single copies in mirrored positions, and the $tRNA^{SER1}$ and $tRNA^{LYS}$, that are present in single copies in the AT1 and AT2 regions, respectively (Figs. 2 and 3).

The symmetry in both copies of the duplicated mitogenome identified here, in which each mirrored mitogenome shares practically identical bases, is unexpected, particularly as it challenges the Tandem Duplication – Random Loss mechanism (TD-RL) [37]. In this model, gene duplication should enable loosening of selection on one of the copies, and consequently result in mutation accumulation and even random loss of one copy of each gene [37–39]. An example of this TD-RL mechanism was described in the parthenogenetic geckos from the *Heteronotia binoei* complex. These species have mitogenomes that range in size from 17 to 27 kb as a result of tandem replications (Moritz 1991, Zevering et al. 1991, Fujita et al. 2007). In one parthenogenic lineage (3N2), mitochondrial genes evolve as expected when they exist in duplicate: one gene copy remains active and functional, whereas deletions and base-pair mutations render the other copy a pseudogene [40–42]. In contrast to *Heteronotia* geckos, the mitogenome structure of *Tetragonula* suggests that an efficient repair system between duplicated copies must exist. A possible explanation for this gene conservation within the duplicated mitogenome is the homogenizing effect of gene conversion [42,43]. Gene conversion has been observed between duplicated control regions of the mitogenomes of several taxa, including birds [44], snakes [12], ticks [45], Australasian agamid lizards [46] and sea cucumbers [47]. In plants, gene conversion is commonly observed in their plastidial genomes. Nearly all land plants have a plastome composed of two copies of an inverted repeat (IR) region, and two different single-copy (SC) regions. The two copies of the IR region normally have low divergence when comparing to the two SC regions, suggesting that the duplicative nature of the IR reduces the substitution rate within this region [48]. Perry & Wolfe [49] observed that when this IR is lost, as in the IR-lacking clade of legumes, the synonymous substitution rate of the remaining IR copy increased to a value similar to that of other SC genes. These findings suggest that the reduced substitution rate is linked to the duplicative nature of the IR because of a copy-dependent repair mechanism. However, while gene conversion could explain the homogenization of genes within each duplicated mitogenome of *Tetragonula*, it itself does not explain the high rate of nucleotide substitutions in this genus. Nevertheless, as observed in *Tetragonula*,

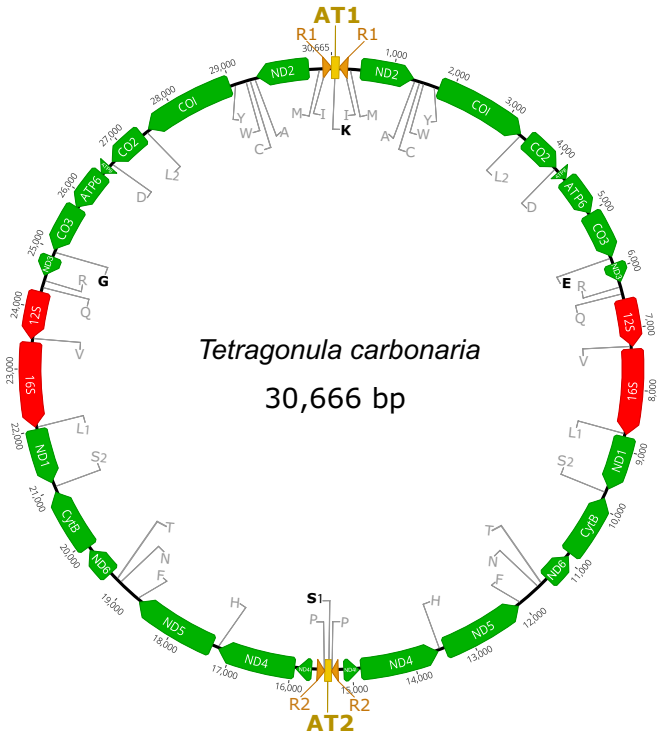

Fig. 2. The complete mitochondrial genome of *Tetragonula carbonaria,* comprised of two mirrored identical single mitogenomes (excepted tRNAs present in single copies, in bold), separated from each other by two distinct fragments of the original AT-rich region (AT1 and AT2) inside of two repeated regions (R1and R2). **A**: Alanine; **C**: Cysteine; **D**: Aspartic acid; **E**: Glutamic acid; **F**: Phenylalanine; **G**: Glycine; **H**: Histidine; **I**: Isoleucine; **K**: Lysine; **L**: Leucine; **M**: Methionine; **N**: Asparagine; **P**: Proline; **Q**: Glutamine; **R**: Arginine; **S**: Serine; **T**: Threonine; **V**: Valine; **W**: Tryptophan; **Y**: Tyrosine.

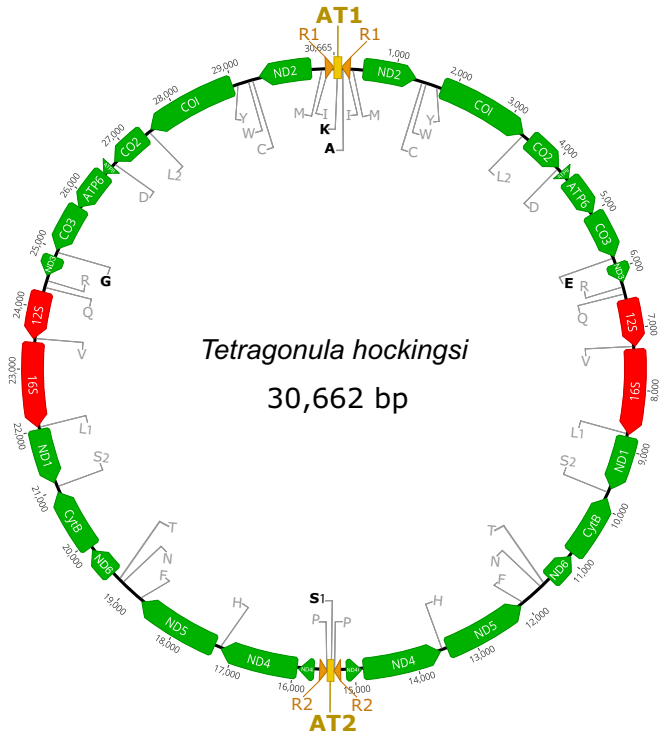

Fig. 3. - The complete mitochondrial genome of *Tetragonula hockingsi,* comprised of two mirrored identical single mitogenomes (excepted tRNAs present in single copies, in bold), separated from each other by two distinct fragments of the original AT-rich region (AT1 and AT2) inside of two repeated regions (R1and R2). **A**: Alanine; **C**: Cysteine; **D**: Aspartic acid; **E**: Glutamic acid; **F**: Phenylalanine; **G**: Glycine; **H**: Histidine; **I**: Isoleucine; **K**: Lysine; **L**: Leucine; **M**: Methionine; **N**: Asparagine; **P**: Proline; **Q**: Glutamine; **R**: Arginine; **S**: Serine; **T**: Threonine; **V**: Valine; **W**: Tryptophan; **Y**: Tyrosine.

some genera of plants such as *Pelargonium*, *Plantago* and *Silene* present highly accelerated synonymous rates despite their IR location, indicating hypermutation [42]. This hypermutation would be induced by a higher level of error-prone double-strand break repair in these regions, which generates substitutional rate variation [42] but still within the constraint of a strong purifying selection.

4.2. Mitogenome duplication and asymmetry in tRNA – a one-way ticket

The only difference found between the mirrored copies of the duplicated mitogenomes of *Tetragonula* outside the AT regions (where the break point occurred and may have resulted in loss of some tRNAs; Figs. 2 and 3) is the second base of the anticodon of a tRNA that encodes for glutamic acid in one side of the mitogenome, and glycine in the other. These ambiguous tRNA genes found in the *Tetragonula* species are presented side by side in both species of *Lepidotrigona* due to a translocation of $tRNA^{GLU}$ to beside the $tRNA^{GLY}$, between *CO3* and *ND3* genes (Fig. 1) [25]. This rearrangement in *Lepidotrigona* suggests, in this order: 1) a whole mitogenome duplication with both $tRNA^{GLY}$ and $tRNA^{GLU}$ in each side of the duplicated mitogenome; 2) a substitution of the second base of the $tRNA^{GLY}$ anticodon in one copy of this gene, transforming it in $tRNA^{GLU}$; and 3) loss of both copies of the original $tRNA^{GLU}$ in the duplicated mitogenome. Corroborating this hypothesis of a stochastic mutation shifting a $tRNA^{GLY}$ into a $tRNA^{GLU}$, the base composition of both tRNAs is identical, except for the mutated base. The sequence size of both tRNAs of *T. carbonaria* and *T. hockingsi* is 67 bp, with just one base difference. For comparisons *L. flavibasis* has 27 divergent sites between $tRNA^{GLU}$ and $tRNA^{GLY}$ in 71 bp; *M. bicolor* has 17 divergent sites in 70 bp; *B. ignitus* has 17 divergent sites in 77 bp; and *A. mellifera* has 16 divergent sites in 68 bp, showing that $tRNA^{GLU}$ and $tRNA^{GLY}$ have

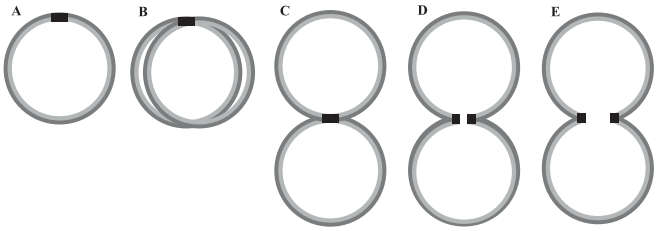

Fig. 4. Hypothesis about the origin of the whole genome duplication in *Tetragonula carbonaria* and *Tetragonula hockingsi*. A, B and C represent the normal stages of mitochondrial genome replication, followed by a break in the control region (D, in black), resulting in two repeated regions before and after each half of control region, and two internal regions between these repeated regions, rich in "A" and "T" (E).

significant divergences in these species, but the same origin in *Tetragonula*.

Although duplication of the whole circular mitogenome appears to be rare, this structure, including the polymorphism in tRNA, is not unique among the Bilateria [21,22]. In this respect, *Tetragonula* mitogenomes share some similarity with the mitogenomes of various lineages of terrestrial isopods (crustaceans of the suborder Oniscidea). In these species, the whole mitogenome is also duplicated, with two atypical conformations of mitogenomes described: one linear of ~14Kb containing the standard animal mitochondrial genes; and one circular of ~28Kb, consisting of two fused mitogenomes, arranged in a mirrored structure [21,22]. Here too, the presence of heteroplasmy in some tRNA loci encoding two alternative tRNAs was described in different oniscid lineages [50,51]. Like *Tetragonula*, all lineages present the same two

expected bases at very similar frequencies (~50 %) at the shared SNPs [51]. In this group, the idea that dimers arise from replication of the linear monomers was excluded, since the duplicated mitogenome expected would have totally identical monomers, and this is not the case with the ambiguous tRNA loci. Using long read sequencing technology, Peccoud et al. [51] concluded that most, if not all, dimers derive from the replication of the other circular dimers, and that monomers would be unable to replicate due to their hairpins. In these oniscids, the two mirrored tRNA genes are thought to have been identical in the ancestral dimeric genome, and to have become asymmetric over time due to an "anticodon shift" (from TGC to TAC via base substitution).

The ambiguous tRNA loci in both oniscids and *Tetragonula* is an example of balancing selection, in which essential heteroplasmy is maintained [50,51]; that is, the maintenance of duplicated mitogenomes would ensure the inheritance of all essential tRNAs [51]. The duplication and the asymmetry therefore represent a one-way ticket, because a regression to the original form of a single mitogenome would never be parsimoniously viable (although some gene loss could occur at some point through the TD-RL mechanism). The unusual duplicated structure of *Tetragonula* mitogenomes indeed may be considered adaptive, since it cannot be reversed.

Small size is a characteristic of mitogenomes that are under selection. Mitogenomes of animals have high rates of replication, and mtDNA transcription puts a limit on metabolic processes in tissues with high energy demands [18,52,53]. Consequently, an asymmetric duplicated and functional genome may provide efficiencies when compared to two single mitogenomes, as it can save energy during replication and transcription [51]. The results reported here provide a framework for testing the functional consequences of having a duplicated mitogenome.

4.3. Rearrangements in the mitogenome

Extreme rearrangement in the mitogenome is considered rare event, but when it happened, they can provide valuable information about mitogenome function and evolution. In both *T. carbonaria* and *T. hockingsi*, an entire block containing the genes *ND6*, *CytB*, *ND1*, *16S*, *12S* and some tRNAs has been translocated and inverted; and another entire block containing the genes *ND5*, *ND4*, *ND4L* and some tRNAs genes has been translocated, relative to the mitogenomes of most other bees (Fig. 1). The same rearrangement (except by some tRNA translocations) was described in the mitogenomes of the stingless bees *Lepidotrigona flavibasis* and *Lepidotrigona terminate* [25], but with no evidence of mitogenome duplication in those species (Fig. 1). Interestingly, the break points of these rearrangements coincide with the three polycistronic transcripts found in the mitogenome of the solitary bee *Tetrapedia diversipes* [54], indicating independence of each block in the transcriptional process, which may have allowed the rearrangements found in *Tetragonula* to occur without compromising mitochondrial activity.

In the stingless bee *F. varia*, a similar reshuffling in gene order was described in one block of genes. In *F. varia* the first break point was after the *ND3* gene, and the entire block with the genes *ND5*, *ND4* and *ND4L* (genes labelled green in the *Tetragonula* mitogenomes in Fig. 1) is inverted and translocated. Unlike *Tetragonula*, however, a second block with the genes *ND1*, *16S*, *12S* and the AT-rich region was also inverted and translocated in *F. varia* [24].

Rare gene rearrangements that are shared by different lineages that have a common ancestor are a valuable tool for phylogenetics studies [55–58]. *Tetragonula* and *Lepidotrigona* belong both to the Indo-Malay/Australasian group inside the Meliponini tribe [59], indicating that this extreme rearrangement in their mitogenomes may be present in the whole Indo-Malay/Australasian group. *Frieseomelitta* belongs to the Neotropical group of Meliponini, and the estimated divergence between this group and the Indo-Malay/Australasian is at least 70MYBP [59]. More studies on mitogenome evolution in meliponines from across their distribution are needed to understand if these rearrangements described in *Tetragonula*, *Lepidotrigona* and *Frieseomelitta* have a common ancestor or if they were independent events in common mitogenome regions.

4.4. Rapid evolution

The high mutation rate in mtDNA is unexpected because of the essential functions of mitochondrial genes [8]. Most mutations accumulating in mitogenomes are usually considered neutral (i.e. mostly synonymous), due to strong purifying selection [7,60–63]. However, it is now recognized that some mtDNA polymorphisms can contribute to key evolutionary processes such as phenotypic variation, adaptation to different climates, and speciation [3,7,60–69].

The mutation rate of mitogenomes ranges from 15 to 34 substitutions per site per billion years for all bilaterian animal groups, which is 9 to 25 times those for the nuclear genomes in the same lineages [4]. For invertebrates, this rate varies between 2 and 6 times [23,70], being 2.4 times faster in *Drosophila*, and 35 times faster in the parasitic wasp *Nasonia* – one of the highest substitution rates known [23]. In the Carbonaria complex, current evidence of an approximately 26 times higher mutation rate in mtDNA relative to nuclear genes [71] points to a rate of mitochondrial substitution that is much higher than that observed in nearly all other insects. Non-coding nuclear sequences flanking *EF-1α* (10 kb) and *Opsin* (26 kb) genes of *T. carbonaria* and *T. hockingsi* revealed very low nucleotide divergence (0.7–1 %) [71], indicating that rapid evolution in this group is a mitochondrial phenomenon. Most of the mutations are synonymous and biased to GC. Given the importance of mitochondrial genes to cell function, this suggests a very strong purifying selection "protecting" the integrity of mitochondrial genes. Mutation accumulation is not uniform across the protein coding genes of the mitogenome. In *Tetragonula*, the *ATP8* gene was the most divergent between species (29.2 %), followed by *ND3* and *ND6* (27.7 % and 27.2 %, respectively). These results are like those described for the parasitic wasp *Nasonia* [23], showing that these genes have the highest accelerated mutation rate among these mitogenomes.

Elevated sequence divergence, frequent rearrangement of mitochondrial genes, and fragmentation of the mitogenome was observed in a genus of booklice, *Liposcelis* [20], despite strong purifying selection. Comparing this genus to other Metazoan clades, Feng et al. [20] observed that species that are more closely related to species with fragmented mitogenomes also exhibit higher sequence divergence. The authors hypothesize that rapid mitogenome evolution might be explained by a faster mitogenome replication rate, whereby high replication rates can decrease the gene order stability, and sometimes sporadically lead to mitogenome fragmentation. In fact, given the long branches in the phylogeny in *Tetragonula* [30] and in *Lepidotrigona* [25], the whole Indo-Malay/Australasian group is likely to experience rapid evolution in their mitogenomes, being particularly elevated in the Australian species of the Carbonaria complex. The mitogenomes of *Lepidotrigona* also presented lower AT content when compared to bees mitogenomes in general (78.34 %) [25], indicating a GC-biased mutation.

Investigating variations in the mitogenome structure and gene order across Indo-Malay/Australasian group will shed further light on the relationship between atypical mitogenome structure and rapid rates of nucleotide substitution. Furthermore, continued work is needed to understand if the high mutation rate and, consequently, the high mitochondrial diversity in this group plays a role in their diversification and speciation [3,7,62].

4.5. Rapid evolution, rearrangements & duplicated mitogenomes in the Carbonaria complex

The mitogenome of *Lepidotrigona* is a key missing link that sheds light on the likely temporal order in which accelerated evolution, rearrangements and mitogenome duplication happened in the Carbonaria complex, since both rapid evolution and the same rare mitogenome

rearrangement pattern are observed in the *Lepidotrigona* mitogenome, but with no evidence of mitogenome duplication [25]. Thus the rapid evolution and the mitogenome rearrangement may be present in the whole Indo-Malay/Australasian group, though with the rapid evolution more intensified in the *Tetragonula* genus, especially in the Carbonaria complex. This exceptional higher rate of mitochondrial evolution in the Carbonaria complex could be explained by a) The founding effect: As discussed by Françoso et al. [30], the ancestor of this complex likely migrated from South-East Asia to Australia [26] and diverged into at least four endemic species (including *T. carbonaria* and *T. hockingsi*) [28]. Rapid evolution could be the result of a founder effect with a population bottleneck followed by expansions [30,72–75]; b) The low effective population size: *Tetragonula* are eusocial bees (in which just queens and males are reproductive) [76,77], and probably monandrous [78,79], which further decreases their effective population size. This very low effective population size could allow the fixation of a mitochondrial haplotype with suboptimal conditions in this new environment during the colonization of Australia. As a compensatory mechanism, the mitochondrial activity increased, intensifying rates of replication and transcription [23]. This increase of mitochondrial activity could lead to hypermutation driven by the release of free radicals as a subproduct of the electron transport chain. c) Mitogenome duplication: The duplicated mitogenome may elevate the rate of nucleotide substitutions that can be tolerated. First, mutations occur in one duplicated copy that are neutral or positive. Then, the repair mechanism, probably through recombination during the replication of the genome can homogenize the new mutations and fix them in both copies – similar to the homogenization mechanism that occurs in the duplicated regions of plant plastomes (gene conversion). This could explain the rapid evolution in the mitogenomes observed between species, while simultaneously explaining why both copies of the duplicated mitogenomes are (practically) identical. Still, the up regulation of mitochondrial activity caused by the duplicated genes would generate unusually high rates of respiration and consequently, increase even more the mutation rate and the rapid evolution.

5. Conclusion

The atypical mitogenomes of *T. carbonaria* and *T. hockingsi* described here are unique and are remarkable in their rapid evolution, rearrangements, and whole genome duplication. They highlight that mitogenomes can be evolutionarily dynamic, and that diversity in their structure and evolution needs to be studied further. Because the features of *Tetragonula* mitogenomes deviate from most mitogenomes described so far, they provide a model system to address questions of mitogenome function and evolution such as replication, transcription, regulation, recombination, repair mechanisms, purifying selection, balancing selection, concerted evolution, mitonuclear ecology, and speciation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Fellowship to EF (grant agreement No. 101023351), Fundação de Amparo `a Pesquisa do Estado de São Paulo (EF - Proc.14/25023-6; and research grant to MCA - 2016/24669-5), CNPq - Conselho Nacional de Desenvolvimento Científico e Tecnológico (research sponsorship to MCA - 306932/2016-4); and CAPES - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brasil (Finance Code 001).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijbiomac.2023.124568.

References

[1] L. Sagan, On the origin of mitosing cells, J. Theor. Biol. 14 (1967) 225–274.
[2] L.M. Van Valen, V.C. Maiorana, The Archaebacteria and eukaryotic origins, Nature 287 (1980) 248–250.
[3] G.E. Hill, Mitonuclear Ecology, Oxford University Press, Oxford, UK, 2019.
[4] M. Lynch, The origins of eukaryotic gene structure, Mol. Biol. Evol. 23 (2006) 450–468.
[5] D.R. Wolstenholme, Animal mitochondrial DNA: structure and evolution, in: D. R. Wolstenholme, K.W. Jeon (Eds.), Mitochondrial Genomes, Academic Press, San Diego, 1992, pp. 173–216.
[6] R. Zardoya, Recent advances in understanding mitochondrial genome diversity, F1000Res. 9 (2020) 270.
[7] J.N. Wolff, E.D. Ladoukakis, J.A. Enríquez, D.K. Dowling, Mitonuclear interactions: evolutionary consequences over multiple biological scales, Philos. Trans. R. Soc. B 369 (2014) 20130443.
[8] W.M. Brown, M. George Jr., A.C. Wilson, Rapid evolution of animal mitochondrial DNA, Proc. Natl. Acad. Sci. 76 (1979) 1967–1971.
[9] T.R. Singh, G. Tsagkogeorga, F. Delsuc, S. Blanquart, N. Shenkar, et al., Tunicate mitogenomics and phylogenetics: peculiarities of the Herdmania momus mitochondrial genome and support for the new chordate phylogeny, BMC Genomics 10 (2009) 534.
[10] C. Gissi, G. Pesole, F. Mastrototaro, F. Iannelli, V. Guida, et al., Hypervariability of ascidian mitochondrial gene order: exposing the myth of deuterostome organelle genome stability, Mol. Biol. Evol. 27 (2010) 211–215.
[11] D.P. Mindell, M.D. Sorenson, D.E. Dimcheff, Multiple independent origins of mitochondrial gene order in birds, Proc. Natl. Acad. Sci. 98 (1998) 10693–10697.
[12] Y. Kumazawa, H. Ota, M. Nishida, T. Ozawa, Gene rearrangements in snake mitochondrial genomes: highly concerted evolution of control-region-like sequences duplicated and inserted into a tRNA gene cluster, Mol. Biol. Evol. 13 (1996) 1242–1254.
[13] J.R. Macey, A. Larson, N.B. Ananjeva, Z. Fang, T.J. Papenfuss, Two novel gene orders and the role of light-strand replication in rearrangement of the vertebrate mitochondrial genome, Mol. Biol. Evol. 14 (1997) 91–104.
[14] J.S. Rest, J.C. Ast, C.C. Austin, P.J. Waddell, E.A. Tibbetts, J.M. Hay, D.P. Mindell, Molecular systematics of primary reptilian lineages and the tuatara mitochondrial genome, Mol. Phylogenet. Evol. 29 (2003) 289–296.
[15] Z.Q. Liu, Y.Q. Wang, B. Su, The mitochondrial genome organization of the rice frog, Fejervarya limnocharis (Amphibia: Anura): a new gene order in the vertebrate mtDNA, Gene 346 (2005) 145–151.
[16] R.L. Mueller, J.L. Boore, Molecular mechanism of extensive mitochondrial gene rearrangement in plethodontid salamanders, Mol. Biol. Evol. 22 (2005) 2104–2112.
[17] D. San Mauro, D.J. Gower, R. Zardoya, M. Wilkinson, A hotspot of gene order rearrangement by tandem duplication and random loss in the vertebrate mitochondrial genome, Mol. Biol. Evol. 23 (2006) 227–234.
[18] A. Sayadi, E. Immonen, C. Tellgren-Roth, G. Arnqvist, The evolution of dark matter in the mitogenome of seed beetles, Genome Biol. Evol. 9 (2017) 2697–2706.
[19] Y.G. Liu, T. Kurokawa, M. Sekino, T. Tanabe, K. Watanabe, Complete mitochondrial DNA sequence of the ark shell Scapharca broughtonii: an ultra-large metazoan mitochondrial genome, Comp. Biochem. Physiol. Part D Genomics Proteomics 8 (2013) 72–81.
[20] S. Feng, A. Pozzi, V. Stejskal, G. Opit, Q. Yang, R. Shao, D.K. Dowling, Z. Li, Fragmentation in mitochondrial genomes in relation to elevated sequence divergence and extreme rearrangements, BMC Biol. 20 (2022) 7.
[21] R. Raimond, I. Marcadé, D. Bouchon, T. Rigaud, J.P. Bossy, et al., Organization of the large mitochondrial genome in the isopod Armadillidium vulgare, Genetics 151 (1999) 203–210.
[22] I. Marcadé, R. Cordaux, V. Doublet, C. Debenest, D. Bouchonet, et al., Structure and evolution of the atypical mitochondrial genome of Armadillidium vulgare (Isopoda, crustacea), J. Mol. Evol. 65 (2007) 651–659.
[23] D.C.S.G. Oliveira, R. Raychoudhury, D.V. Lavrov, J.H. Werren, Rapidly evolving mitochondrial genome and directional selection in mitochondrial genes in the parasitic wasp Nasonia (Hymenoptera: Pteromalidae), Mol. Biol. Evol. 25 (2008) 2167–2180.
[24] F.C.P. Freitas, A.P. Lourenço, F.M.F. Nunes, A.R. Paschoal, F.C.P. Abreu, F. O. Barbin, et al., The nuclear and mitochondrial genomes of Frieseomelitta varia – a highly eusocial stingless bee (Meliponini) with a permanently sterile worker caste, BMC Genomics 21 (2020) 386.
[25] C.Y. Wang, M. Zhao, S.J. Wang, H.L. Xu, Y.M. Yang, L.N. Liu, Y. Feng, The complete mitochondrial genome of Lepidotrigona flavibasis (Hymenoptera: Meliponini) and high gene rearrangement in lepidotrigona mitogenomes, J. Insect Sci. 21 (2021) 1–7.
[26] P. Franck, E. Cameron, G. Good, J.Y. Rasplus, B.P. Oldroyd, Nest architecture and genetic differentiation in a species complex of Australian stingless bees, Mol. Ecol. 13 (2004) 2317–2331.
[27] T. Heard, The Australian Native Bee Book, Keeping Stingless Bee Hives for Pets, Pollination and Sugarbag Honey, Sugarbeg Bees, Brisbane, 2016.

[28] A.E. Dollin, L.J. Dollin, S.F. Sakagami, Australian stingless bees of the genus Trigona (Hymenoptera: Apidae), Invertebr. Taxon. 11 (1997) 861–896.
[29] R.M. Brito, T.M. Schaerf, M.R. Myerscough, T.A. Heard, B.P. Oldroyd, Brood comb construction by the sting-less bees Tetragonula hockingsi and Tetragonula carbonaria, Swarm Intelligence 6 (2012) 151–176.
[30] E. Françoso, A. Zuntini, P. Cseri de Oliveira, J.P.N. Silva, R. Brito, B. Oldroyd, M. C. Arias, Conserved numts mask a highly divergent mitochondrial-CO1 gene in a species complex of Australian stingless bee Tetragonula (Hymenoptera: Apidae), Mitochondrial DNA Part A 30 (2019) 806–817.
[31] E. Françoso, F. Gomes, M.C. Arias, A protocol for isolating insect mitochondrial genomes: a case study of NUMT in Melipona flavolineata (Hymenoptera: Apidae), Mitochondrial DNA 10 (2015) 1–4.
[32] M. Kearse, R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, A. Cooper, S. Markowitz, C. Duran, T. Thierer, B. Ashton, P. Mentjies, A. Drummond, Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data, Bioinformatics 28 (2012) 1647–1649.
[33] M. Bernt, A. Donath, F. Jühling, et al., MITOS: improved de novo metazoan mitochondrial genome annotation, Mol. Phylogenet. Evol. 69 (2013) 313–319.
[34] CE, Q. Yue, M. Akam, Mitochondrial genomes suggest that hexapods and crustaceans are mutually paraphyletic, Proc. R. Soc. B Biol. Sci. 272 (2005) 1295–1304.
[35] E. Françoso, M.C. Arias, Primers for corbiculate bees: DNA barcode and mini-barcode, Mol. Ecol. Resour. 13 (2013) 844–850.
[36] E. Rayko, Organization, generation and replication of amphimeric genomes: a review, Gene 199 (1997) 1–18.
[37] Z.N. Yu, Z.P. Wei, X.Y. Kong, W. Shi, Complete mitochondrial DNA sequence of oyster Crassostrea hongkongensis -a case of "Tandem duplication-random loss" for genome rearrangement in Crassostrea? BMC Genomics 9 (2008) 477.
[38] J. Ren, X. Liu, G. Zhang, B. Liu, X. Guo, Tandem duplication-random loss" is not a real feature of oyster mitochondrial genomes, BMC Genomics 10 (2009) 84.
[39] T. Kim, E. Kern, C. Park, S.A. Nadler, Y.J. Bae, J.K. Park, The bipartite mitochondrial genome of Ruizia karukerae (Rhigonematomorpha, Nematoda), Sci. Rep. 8 (2018) 7482.
[40] C. Moritz, Evolutionary dynamics of mitochondrial DNA duplications in parthenogenetic geckos, Heteronotia binoei, Genetics 129 (1991) 221–230.
[41] C.E. Zevering, C. Moritz, A. Heideman, R.A. Sturm, Parallel origins of duplications and the formation of pseudogenes in mitochondrial DNA from parthenogenetic lizards (Heteronotia binoei; Gekkonidae), J. Mol. Evol. 33 (1991) 431–441.
[42] M.K. Fujita, J.L. Boore, C. Moritz, Multiple origins and rapid evolution of duplicated mitochondrial genes in parthenogenetic geckos (Heteronotia binoei; Squamata, Gekkonidae), Mol. Biol. Evol. 12 (2007) 2775–2786.
[43] C.W. Birky, J. Walsh, Biased gene conversion, copy number, and apparent mutation rate differences within chloroplast and bacterial genomes, Genetics 130 (1992) 677–683.
[44] J.R. Eberhard, T.F. Wright, E. Bermingham, Duplication and concerted evolution of the mitochondrial control region in the parrot genus Amazona, Mol. Biol. Evol. 18 (2001) 1330–1342.
[45] W.C. Black, R.L. Roehrdanz, Mitochondrial gene order is not conserved in arthropods: prostriate and metastriate tick mitochondrial genomes, Mol. Biol. Evol. 15 (1998) 1772–1785.
[46] S.A.M. Amer, Y. Kumazawa, Mitochondrial DNA sequences of the Afro-Arabian spiny-tailed lizards (genus Uromastyx; family Agamidae): phylogenetic analyses and evolution of gene arrangements, Biol. J. Linn. Soc. 85 (2005) 247–260.
[47] A. Arndt, M.J. Smith, Mitochondrial gene rearrangement in the sea cucumber genus Cucumaria, Mol. Biol. Evol. 15 (1998) 1009–1016;
J.W. Ballard, M. Kreitman, Unraveling selection in the mitochondrial genome of Drosophila, Genetics 138 (1994) 757–772.
[48] A. Zhu, W. Guo, S. Gupta, W. Fan, J.P. Mower, Evolutionary dynamics of the plastid inverted repeat: the effects of expansion, contraction, and loss on substitution rates, New Phytol. 209 (2016) 1747–1756.
[49] A.S. Perry, K.H. Wolfe, Nucleotide substitution rates in legume chloroplast DNA depend on the presence of the inverted repeat, J. Mol. Evol. 55 (2002) 501–508.
[50] V. Doublet, C. Souty-Grosset, D. Bouchon, R. Cordaux, I. Marcadé, A thirty million year-old inherited heteroplasmy, PLoS One 3 (2008), e2938.
[51] J. Peccoud, M.A. Chebbi, A. Cormier, B. Moumen, C. Gilbert, I. Marcadé, C. Chandler, R. Cordaux, Untangling heteroplasmy, structure, and evolution of an atypical mitochondrial genome by PacBio sequencing, Genetics 207 (2017) 269–280.
[52] D.M. Rand, Endotherms, ectotherms, and mitochondrial genome-size variation, J. Mol. Evol. 37 (1993) 281–295.
[53] D.X. Zhang, G.M. Hewitt, Insect mitochondrial control region: a review of its structure, evolution and usefulness in evolutionary studies, Biochem. Syst. Ecol. 25 (1997) 99–120.
[54] E. Françoso, N.S. Araujo, P.C. Ricardo, P.K.F. Santos, A.R. Zuntini, M.C. Arias, Evolutionary perspectives on bee mtDNA from mito-OMICS analyses of a solitary species, Apidologie 51 (2020) 531–544.
[55] J.L. Boore, W.M. Brown, Big trees from little genomes: mitochondrial gene order as a phylogenetic tool, Curr. Opin. Genet. Dev. 8 (1998) 668–674.
[56] D. Silvestre, M. Dowton, M.C. Arias, The mitochondrial genome of the stingless bee Melipona bicolor (Hymenoptera, Apidae, Meliponini): sequence, gene organization and a unique tRNA translocation event conserved across the tribe Meliponini, Genet. Mol. Biol. 31 (2008) 451–460.
[57] L. Podsiadlowski, A. Braband, T.H. Struck, J. von Döhren, T. Bartolomaeus, Phylogeny and mitochondrial gene order variation in Lophotrochozoa in the light of new mitogenomic data from Nemertea, BMC Genomics 10 (2009) 364.
[58] A. Mwinyi, A. Meyer, C. Bleidorn, B. Lieb, T. Bartolomaeus, L. Podsiadlowski, Mitochondrial genome sequence and gene order of Sipunculus nudus give additional support for an inclusion of Sipuncula into Annelida, BMCGenomics 10 (2022) 27.
[59] C. Rasmussen, E.C. Cameron, Global stingless bee phylogeny supports ancient divergence, vicariance, and long distance dispersal, Biol. J. Linn. Soc. 99 (2010) 206–232.
[60] J.W. Ballard, M.C. Whitlock, The incomplete natural history of mitochondria, Mol. Ecol. 13 (2004) 729–744.
[61] D.M. Rand, The units of selection on mitochondrial DNA, Annu. Rev. Ecol. Syst. 32 (2001) 415–448.
[62] D.K. Dowling, U. Friberg, J. Lindell, Evolutionary implications of non-neutral mitochondrial genetic variation, Trends Ecol. Evol. 23 (2008) 546–554.
[63] M.F. Camus, J.N. Wolff, C.M. Sgrò, D.K. Dowling, Experimental support that natural selection has shaped the latitudinal distribution of mitochondrial haplotypes in Australian Drosophila melanogaster, Mol. Biol. Evol. 34 (2017) 2600–2612.
[64] P.U. Bier, F. Dufresne, R.S. Burton, Natural selection and the evolution of mtDNA-encoded peptides: evidence for intergenomic co-adaptation, Trends Genet. 17 (2001) 400–406.
[65] D. Mishmar, E. Ruiz-Pesini, P. Golik, V. Macaulay, A.G. Clark, S. Hosseini, M. Brandon, K. Easley, E. Chen, M.D. Brown, et al., Natural selection shaped regional mtDNA variation in humans, Proc. Natl. Acad. Sci. 100 (2003) 171–176.
[66] N. Galtier, B. Nabholz, S. Glémin, G.D.D. Hurst, Mitochondrial DNA as a marker of molecular diversity: a reappraisal, Mol. Ecol. 18 (2009) 4541–4550.
[67] R.S. Burton, R.J. Pereira, F.S. Barreto, Cytonuclear genomic interactions and hybrid breakdown, Annu. Rev. Ecol. Evol. Syst. 44 (2013) 281–302.
[68] R. Dobler, B. Rogell, F. Budar, D.K. Dowling, A meta-analysis of the strength and nature of cytoplasmic genetic effects, J. Evol. Biol. 27 (2014) 2021–2034.
[69] D.K. Dowling, Evolutionary perspectives on the links between mitochondrial genotype and disease phenotype, Biochim. Biophys. Acta 1840 (2014) 1393–1403.
[70] R. Allio, S. Donega, N. Galtier, B. Nabholz, Large variation in the ratio of mitochondrial to nuclear mutation rate across animals: implications for genetic diversity and the use of mitochondrial DNA as a molecular marker, Mol. Biol. Evol. 34 (2017) 2762–2772.
[71] P. Hereward, J.S. Tobias, R. Gloag, D.R. Brookes, Tests of Hybridisation in Tetragonula Stingless Bees Using Multiple Genetic Markers, 2020, https://doi.org/10.1101/2020.03.08.982546. BioRXIV.
[72] E. Mayr, Changes in genetic environment and evolution, in: J. Huxley, A.C. Hardy, E.B. Ford (Eds.), Evol as a Process, Allen and Unwin, London, 1954, pp. 157–180.
[73] A.R. Templeton, The theory of speciation via the founder principle, Genetics 94 (1980) 1011–1038.
[74] P.R. Grant, Evolution on Islands, Oxford University Press, Oxford, 1998.
[75] G. Velo-Antón, K.R. Zamudio, A. Cordero-Rivera, Genetic drift and rapid evolution of viviparity in insular fire salamanders (Salamandra salamandra), Heredity 108 (2012) 410–418.
[76] J. Romiguier, J. Lourenco, P. Gayral, N. Faivre, L.A. Weinert, S. Ravel, M. Ballenghien, V. Cahais, A. Bernard, E. Loire, L. Keller, N. Galtier, Population genomics of eusocial insects: the costs of a vertebrate-like effective population size, Evol. Biol. 27 (2014) 593–603.
[77] A, J. Romiguier, Relaxation of purifying selection suggests low effective population size in eusocial Hymenoptera and solitary pollinating bees, Peer Community J. 1 (2021) e2.
[78] K.A. Palmer, B.P. Oldroyd, J.J.G. Quezada-Euán, R.J. Paxton, W.J. May-Itza, Paternity frequency and maternity of males in some stingless bee species, Mol. Ecol. 11 (2002) 2107–2113.
[79] T.J. Smith, Evidence for male genitalia detachment and female mate choice in the Australian stingless bee Tetragonula carbonaria, Insect. Soc. 67 (2020) 189–193.