NRC Research Press

ARTICLE

First mitochondrial genomes of five hoverfly species of the genus *Eristalinus* (Diptera: Syrphidae)

Gontran Sonet, Yannick De Smet, Min Tang, Massimiliano Virgilio, Andrew Donovan Young, Jeffrey H. Skevington, Ximo Mengual, Thierry Backeljau, Shanlin Liu, Xin Zhou, Marc De Meyer, and Kurt Jordaens

Abstract: The hoverfly genus *Eristalinus* (Diptera, Syrphidae) contains many widespread pollinators. The majority of the species of *Eristalinus* occur in the Afrotropics and their molecular systematics still needs to be investigated. This study presents the first complete and annotated mitochondrial genomes for five species of *Eristalinus*. They were obtained by high-throughput sequencing of total genomic DNA. The total length of the mitogenomes varied between 15 757 and 16 245 base pairs. Gene composition, positions, and orientation were shared across species, and were identical to those observed for other Diptera. Phylogenetic analyses (maximum likelihood and Bayesian inference) based on the 13 protein coding and both rRNA genes suggested that the subgenus *Eristalinus* was paraphyletic with respect to the subgenus *Eristalodes*. An analysis of the phylogenetic informativeness of all protein coding and rRNA genes suggested that NADH dehydrogenase subunit 5 (nad5), cytochrome *c* oxidase subunit 1, nad4, nad2, cytochrome *b*, and 16S rRNA genes are the most promising mitochondrial molecular markers to result in supported phylogenetic hypotheses of the genus. In addition to the five complete mitogenomes currently available for hoverflies, the five mitogenomes published here will be useful for broader molecular phylogenetic analyses among hoverflies.

Key words: mitogenome, phylogeny, phylogenetic informativeness, flower fly.

Résumé : Les syrphes appartenant au genre *Eristalinus* (Diptera, Syrphidae) comptent plusieurs pollinisateurs répandus. La majorité des espèces d'*Eristalinus* sont retrouvées dans l'écozone afrotropicale et leur systématique moléculaire n'a pas été étudiée. Cette étude présente les premiers génomes mitochondriaux complets et annotés pour cinq espèces d'*Eristalinus*. Ils ont été obtenus par séquençage à haut débit d'ADN génomique total. La longueur totale des mitogénomes variait entre 15 757 et 16 245 pb. Les catalogues de gènes, leurs positions et orientation étaient identiques chez ces espèces et identiques à ce qui a été observé chez d'autres diptères. Des analyses phylogénétiques (vraisemblance maximale et inférence bayésienne) fondées sur 13 gènes codant pour des protéines et les deux gènes d'ARNr suggèrent que le sous-genre *Eristalinus* est paraphylétique par rapport au sous-genre *Eristalodes*. Une analyse de la valeur informative phylogénétique de tous les gènes codant pour des protéines ou des ARNr suggère que les gènes codant pour la sous-unité 5 de la NADH déshydrogénase (nad5), la sous-unité 1 de la cytochrome *c* oxydase, nad4, nad2, le cytochrome *b* et les ARNr 16S seraient les marqueurs moleculaires mito-

Received 18 January 2019. Accepted 8 June 2019.

Corresponding Editor: T. Crease.

G. Sonet. Royal Belgian Institute of Natural Sciences, JEMU and BopCo, Vautierstraat 29, B-1000 Brussels, Belgium.
Y. De Smet, M. Virgilio, and M. De Meyer. Royal Museum for Central Africa, JEMU and BopCo, Leuvensesteenweg 13, B-3080 Tervuren, Belgium.
M. Tang and X. Zhou. Department of Entomology, China Agricultural University, Beijing 100193, China.
A.D. Young. Plant Pest Diagnostics Center, California Department of Food and Agriculture, 3294 Meadowview Road, Sacramento, CA 95832, USA; Department of Entomology and Nematology, University of California, Davis, Briggs Hall, Davis, CA 95616-5270, USA.
J.H. Skevington. Canadian National Collection of Insects, Arachnids and Nematodes, Agriculture and Agri-Food Canada, K.W. Neatby Building, 960 Carling Avenue, Ottawa, ON K1A 0C6, Canada; Department of Biology, Carleton University, 1125 Colonel By Drive, Ottawa, ON K1S 5B6, Canada.
X. Mengual. Zoologisches Forschungsmuseum Alexander Koenig, Leibniz-Institut für Biodiversität der Tiere, Adenauerallee 160, D-53113 Bonn, Germany.
T. Backeljau. Royal Belgian Institute of Natural Sciences, JEMU and BopCo, Vautierstraat 29, B-1000 Brussels, Belgium; University of Antwerp, Evolutionary Ecology Group, Universiteitsplein 1, B-2610 Antwerp, Belgium.
S. Liu. China National GeneBank, BGI-Shenzhen, Shenzhen, Guangdong 518083, China.
K. Jordaens. Royal Museum for Central Africa, JEMU and BopCo, Leuvensesteenweg 13, B-3080 Tervuren, Belgium; University of Antwerp, Evolutionary Ecology Group, Universiteitsplein 1, B-2610 Antwerp, Belgium.

Corresponding author: Gontran Sonet (email: gsonet@naturalsciences.be).

chondriaux les plus prometteurs pour obtenir des hypothèses phylogénétiques supportées pour ce genre. Avec les cinq mitogénomes complets déjà disponibles pour les syrphes, les cinq mitogénomes rapportés ici seront utiles pour des analyses phylogénétiques moléculaires plus larges au sein des syrphes. [Traduit par la Rédaction]

Mots-clés : mitogénome, phylogénie, valeur informative en phylogénie, mouche de fleurs.

Introduction

Four subfamilies are recognized in the family Syrphidae (Insecta: Diptera), also known as hoverflies or flower flies: Microdontinae, Syrphinae, Eristalinae, and Pipizinae (Mengual et al. 2015; Young et al. 2016). Yet, the subfamily Eristalinae appears paraphyletic in relation with the three other subfamilies (Hippa and Ståhls 2005; Skevington and Yeates 2000; Ståhls et al. 2003). A widespread genus within the Eristalinae is *Eristalinus* Rondani, 1845. It is naturally present in all biogeographic regions except in the Neotropics, although a few records indicate that it was introduced in Chile (Thompson 1999). Species of this genus are large hoverflies (4–16 mm), and most of them are imperfect bee mimics with punctate (spotted) and (or) fasciate (striped) eyes. They occur in a wide variety of habitats including open grasslands, shrub lands, river valleys, forest margins, wetlands, river banks, lake shores, and even urban areas. Larvae can be found in small temporary waterbodies. The genus comprises approximately 75 species worldwide, of which 54 occur in the Afrotropical Region. It is divided into five subgenera: *Eristalinus* Rondani, 1845 (including *Lathyrophthalmus* Mik, 1897), *Eristalodes* Mik, 1897, *Helophilina* Becker, 1923, *Merodonoides* Curran, 1931, and *Oreristalis* Séguy, 1951. So far, the phylogenetic relationships among the species of *Eristalinus* of the Afrotropical Region remain unknown. A preliminary phylogenetic study on species of *Eristalinus* from the Afrotropics using three mitochondrial (cytochrome *c* oxidase subunit 1, cytochrome *b*, and 12S rRNA) and two nuclear (18S and 28S rRNA) gene fragments showed low support for many nodes in the phylogenetic trees (De Smet et al. unpublished results) prompting the usage of additional DNA markers to resolve species and subgenus relationships within *Eristalinus*. Phylogenetic studies using whole mitochondrial genomes (mitogenomes) have shown the potential to tackle phylogenetic issues at varying taxonomic levels (e.g., Cameron 2013; Cameron et al. 2007, 2009; Ma et al. 2012; Nelson et al. 2012; Yong et al. 2015). The objectives of the current study are to assemble the mitogenomes of five Afrotropical species of *Eristalinus* (belonging to two subgenera), infer their phylogenetic relationships, and measure the informativeness of each mitochondrial protein coding gene (PCG) and rRNA gene for the resolution of phylogenetic relationships within *Eristalinus*.

Materials and methods

Total genomic DNA was extracted from five Afrotropical specimens of *Eristalinus* of the subgenera *Eristalinus* and *Eristalodes*, viz. *Eristalinus* (*Eristalinus*) *aeneus* (Scopoli, 1763), *Eristalinus* (*Eristalinus*) *tabanoides* (Jaennicke, 1867), *Eristalinus* (*Eristalinus*) *vicarians* (Bezzi, 1915), *Eristalinus* (*Eristalodes*) *barclayi* (Bezzi, 1915), and *Eristalinus* (*Eristalodes*) *fuscicornis* (Karsch, 1887) (Table 1), using the DNeasy® Blood &Tissue kit (Qiagen Inc., Hilden, Germany). Specimens were collected between October 2012 and May 2014 and stored in absolute ethanol. To minimize contamination with exogenous DNA, three legs were extracted from each specimen, rinsed in absolute ethanol, and air dried for 10 min at 50 °C before DNA extraction. DNA concentrations of the extracts were measured with Qubit 2.0 (ThermoFisher Scientific, Waltham, Massachusetts, USA) and ranged between 13 and 60 ng/μL.

One DNA library was prepared following the Illumina® TruSeq® DNA Sample Preparation Kit. An insert size of 250 base pairs (bp) was targeted after pooling 100 ng of the genomic DNA of each specimen. The DNA library was sequenced (150 bp paired-end reads) on an Illumina (San Diego, California, USA) HiSeq4000 platform at BGI-Shenzhen (China). The raw data were cleaned using a custom Perl script (Zhou et al. 2013) to remove Illumina adapters, reads with >10% of low quality bases (Phred-scores <20), or with more than five unresolved bases. The remaining reads were processed in the mitogenome assembly pipeline described by Tang et al. (2015). To achieve the final mitogenome assembly, the source code of the assembler SOAPdenovo-Trans(-K 71, -t 1) (Xie et al. 2014) was modified to remove scaffold connections with read supports ≤10. Mitogenomes were circularized and visualized in Geneious 10.2.2 (Biomatters, Auckland, New Zealand). A first draft annotation of the mitogenomes was obtained in MITOS (Bernt et al. 2013) with default settings. Then, open reading frames (ORF) of the mitochondrial coding genes were inferred using the transfer RNA (tRNA) punctuation principle as proposed by Cameron (2014). For this, annotations obtained from MITOS were manually edited in Geneious 10.2.2. Largest ORF found between tRNAs were selected, allowing truncated (incomplete) stop codons (which are completed by RNA polyadenylation) and overlap between adjacent ORF. Limits of the currently adopted routines for mitogenomic assemblage and annotation are presented by Velozo Timbó et al. (2017). The cloverleaf structure of the 22 inferred tRNAs were visualized in MITOS. The assembled and annotated mitogenomes were submitted to GenBank (accession numbers from MH321204 to MH321208). Gene symbols and their concordance with the nomenclature of FlyBase (Gramates et al. 2017) are given in Table 2. All mitogenomes were aligned in MUSCLE (Edgar 2004).

Using Geneious 10.2.2, uncorrected pairwise p-distances (proportion of nucleotide sites at which two sequences are

Table 1. Voucher numbers and collection information of the five Afrotropical species of *Eristalinus* (Diptera, Syrphidae) used in this study.

Species	Voucher no.	Sex	Country	Region	Collection date
Subgenus *Eristalinus*					
Eristalinus aeneus	RMCA ENT000033538	Female	Ethiopia	Holeta	October 2012
Eristalinus tabanoides	RMCA ENT000033541	Male	Benin	Calavi	April 2014
Eristalinus vicarians	RMCA ENT000033542	Female	Benin	Calavi	April 2014
Subgenus *Eristalodes*					
Eristalinus barclayi	RMCA ENT000033539	Male	Benin	Calavi	April 2014
Eristalinus fuscicornis	RMCA ENT000033540	Female	Benin	Cotonou	November 2013

different) were calculated among the mitogenomes of the five specimens of *Eristalinus* sequenced here and the five mitochondrial genomes of Syrphidae available in GenBank: *Episyrphus balteatus* (de Geer, 1776) (GenBank accession number: NC_036481), *Eristalis tenax* (Linnaeus, 1758) (MH159199), *Eupeodes corollae* (Fabricius, 1794) (NC_036482), *Ocyptamus sativus* (Curran, 1941) (KT272862), and *Simosyrphus grandicornis* (Macquart, 1842) (NC_008754). Graphs representing GC content and the similarity among the mitogenomes sequenced here were made in Geneious 10.2.2 using sliding windows of 99 and 50 bp, respectively.

Phylogenetic analyses were performed using both maximum likelihood (ML) and Bayesian inference (BI). For this, all PCGs and rRNA genes of the five mitogenomes obtained here and the five mitogenomes of Syrphidae available in GenBank (see above) were concatenated and aligned using the default parameters of ClustalW (Thompson et al. 1994). The dataset was then partitioned in 41 character sets corresponding to the different codon positions of the 13 PCGs (39) and the two rRNA genes. Additional ML and BI analyses were performed on the basis of cox1 and including additional samples of *Eristalinus* sequenced by Pérez-Bañón et al. (2003). This latter dataset was partitioned according to codon position. For each partition of all analyses, the General Time Reversible (GTR) model was applied, and a gamma distribution was used to approximate the heterogeneity of substitution rates among different sites. In all analyses (ML and BI), members of the subfamily Eristalinae represented the ingroup and members of the subfamily Syrphinae were used as outgroup to root the tree. ML analyses and BI were performed on the CIPRES Science Gateway (Miller et al. 2010) using RAxML v. 8 (Stamatakis 2014) and MrBayes 3.2.6 (Ronquist and Huelsenbeck 2003), respectively. ML analyses were implemented with autoMRE bootstrapping (setting a maximum of 1000 bootstraps) (Stamatakis 2014). The Bayesian analyses consisted in two runs, each with a cold chain and three incrementally heated chains. Starting trees for each chain were random and the default values of MrBayes were chosen for all settings (including prior distributions). MrBayes metropolis coupled Markov Chains Monte Carlo (MCMC) were run for 20 million generations with heating temperature of 0.1. Trees were sampled every 1000 generations with 50% of trees discarded as burn-in. Run convergence was verified by considering the average standard deviation of split frequencies (Ronquist and Huelsenbeck 2003). Only nodes with posterior probabilities >0.95 (BI) and bootstrap support >70% (ML) were considered.

Phylogenetic Informativeness (PI) is a quantitative measure of the ability of characters to resolve phylogenetic relationships among taxa over a specified historical time scale (Townsend 2007). To identify the best mitochondrial markers for the resolution of phylogenetic hypotheses within the genus *Eristalinus*, we estimated the net PI of each PCG and rRNA gene. In contrast with the PI per site, the net PI summarizes the PI for a set of characters and quantifies signal as a whole. The PI of single genes can be evaluated using, as prior information, more complete or comparative genomic data from within or outside the taxonomic group of interest (López-Giráldez and Townsend 2011). Here, we considered the concatenation of all PCGs and rRNA genes as prior information. The analysis was performed by applying the HyPhy algorithm (http://phydesign.townsend.yale.edu) (López-Giráldez and Townsend 2011) to the PCGs and rRNA genes of the five species of *Eristalinus* and the five mitochondrial genomes of Syrphidae available in GenBank (see above). The reference ultrametric tree required by the algorithm as prior information was reconstructed as a strict molecular clock tree using MrBayes v. 3.2.6 (Ronquist and Huelsenbeck 2003) available on the CIPRES Science Gateway (Miller et al. 2010). Convergence was checked as explained above.

Results

An average of 11.2 million reads were obtained per sample (SD = 0.8 × 10^6) with high proportions of reads with a Phred score ≥20 (96.6% ± 0.3% for read 1; 90.6% ± 1.0% for read 2). The aligned mitogenomes (File S1[1]) ranged from 15 757 (*E. barclayi*; N.B. abbreviation "*E.*" is only used for the genus *Eristalinus*) to 16 245 bp (*E. aeneus*), with most of the interspecific length variation situated in the AT-rich control region (Tables 2 and 3), which, in

[1]Supplementary data are available with the article through the journal Web site at http://nrcresearchpress.com/doi/suppl/10.1139/gen-2019-0009.

Table 2. Mitogenome annotations for the five Afrotropical species of *Eristalinus* sequenced here.

(A)		*Eristalinus aeneus*						*Eristalinus barclayi*						*Eristalinus fuscicornis*					
Gene	FlyBase Symbol*	Start position	End position	Size (bp)	Strand†	Intergenic‡	Start/stop codon	Start position	End position	Size (bp)	Strand†	Intergenic‡	Start/stop codon	Start position	End position	Size (bp)	Strand†	Intergenic‡	Start/stop codon
trnI(gat)	tRNA:Ile-GAT	1	66	66	H	—	—	1	66	66	H	—	—	1	66	66	H	—	—
trnQ(ttg)	tRNA:Gln-TTG	70	138	69	L	3	—	64	132	69	L	−3	—	64	132	69	L	−3	—
trnM(cat)	tRNA:Met-CAT	151	219	69	H	13	—	153	221	69	H	20	—	151	219	69	H	18	—
nad2	ND2	220	1240	1021	H	0	ATC/T(AA)	222	1242	1021	H	0	ATC/T(AA)	220	1240	1021	H	0	ATC/T(AA)
trnW(tca)	tRNA:Trp-TCA	1241	1308	68	H	0	—	1243	1310	68	H	0	—	1241	1308	68	H	0	—
trnC(gca)	tRNA:Cys-GCA	1301	1369	69	L	−8	—	1303	1371	69	L	−8	—	1301	1369	69	L	−8	—
trnY(gta)	tRNA:Tyr-GTA	1399	1465	67	L	29	—	1388	1454	67	L	16	—	1386	1452	67	L	16	—
cox1	CoI	1464	2997	1534	H	−2	TCG/T(AA)	1453	2986	1534	H	−2	TCG/T(AA)	1451	2984	1534	H	−2	TCG/T(AA)
trnL2(taa)	tRNA:Leu-TAA	2998	3063	66	H	0	—	2987	3052	66	H	0	—	2985	3050	66	H	0	—
cox2	CoII	3067	3750	684	H	3	ATG/TAA	3059	3742	684	H	6	ATG/TAA	3057	3740	684	H	6	ATG/TAA
trnK(ctt)	tRNA:Lys-CTT	3754	3811	58	H	3	—	3744	3814	71	H	1	—	3742	3812	71	H	1	—
trnD(gtc)	tRNA:Asp-GTC	3874	3940	67	H	>62	—	3820	3886	67	H	5	—	3826	3892	67	H	13	—
atp8	ATPase8	3941	4102	162	H	0	ATT/TAA	3887	4048	162	H	0	ATT/TAA	3893	4054	162	H	0	ATT/TAA
atp6	ATPase6	4096	4773	678	H	−7	ATG/TAA	4042	4719	678	H	−7	ATG/TAA	4048	4725	678	H	−7	ATG/TAA
cox3	CoIII	4779	5567	789	H	5	ATG/TAA	4724	5512	789	H	4	ATG/TAA	4730	5518	789	H	4	ATG/TAA
trnG(tcc)	tRNA:Gly-TCC	5572	5637	66	H	4	—	5516	5582	67	H	3	—	5522	5588	67	H	3	—
nad3	ND3	5638	5989	352	H	0	ATT/T(AA)	5583	5936	354	H	0	ATT/TAG	5589	5942	354	H	0	ATT/TAG
trnA(tgc)	tRNA:Ala-TGC	5990	6057	69	H	0	—	5938	6006	69	H	1	—	5944	6012	69	H	1	—
trnR(tcg)	tRNA:Arg-TCG	6058	6121	64	H	0	—	6007	6070	64	H	0	—	6013	6076	64	H	0	—
trnN(gtt)	tRNA:Asn-GTT	6136	6203	68	H	14	—	6076	6142	67	H	5	—	6082	6148	67	H	5	—
trnS1(gct)	tRNA:Ser-GCT	6204	6270	67	H	0	—	6143	6209	67	H	0	—	6149	6215	67	H	0	—
trnE(ttc)	tRNA:Glu-TTC	6273	6339	67	H	2	—	6210	6277	68	H	0	—	6216	6283	68	H	0	—
trnF(gaa)	tRNA:Phe-GAA	6368	6435	68	L	28	—	6317	6383	67	L	39	—	6324	6390	67	L	40	—
nad5	ND5	6436	8167	1732	L	0	GTG/T(AA)	6384	8115	1732	L	0	GTG/T(AA)	6391	8122	1732	L	0	GTG/T(AA)
trnH(gtg)	tRNA:His-GTG	8168	8233	66	L	0	—	8116	8181	66	L	0	—	8123	8188	66	L	0	—
nad4	ND4	8236	9576	1341	L	2	ATG/TAA	8182	9521	1340	L	0	ATG/TA(A)	8189	9528	1340	L	0	ATG/TA(A)
nad4l	ND4L	9570	9866	297	L	−7	ATG/TAA	9515	9811	297	L	−7	ATG/TAA	9522	9818	297	L	−7	ATG/TAA
trnT(tgt)	tRNA:Thr-TGT	9869	9934	66	H	2	—	9814	9879	66	H	2	—	9821	9886	66	H	2	—
trnP(tgg)	tRNA:Pro-TGG	9935	10 001	67	L	0	—	9880	9946	67	L	0	—	9887	9953	67	L	0	—
nad6	ND6	10 004	10 525	522	H	2	ATT/TAA	9949	10 473	525	H	2	ATC/TA(A)	9956	10 480	525	H	2	ATT/TA(A)
cob	Cyt-b	10 533	11 669	1137	H	7	ATG/TAA	10 473	11 609	1137	H	−1	ATG/TAA	10 480	11 616	1137	H	−1	ATG/TAA
trnS2(tga)	tRNA:Ser-TGA	11 676	11 743	68	H	6	—	11 622	11 689	68	H	12	—	11 629	11 696	68	H	12	—
nad1	ND1	11 765	12 706	942	L	25	TTG/TAA	11 711	12 652	942	L	21	TTG/TAA	11 718	12 659	942	L	21	TTG/TAA
trnL1(tag)	tRNA:Leu-TAG	12 708	12 772	65	L	1	—	12 654	12 718	65	L	1	—	12 661	12 725	65	L	1	—
rrnL(=16S)	lrRNA	12 773	14 120	1348	L	0	—	12 719	14 056	1338	L	0	—	12 726	14 063	1338	L	0	—
trnV(tac)	tRNA:Val-TAC	14 121	14 192	72	L	0	—	14 057	14 128	72	L	0	—	14 064	14 135	72	L	0	—
rrnS(=12S)	srRNA	14 193	14 982	790	L	0	—	14 129	14 918	790	L	0	—	14 136	14 925	790	L	0	—
control region	ori	14 983	16 245	1263	H	0	—	14 919	15 757	839	H	0	—	14 926	15 815	890	H	0	—

Table 2 (*concluded*).

(B)		*Eristalinus tabanoides*						*Eristalinus vicarians*					
Gene	FlyBase Symbol*	Start position	End position	Size (bp)	Strand†	Intergenic‡	Start/stop codon	Start position	End position	Size (bp)	Strand†	Intergenic‡	Start/stop codon
trnI(gat)	tRNA:Ile-GAT	1	66	66	H	—	—	1	66	66	H	—	—
trnQ(ttg)	tRNA:Gln-TTG	67	135	69	L	0	—	64	132	69	L	–3	—
trnM(cat)	tRNA:Met-CAT	174	242	69	H	38	—	156	224	69	H	23	—
nad2	ND2	243	1266	1024	H	0	ATC/T(AA)	225	1245	1021	H	0	ATC/T(AA)
trnW(tca)	tRNA:Trp-TCA	1267	1334	68	H	0	—	1246	1313	68	H	0	—
trnC(gca)	tRNA:Cys-GCA	1327	1395	69	L	–8	—	1306	1371	66	L	–8	—
trnY(gta)	tRNA:Tyr-GTA	1446	1512	67	L	50	—	1382	1448	67	L	10	—
cox1	CoI	1511	3044	1534	H	–2	TCG/T(AA)	1447	2980	1534	H	–2	TCG/T(AA)
trnL2(taa)	tRNA:Leu-TAA	3045	3110	66	H	0	—	2981	3046	66	H	0	—
cox2	CoII	3117	3800	684	H	6	ATG/TAA	3055	3738	684	H	8	ATG/TAA
trnK(ctt)	tRNA:Lys-CTT	3802	3872	71	H	1	—	3740	3810	71	H	1	—
trnD(gtc)	tRNA:Asp-GTC	3943	4009	67	H	70	—	3881	3947	67	H	70	—
atp8	ATPase8	4010	4171	162	H	0	ATT/TAA	3948	4109	162	H	0	ATT/TAA
atp6	ATPase6	4165	4842	678	H	–7	ATG/TAA	4103	4780	678	H	–7	ATG/TAA
cox3	CoIII	4849	5637	789	H	6	ATG/TAA	4784	5572	789	H	3	ATG/TAA
trnG(tcc)	tRNA:Gly-TCC	5641	5707	67	H	3	—	5576	5642	67	H	3	—
nad3	ND3	5708	6059	352	H	0	ATT/T(AA)	5643	5994	352	H	0	ATT/T(AA)
trnA(tgc)	tRNA:Ala-TGC	6060	6128	69	H	0	—	5995	6063	69	H	0	—
trnR(tcg)	tRNA:Arg-TCG	6129	6191	63	H	0	—	6064	6127	64	H	0	—
trnN(gtt)	tRNA:Asn-GTT	6201	6266	66	H	9	—	6152	6218	67	H	24	—
trnS1(gct)	tRNA:Ser-GCT	6267	6333	67	H	0	—	6219	6285	67	H	0	—
trnE(ttc)	tRNA:Glu-TTC	6334	6401	68	H	0	—	6286	6353	68	H	0	—
trnF(gaa)	tRNA:Phe-GAA	6458	6524	67	L	56	—	6382	6448	67	L	28	—
nad5	ND5	6525	8256	1732	L	0	GTG/T(AA)	6449	8180	1732	L	0	GTG/T(AA)
trnH(gtg)	tRNA:His-GTG	8257	8322	66	L	0	—	8181	8246	66	L	0	—
nad4	ND4	8323	9662	1340	L	0	ATG/TA(A)	8247	9586	1340	L	0	ATG/TA(A)
nad4l	ND4L	9656	9952	297	L	–7	ATG/TAA	9580	9876	297	L	–7	ATG/TAA
trnT(tgt)	tRNA:Thr-TGT	9955	10 020	66	H	2	—	9879	9944	66	H	2	—
trnP(tgg)	tRNA:Pro-TGG	10 021	10 087	67	L	0	—	9945	10 011	67	L	0	—
nad6	ND6	10 090	10 614	525	H	2	ATC/TA(A)	10 014	10 538	525	H	0	ATT/TA(A)
cob	Cyt-b	10 614	11 750	1137	H	–1	ATG/TAA	10 538	11 674	1137	H	–1	ATG/TAA
trnS2(tga)	tRNA:Ser-TGA	11 761	11 828	68	H	10	—	11 693	11 760	68	H	18	—
nad1	ND1	11 850	12 791	942	L	21	TTG/TAG	11 782	12 723	942	L	21	TTG/TAG
trnL1(tag)	tRNA:Leu-TAG	12 793	12 857	65	L	1	—	12 725	12 789	65	L	1	—
rrnL(=16S)	lrRNA	12 858	14 195	1338	L	0	—	12 790	14 129	1340	L	0	—
trnV(tac)	tRNA:Val-TAC	14 196	14 267	72	L	0	—	14 130	14 201	72	L	0	—
rrnS(=12S)	srRNA	14 268	15 057	790	L	0	—	14 202	14 990	789	L	0	—
control region	ori	15 058	15 792	735	H	0	—	14 991	15 966	976	H	0	—

*Gene abbreviations: ATP, ATP synthase membrane subunits; Co, cytochrome oxidase subunits; Cytb, cytochrome *b*; ND, NADH dehydrogenase subunits; rRNA, ribosomal RNA gene (12S and 16S); tRNA, transfer RNA gene.

†Strand: H, H-strand; L, L-strand.

‡Intergenic: intergenic region where negative values indicate an overlap between the genes.

Table 3. Genome size (base pairs), base composition (%), and GC content (%) of the syrphid mitogenomes sequenced here and retrieved from GenBank.

Species	Genome size (base pairs)	A (%)	T (%)	G (%)	C (%)	GC content (%)	Reference
Subfamily Eristalinae							
Eristalinus tabanoides	15 792	41.2	38.8	8.3	11.7	20.0	This study
Eristalinus vicarians	15 966	41.1	38.9	8.2	11.8	20.0	This study
Eristalinus aeneus	16 245	47.8	32.0	7.1	13.1	20.2	This study
Eristalinus fuscicornis	15 815	41.0	38.9	8.4	11.7	20.1	This study
Eristalinus barclayi	15 757	40.9	39.0	8.4	11.8	20.2	This study
Eristalis tenax	16 091	40.0	40.1	8.7	11.2	19.9	Li et al. (2017)
Subfamily Syrphinae							
Episyrphus balteatus	16 175	39.5	40.2	8.9	11.4	20.4	Pu et al. (2017)
Eupeodes corollae	15 326	40.5	39.7	8.6	11.2	19.8	Pu et al. (2017)
Ocyptamus sativus	15 214	40.2	40.2	8.5	11.1	19.6	Junqueira et al. (2016)
Simosyrphus grandicornis	16 141	40.3	40.6	8.3	10.9	19.2	Junqueira et al. (2016)

all species, was between the trnS and trnI genes (conversions of gene name abbreviations in Table 2). Base composition of mitogenomes was heavily skewed towards A+T, with a GC content of 20%–20.2% (Table 3). Every mitogenome contained the same 37 genes, ordered identically and comprising 13 PCGs, 2 rRNA, and 22 tRNA genes. The number of intergenic regions, and their cumulative length, was variable among species (Table 2).

Nine of the 13 PCGs were situated on the H-strand (represented clockwise in Fig. 1) and four on the L-strand (represented counter clockwise in Fig. 1). Lengths of coding regions and start/stop codons are given in Table 2. Each PCG had the same start and stop codons in the five species with the following exceptions: ATT instead of ATC as start codon for nad6 in *E. aeneus* and TAA stop codon in *E. aeneus*, *E. barclayi*, and *E. fuscicornis* instead of TAG in *E. tabanoides* and *E. vicarians* for nad1. The most common start codon was ATG (observed in 46% of the PCG/species combinations), followed by ATT (20%) and ATC (10%). Other start codons, TCG, TTG, and GTG represented 8% each. The most common stop codon was TAA (54% of the PCG/species combinations), while TAG was observed in 6% of the cases. An incomplete stop codon T(AA) was observed in three PCGs (nad2, cox1, and nad5) of all five species, with an additional T(AA) stop codon in nad3 of *E. aeneus*, *E. tabanoides*, and *E. vicarians*. Another incomplete stop codon TA(A) was observed in two PCGs (nad4 and nad6) of four species (all except *E. aeneus*). These incomplete stop codons represented thus 28% [T(AA)] and 12% [TA(A)] of PCG/species combinations (Table 2).

Fourteen of the 22 tRNA genes were situated on the H-strand and eight on the L-strand (Table 2; Fig. 1). The cloverleaf structure of trnS1 lacked the D-loop in all species, while trnK and trnR lacked the TΨC-loop in *E. aeneus* and *E. tabanoides*, respectively (File S2[1]). The two rRNA genes were situated on the H-strand (Table 2; Fig. 1).

Overlapping genes were found two times on the same strand (atp8/atp6 and nad4/nad4l), and two times on different strands (trnW/trnC and trnY/cox1) in all five species. Additional overlaps were observed on the same strand for nad6/cob in all species except for *E. aeneus*, and on different strands for trnI/trnQ in *E. barclayi*, *E. fuscicornis*, and *E. vicarians* (Table 2).

Uncorrected p-distances measured among the five mitogenomes of *Eristalinus* (File S3[1]) ranged from 0.006 (*E. fuscicornis* – *E. barclayi*) to 0.112 (*E. aeneus* – *E. tabanoides*). Surprisingly, intergeneric distances measured between *Episyrphus* and *Simosyrphus* (0.019) were relatively small compared to the interspecific divergences measured within *Eristalinus*. They were also much smaller than the other intergeneric distances measured within Syrphidae (ranging from 0.111 between *Eristalis tenax* and *Eristalinus fuscicornis* to 0.198 between *Ocyptamus sativus* and *Eristalinus aeneus*).

The phylogenetic trees reconstructed using ML and BI had almost the same topology, the only difference between the two analyses being the sister relationship between *E. tabanoides* and *E. vicarians* that is suggested by the BI but remained unresolved in the ML analysis (Fig. 2). These analyses suggested that *Eristalinus aeneus* was the sister species of the other four species of *Eristalinus* sampled here. It also suggested that the two species belonging to the subgenus *Eristalodes*, *E. barclayi* and *E. fuscicornis*, were sister species and formed a clade that was nested within the genus *Eristalinus*, making the subgenus *Eristalinus* paraphyletic with respect to the subgenus *Eristalodes*. The phylogenetic analyses (ML and BI) based on cox1 (alignment in File S4[1]) and including the additional samples of *Eristalinus* sequenced by Pérez-Bañón et al. (2003) were not sufficiently resolved to evaluate the reciprocal monophyly of the two subgenera (Fig. 3). In these trees, *E. tabanoides* and *E. megacephalus* were sister species, and *E. vicarians* and *E. dubiosus* were sister species. In contrast, the two specimens of *E. aeneus* (one of this study, one of Pérez-Bañón et al. (2003)) did not cluster with one another.

A plot of the net PI profiles on the same time scale (*x*-axis) as the Bayesian phylogenetic ultrametric tree obtained for the dataset including all PCGs and rRNA genes

Fig. 1. (A) Schematic representation of the complete mitogenome of *Eristalinus aeneus*. Black circles represent the DNA strands with numbers indicating positions in base pairs. Protein coding, rRNA, and tRNA genes are represented by green arrows, red arrows, and pink triangles, respectively. The inner blue ring depicts the GC content (with lower values towards the center of the circle). (B) Similarity among the mitogenomes of the five species of *Eristalinus* sequenced here (mean pairwise similarity over all pairs of sequences at positions covered by a sliding window of 99 bp) and GC content for each species (sliding window of 50 bp).

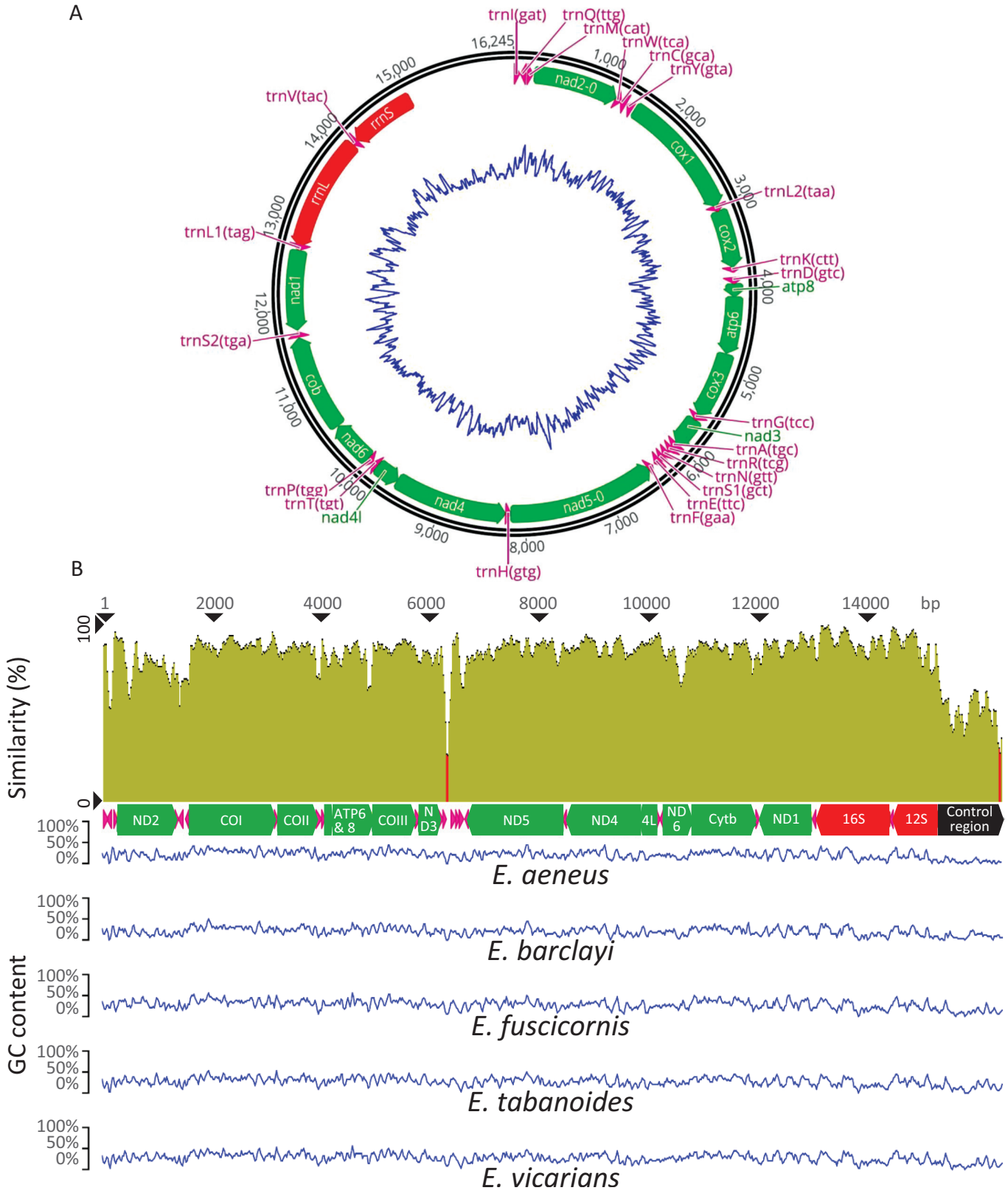


Fig. 2. Phylogenetic trees of the six species of Eristalinae (five *Eristalinus* sequenced here and one *Eristalis* from GenBank) based on the concatenation of all protein coding and rRNA genes. Both maximum likelihood analysis (ML) and Bayesian inference (BI) were rooted using the four representatives of Syrphinae (*) available in GenBank as outgroup. Bootstrap support (ML) and posterior probabilities (BI) are indicated at nodes and were collapsed when <70% and <0.95, respectively.

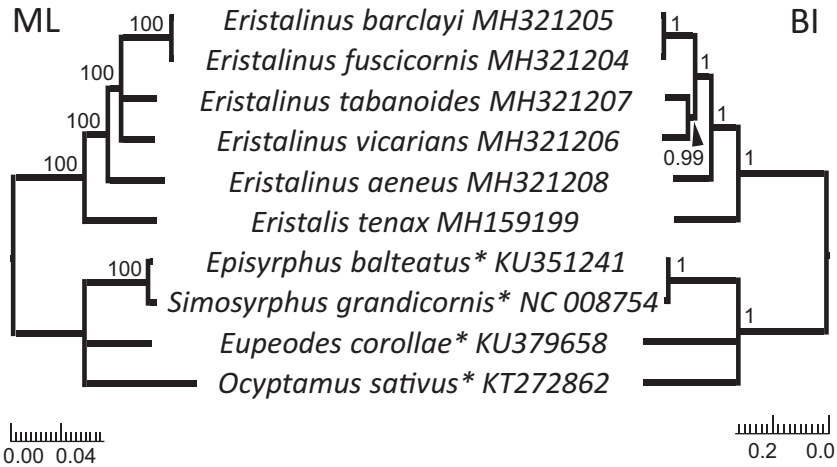
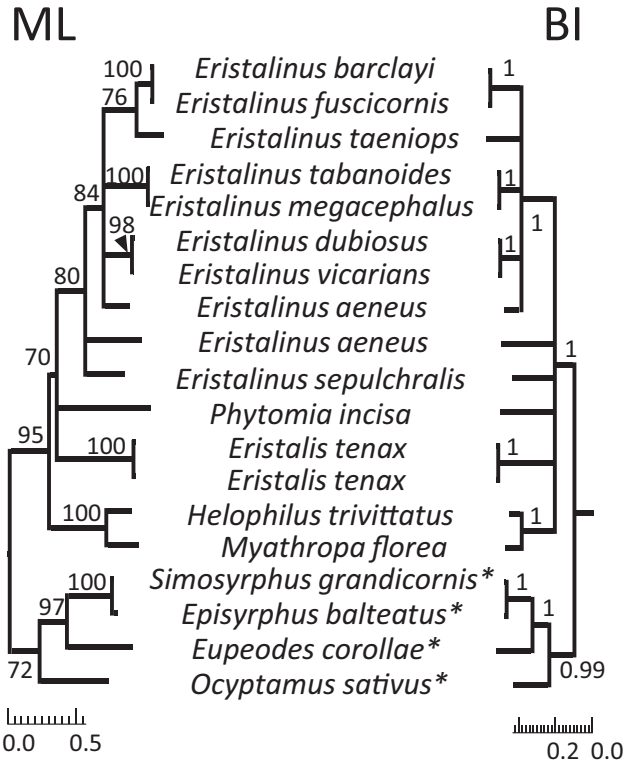


Fig. 3. Phylogenetic trees of the species of Eristalinae sequenced here and by Pérez-Bañón et al. (2003), and based on cox1. Both maximum likelihood analysis (ML) and Bayesian inference (BI) were rooted using the representatives of Syrphinae (*) as outgroup. Bootstrap support (ML) and posterior probabilities (BI) are indicated at nodes and were collapsed when <70% and <0.95, respectively.

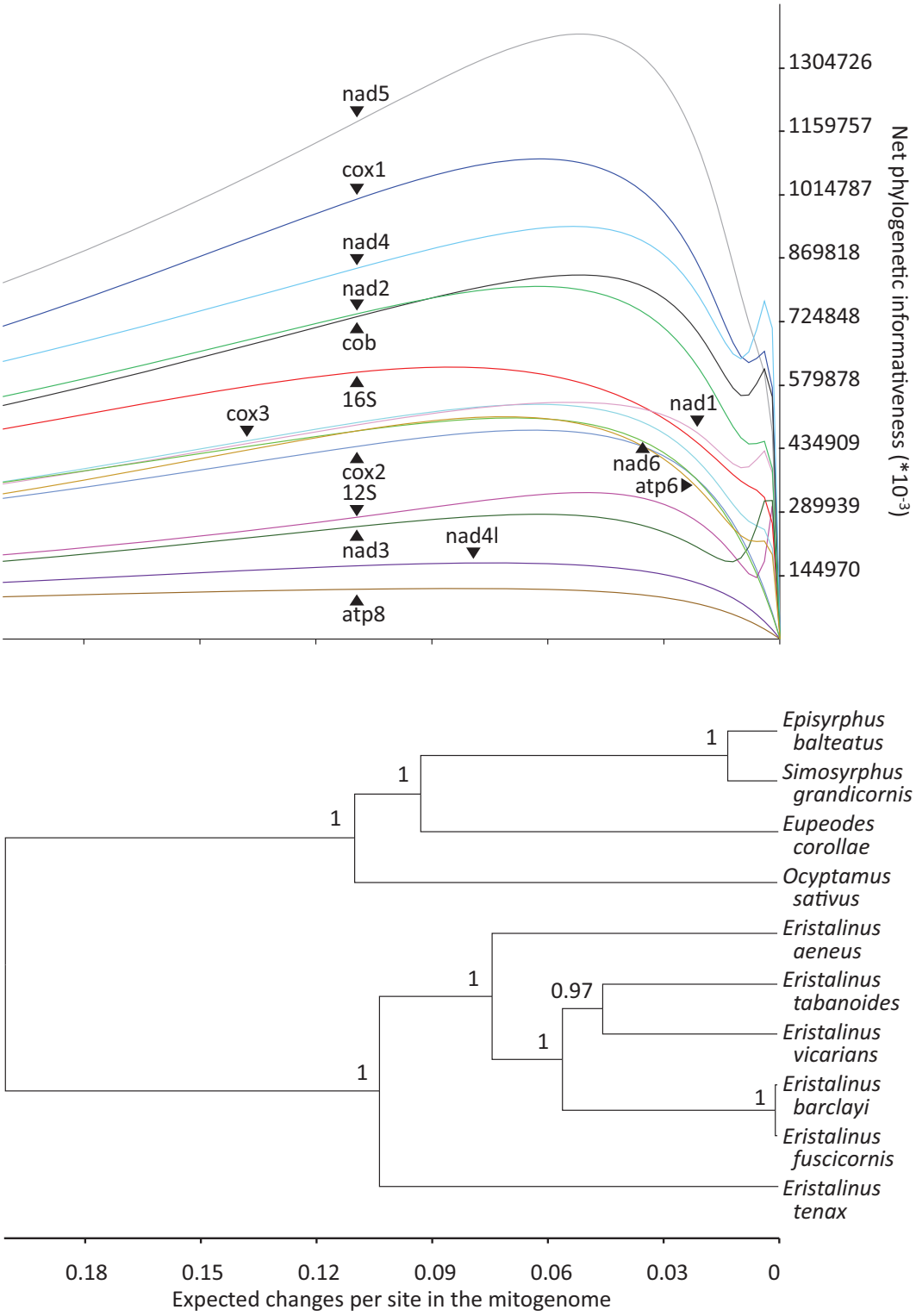


is given in Fig. 4. The genes showing the highest PI profiles are nad5, cox1, and nad4, followed by nad2, cob, and 16S rRNA, indicating an overall higher utility of these genes for phylogenetic inference compared to the other genes such as atp8, nad4l, nad3, and 12S rRNA, which had flatter PI profiles (Fig. 4). For most genes, the optima of the PI profiles were situated after the divergence between *E. aeneus* and the other representatives of *Eristalinus*. This means that, for these genes, character changes (mutations) were estimated to mainly occur along the branches of the subtree corresponding to the genus *Eristalinus*. These genes are therefore more promising to resolve intrageneric relationships. For several genes, especially for 16S rRNA, the PI profile also showed a narrow peak next to the tips of the tree (between the bipartition leading to *Episyrphus balteatus* and *Simosyrphus grandicornis* and the bipartition leading to *Eristalinus barclayi* and *E. fuscicornis*). These genes may therefore have a relatively higher power to resolve more recent divergences.

Discussion

This study presents the first complete and annotated mitogenomes for the hoverfly genus *Eristalinus*. The only other published mitogenomes of Syrphidae are from one species of the subfamily Eristalinae, *Eristalis tenax* (Li et al. 2017), and four species of the subfamily Syrphinae: *Episyrphus balteatus* and *Eupeodes corollae* (Pu et al. 2017), *Ocyptamus sativus* and *Simosyrphus grandicornis* (Junqueira et al. 2016), and the partial mitogenome of an unknown species "Syrphidae sp." (Tang et al. 2014, accession number: KM244713). Mitogenome sizes of *Eristalinus* (15 757 – 16 245 bp) are comparable to those already published for the family (15 214 – 16 175 bp, see Table 3), with *E. aeneus* showing the largest published genome of Syrphidae to date (16 245 bp). Gene order is identical in all Syrphidae and in line with previously reported dipteran mitogenomes (e.g., Li et al. 2017; Pu et al. 2017; Yong et al. 2015). This result is not surprising because gene composition and order is well conserved within Diptera (Wolstenholme 1992) even if cases of tRNA gene duplica-

Fig. 4. Net phylogenetic informativeness (PI) of each mitochondrial gene. The net PI (above) is plotted in reference to the indirect measure of time of the ultrametric tree (below) constructed using the concatenated sequences of all protein coding and rRNA genes and where posterior probabilities are given at nodes.



tion have been observed within the family Calliphoridae (Duarte et al. 2008; Junqueira et al. 2004).

The mitogenomes obtained in this study show a promising array of diversity across the genus *Eristalinus*, with both large p-distances (between *E. aeneus* and the other species of *Eristalinus*) and remarkably low pairwise p-distance between the two representatives of the subgenus *Eristalodes* (File S3[1]), suggesting a close relationship between both species. The latter is corroborated with species delimitation methods applied to mitochondrial

and nuclear markers which show that *E. barclayi* and *E. fuscicornis* form a species complex, also comprising *Eristalinus quinquelineatus* (Fabricius, 1797) (De Smet et al., unpublished data). These results either question the taxonomic value of the morphological characters used to distinguish the three species, or illustrates a recent divergence, ongoing speciation, hybridization, or introgression. The phylogenetic analyses of the mitogenomes of the five species of *Eristalinus* suggest that the subgenera *Eristalinus* and *Eristalodes* are not reciprocally monophyletic. Indeed, the subgenus *Eristalodes* renders the subgenus *Eristalinus* paraphyletic (Fig. 2). In a phylogenetic study (parsimony analysis) of five species of *Eristalinus* based on cox1 and the 28S rRNA gene (Pérez-Bañón et al. 2003), *E. taeniops*, the species type of the subgenus *Eristalodes* was nested within the subgenus *Eristalinus*. The lack of resolution of the phylogenetic trees obtained here using cox1 only (Fig. 3) does not contribute to solve this question. However, since our study and that of Pérez-Bañón et al. (2003) only included a limited number of species of the subgenus *Eristalodes*, more comprehensive phylogenetic analyses, including more species of both subgenera and species of the remaining three subgenera, are needed to test for the subgeneric rank of the different subgenera in this genus. The phylogenetic trees based on cox1 (Fig. 3) also revealed that the specimens identified as *E. aeneus* here in and in the study of Pérez-Bañón et al. (2003) may belong to different lineages. This species has a very wide distribution and probably may comprise cryptic species.

Assessing the PI profiles for all PCGs within the genus *Eristalinus* revealed substantial difference in their suggested utility for phylogeny reconstruction (Fig. 4). Some mitochondrial markers (nad5, cox1, nad4, nad2, cob, and the 16S rRNA gene) exhibit higher PI profiles, especially for the time range corresponding to the diversification of the species of *Eristalinus* considered here. The utility of mitogenomics for the study of insect evolution and phylogeny has been amply demonstrated (Cameron 2013). Mitogenomes are available for all insect orders and have been shown to hold phylogenetic information over extensive taxonomic scales (e.g., Cameron et al. 2007, 2009; Logue et al. 2013; Ma et al. 2012; Nelson et al. 2012; Zhao et al. 2013). These results suggest that the sequencing of the mitogenomes of additional syrphid species would be useful to investigate the systematics of the family and clarify the classification within *Eristalinus*.

Acknowledgements

G.S., Y.D.S., M.V., T.B., M.D.M., and K.J. acknowledge financial support of the Belgian Science Policy (BELSPO) for part of this work through the Joint Experimental Molecular Unit (JEMU), the Barcoding Facility for Organisms and Tissues of Policy Concern (BopCo), and the project BR/314//PI/SYRPINTINE to K.J. Research work was done with IITA and is based on bilateral agreements in form of memorandums of understanding (MoU) signed by the ministries of agriculture of the respective countries (more information can be found on http:/:www.iita.org/). In this framework, research work in the field is an integral part of IITA's contracted mandate. Therefore, no specific permissions were required for the collected hoverfly material. None of the hoverfly species figure in any red list, are endangered, threatened, or considered to be endangered in the involved countries. No species collected in the present study are ranked in any IUCN list or are protected by CITES. The authors thank the reviewers for their pertinent suggestions.

References

Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsch, G. et al. 2013. MITOS: improved de novo metazoan mitochondrial genome annotation. Mol. Phylogenet. Evol. **69**(2): 313–319. doi:10.1016/j.ympev.2012.08.023. PMID:22982435.

Cameron, S.L. 2013. Insect mitochondrial genomics: implications for evolution and phylogeny. Annu. Rev. Entomol. **59**(1): 95–117. doi:10.1146/annurev-ento-011613-162007. PMID:24160435.

Cameron, S.L. 2014. How to sequence and annotate insect mitochondrial genomes for systematic and comparative genomics research. Syst. Entomol. **39**(3): 400–411. doi:10.1111/syen.12071.

Cameron, S.L., Lambkin, C.L., Barker, S.C., and Whiting, M.F. 2007. A mitochondrial genome phylogeny of Diptera: whole genome sequence data accurately resolve relationships over broad timescales with high precision. Syst. Entomol. **32**(1): 40–59. doi:10.1111/j.1365-3113.2006.00355.x.

Cameron, S.L., Sullivan, J., Song, H., Miller, K.B., and Whiting, M.F. 2009. A mitochondrial genome phylogeny of the Neuropterida (lace-wings, alderflies and snakeflies) and their relationship to the other holometabolous insect orders. Zool. Scr. **38**(6): 575–590. doi:10.1111/j.1463-6409.2009.00392.x.

Duarte, G.T., De Azeredo-Espin, A.M., and Junqueira, A.C. 2008. The mitochondrial control region of blowflies (Diptera: Calliphoridae): a hot spot for mitochondrial genome rearrangements. J. Med. Entomol. **45**(4): 667–676. doi:10.1093/jmedent/45.4.667. PMID:18714866.

Edgar, R.C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. **32**(5): 1792–1797. doi:10.1093/nar/gkh340. PMID:15034147.

Gramates, L.S., Marygold, S.J., dos Santos, G., Urbano, J.-M., Antonazzo, G., Matthews, B.B. et al. 2017. FlyBase at 25: looking to the future. Nucleic Acids Res. **45**(Database issue): D663–D671. doi:10.1093/nar/gkw1016. PMID:27799470.

Hippa, H., and Ståhls, G. 2005. Morphological characters of adult Syrphidae: Descriptions and phylogenetic utility. Acta Zool. Fenn. **215**: 1–72.

Junqueira, A.C., Azeredo-Espin, A.M., Paulo, D.F., Marinho, M.A., Tomsho, L.P., Drautz-Moses, D.I. et al. 2016. Large-scale mitogenomics enables insights into Schizophora (Diptera) radiation and population diversity. Sci. Rep. **6**: 21762. doi:10.1038/srep21762. PMID:26912394.

Junqueira, A.C., Lessinger, A.C., Torres, T.T., da Silva, F.R., Vettore, A.L., Arruda, P., and Azeredo Espin, A.M. 2004. The mitochondrial genome of the blowfly *Chrysomya chloropyga* (Diptera: Calliphoridae). Gene, **339**(15): 7–15. doi:10.1016/j.gene.2004.06.031. PMID:15363841.

Li, X., Ding, S., Li, X., Peng, H., Tang, C., and Yang, D. 2017. The complete mitochondrial genome analysis of *Eristalis tenax* (Diptera, Syrphidae). Mitochondrial DNA, Part B, **2**(2): 654–655. doi:10.1080/23802359.2017.1375875.

Logue, K., Chan, E.R., Phipps, R., Small, S., Reimer, L., Henry-Halldi, C. et al. 2013. Mitochondrial genome sequences reveal deep divergences among *Anopheles punctulatus* sibling

species in Papua New Guinea. Malaria J. **12**: 64. doi:10.1186/1475-2875-12-64. PMID:23405960.

López-Giráldez, F., and Townsend, J.P. 2011. PhyDesign: an online application for profiling phylogenetic informativeness. BMC Evol. Biol. **11**: 152. doi:10.1186/1471-2148-11-152. PMID:21627831.

Ma, C., Yang, P.C., Jiang, F., Chapuis, M.-P., Shall, Y., Sword, G.A., and Kang, L. 2012. Mitochondrial genomes reveal the global phylogeography and dispersal routes of the migratory locust. Mol. Ecol. **21**(17): 4344–4358. doi:10.1111/j.1365-294X.2012.05684.x. PMID:22738353.

Mengual, X., Ståhls, G., and Rojo, S. 2015. Phylogenetic relationships and taxonomic ranking of pipizine flower flies (Diptera: Syrphidae) with implications for the evolution of aphidophagy. Cladistics, **31**(5): 491–508. doi:10.1111/cla.12105.

Miller, M.A., Pfeiffer, W., and Schwartz, T. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. *In* Proceedings of the Gateway Computing Environments Workshop (GCE), New Orleans, Louisiana, 14 November 2010. Institute of Electrical and Electronics Engineers (IEEE), New Orleans, Louisiana. pp. 1–8.

Nelson, L.A., Lambkin, C.L., Batterham, P., Wallman, J.F., Dowton, M., Whiting, M.F. et al. 2012. Beyond barcoding: A mitochondrial genomics approach to molecular phylogenetics and diagnostics of blowflies (Diptera: Calliphoridae). Gene, **511**(2): 131–142. doi:10.1016/j.gene.2012.09.103. PMID:23043935.

Pérez-Bañón, C., Rojo, S., Ståhls, G., and Marcos-García, M.A. 2003. Taxonomy of European *Eristalinus* (Diptera: Syrphidae) based on larval morphology and molecular data. Eur. J. Entomol. **100**(3): 417–428. doi:10.14411/eje.2003.064.

Pu, D.-q., Liu, H.-l., Gong, Y.-y., Ji, P.-c., Li, Y.-j., Mou, F.-S., and Wei, S.-j. 2017. Mitochondrial genomes of the hoverflies *Episyrphus balteatus* and *Eupeodes corollae* (Diptera: Syrphidae), with a phylogenetic analysis of Muscomorpha. Sci. Rep. **7**: 44300. doi:10.1038/srep44300. PMID:28276531.

Ronquist, F., and Huelsenbeck, J.P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics, **19**(12): 1572–1574. doi:10.1093/bioinformatics/btg180. PMID:12912839.

Skevington, J.H., and Yeates, D.K. 2000. Phylogeny of the Syrphoidea (Diptera) inferred from mtDNA sequences and morphology with particular reference to classification of the Pipunculidae (Diptera). Mol. Phylogenet. Evol. **16**(2): 212–224. doi:10.1006/mpev.2000.0787. PMID:10942608.

Ståhls, G., Hippa, H., Rotheray, G.E., Muona, J., and Gilbert, F. 2003. Phylogeny of Syrphidae (Diptera) inferred from combined analysis of molecular and morphological characters. Syst. Entomol. **28**(4): 433–450. doi:10.1046/j.1365-3113.2003.00225.x.

Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics, **30**(9): 1312–1313. doi:10.1093/bioinformatics/btu033. PMID:24451623.

Tang, M., Tan, M., Meng, G., Yang, S., Su, X., Liu, S. et al. 2014. Multiplex sequencing of pooled mitochondrial genomes — a crucial step toward biodiversity analysis using mitogenomics. Nucleic Acids Res. **42**(22): e166. doi:10.1093/nar/gku917. PMID:25294837.

Tang, M., Hardman, C.J., Ji, Y., Meng, G., Liu, S., Tan, M. et al. 2015. High-throughput monitoring of wild bee diversity and abundance via mitogenomics. Methods Ecol. Evol. **6**(9): 1034–1043. doi:10.1111/2041-210X.12416. PMID:27867467.

Thompson, F.C. 1999. A key to the genera of the flower flies of the Neotropical Region including the descriptions of genera and species and a glossary of taxonomic terms. Contributions on Entomology, International, **3**(3): 321–378.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22**(22): 4673–4680. doi:10.1093/nar/22.22.4673. PMID:7984417.

Townsend, J.P. 2007. Profiling phylogenetic informativeness. Syst. Biol. **56**(2): 222–231. doi:10.1080/10635150701311362. PMID:17464879.

Velozo Timbó, R., Coiti Togawa, R., Costa, M.A., Andow, D., and Paula, D.P. 2017. Mitogenome sequence accuracy using different elucidation methods. PLoS ONE, **12**(6): e0179971. doi:10.1371/journal.pone.0179971. PMID:28662089.

Wolstenholme, D.R. 1992. Animal mitochondrial DNA: structure and evolution. Int. Rev. Cytol. **141**: 173–216. doi:10.1016/S0074-7696(08)62066-5. PMID:1452431.

Xie, Y., Wu, G., Tang, J., Luo, R., Patterson, J., Liu, S. et al. 2014. SOAPdenovo-Trans: *de novo* transcriptome assembly with short RNA-Seq reads. Bioinformatics, **30**(12): 1660–1666. doi:10.1093/bioinformatics/btu077. PMID:24532719.

Yong, H.-S., Song, S.-L., Lim, P.-E., Chan, K.-G., Chow, W.-L., and Eamsobhana, P. 2015. Complete mitochondrial genome of *Bactrocera arecae* (Insecta: Tephritidae) by next-generation sequencing and molecular phylogeny of Dacini tribe. Sci. Rep. **5**: 15155. doi:10.1038/srep15155. PMID:26472633.

Young, A.D., Lemmon, A.R., Skevington, J.H., Mengual, X., Ståhls, G., Reemer, M. et al. 2016. Anchored enrichment dataset for true flies (order Diptera) reveals insights into the phylogeny of flower flies (family Syrphidae). BMC Evol. Biol. **16**(1): 143. doi:10.1186/s12862-016-0714-0. PMID:27357120.

Zhao, Z., Su, T., Chesters, D., Wang, S., Ho, S.Y.W., Zhu, C., and Chen, X. 2013. The mitochondrial genome of *Elodia flavipalpis* Aldrich (Diptera: Tachinidae) and the evolutionary timescale of tachinid flies. PLoS ONE, **8**(4): e61814. doi:10.1371/journal.pone.0061814. PMID:23626734.

Zhou, X., Li, Y., Liu, S., Yang, Q., Su, X., Zhou, L., Tang, M., Fu, R., Li, J., and Huang, Q. 2013. Ultra-deep sequencing enables high-fidelity recovery of biodiversity for bulk arthropod samples without PCR amplification. Gigascience, **2**(1): 4. doi:10.1186/2047-217X-2-4. PMID:23587339.