Evolutionary Ecology Research, 2017, **18**: 443–457

Ambient temperature and host specialization drive mitochondrial genome evolution in fruit flies of the genus *Bactrocera* (Diptera: Tephritidae)

Hichem Ben Slimen, A. Awadi and M. Makni

Unité de Recherche Génomique des Insectes Ravageurs des Cultures d'Intérêt Agronomique, Faculté des Sciences de Tunis, Université de Tunis El-Manar, El Manar, Tunis, Tunisia

ABSTRACT

Background: Recent studies of selection on mitochondrial coding genes suggest adaptation due mainly to environmental variation.

Question: What is the extent of positive selection on the mitochondrial genomes of fruit flies? Are amino acid substitutions on these genomes affected by climate variation and/or host type (i.e. polyphagous, oliphagous or monophagous)?

Data incorporated: Seventeen complete mitochondrial genomes of fruit flies of the genus *Bactrocera* (retrieved from Genbank).

Method of analysis: Site and branch-site methods were applied to detect amino acid positions and lineages under positive selection. Relationships between genotypes and bioclimatic data were assessed using correlation tests, while genetic variation between polyphagous, oliphagous, and monophagous insects was assessed using a Kruskal-Wallis *H*-test. In addition, we applied a multiple regression analysis, with the genetic parameters as the dependent (outcome) variables, and the bioclimatic factors and host type as the predictor (independent) variables. Finally, we compared rates of evolution of the 13 mitochondrial coding genes with the corresponding predicted ancestral sequences.

Conclusions: Site and branch-site applied tests showed evidence of positive selection in 37 codons in 11 of the 13 mitochondrial coding genes, mostly affecting the oligophagous lineages. We found strong relationships between the genetic parameters and mean diurnal range, suggesting an effect of ambient temperature in the evolution of the genus *Bactrocera*. Moreover, a significant increase of amino acid substitutions from the predicted ancestor suggests an effect of the observed mitochondrial variation in the specialization process from polyphagous to monophagous species. Finally, we showed that NADH dehydrogenase subunit 6, one of the 13 mitochondrial coding genes, evolves independently from all the other genes of the same linkage system.

Keywords: *Bactrocera*, environmental adaptation, host specialization, mtDNA, positive selection

Correspondence: Hichem Ben Slimen, Unité de Recherche Génomique des Insectes Ravageurs des Cultures d'Intérêt Agronomique, Faculté des Sciences de Tunis, Université de Tunis El-Manar, El Manar, 2092 Tunis, Tunisia. email: ben_slimen_hichem@yahoo.fr

INTRODUCTION

Mitochondrial DNA (mtDNA) is a widely used molecular tool in population genetics, phylogeographical, and phylogenetic studies in eukaryotes, with most studies reliant explicitly or implicitly on the neutral mutation theory. Neutral scenarios are commonly advanced to explain population-level phenomena involving mtDNA, such as population structure and dynamics, mtDNA gene flow, genetic breaks in species distributions, and levels of mtDNA polymorphism within a species. However, in recent years, accumulating evidence of the selective constraints on mtDNA has led to intense discussion of whether mainly neutral evolutionary or adaptive evolutionary forces are shaping variation in mtDNA (Nachman *et al.*, 1996; Da Fonseca *et al.*, 2008; Stewart *et al.*, 2008; Ballard and Melvin, 2010; Melo-Ferreira *et al.*, 2014; Morales *et al.*, 2015). In particular, studies of the variation in human mtDNA protein coding genes (Mishmar *et al.*, 2003; Ruiz-Pesini *et al.*, 2004; Ingman and Gyllensten, 2007; Balloux *et al.*, 2009) have suggested non-neutral evolution of the mitochondrial genome, with climate as one selective influence. Additional evidence of selective pressures can be seen in the deleterious effects of mtDNA variation, where a single substitution could be responsible for serious diseases in humans (e.g. Taylor and Turnbull, 2005; Tuppen *et al.*, 2010). Moreover, it has also been suggested that mtDNA is involved in adaptation to high altitude, temperature, colder climate, food availability, and habitat change (see Garvin *et al.*, 2014).

In the genus *Bactrocera* (Diptera: Tephritidae: Dacinae), the mitochondrial genome contains 37 genes, made up of 13 protein-encoding genes, 22 transfer RNA genes, and 2 ribosomal RNA genes. All 13 protein-encoding genes are components of the mitochondrial respiratory chain (OXPHOS): these are 7 subunits of the NADH dehydrogenase complex I (ND1, ND2, ND3, ND4, ND4L, ND5, and ND6), 1 subunit (cytochrome b) of the cytochrome cb1 complex III, 3 subunits of cytochrome c oxidase complex (COX1, COX2, and COX3), and 2 subunits of ATPase complex V (ATP6 and ATP8) (Zhang *et al.*, 2014; Choudhary *et al.*, 2015; Tan *et al.*, 2016; Yong *et al.*, 2016).

Fruit flies of the genus *Bactrocera* have a worldwide distribution. Most are highly polyphagous species that use a wide variety of host plants (Vargas *et al.*, 2015). These fruit fly species display strong behavioural and biological interactions with their host plants. Such a dependent relationship has influenced their evolution and their processes of speciation (Drew and Hancock, 2001). Indeed, Funk (2010) suggested that host plants affect ecological divergence, reproductive isolation, and host-related selection in insect species. In addition, Kelley *et al.* (2000) showed that host specialization affects genetic differentiation in insects, while Nosil (2002) showed that it affects their transition rates.

In this study, we tested for signatures of natural selection, using different statistical methods, on 17 mitochondrial genomes (available from GenBank) of fruit flies of the genus *Bactrocera*. Furthermore, we analysed the effect of different climatic factors and host range on the positively selected sites, and also on genetic variation as estimated by synonymous and non-synonymous changes from the putative ancestral sequences. Finally, we expected a strong correlation between the 13 mitochondrial loci under study as an indication of a significant signal of co-evolution of these genes, in particular because non-recombining mtDNA can be considered a single linkage system.

MATERIALS AND METHODS

Sequences and database searches

We re-analysed 17 mitochondrial genomes of the genus *Bactrocera* studied previously that we retrieved from GenBank (Table 1). We wished to determine whether positive selection has played a significant role in the evolution of the mtDNA coding genes of the *Bactrocera* fruit flies. This aspect was not covered in any of the mtDNA studies of this genus, in which the mitochondrial data were only used to construct phylogenetic relationships.

Molecular evolution analysis

First, we tested for positive selection at specific codons of the 13 protein-coding mitochondrial genes using the site model implemented in CODEML, in the PAML 4 package (Yang, 2007). We evaluated the presence of positive selection ($\omega > 1$) using a likelihood-ratio test (LRT) after comparing the distribution of a null model that does not allow for any codons to be $\omega > 1$ against a model that does. Six different models proposed by Yang *et al.* (2000, 2005) were compared: M0 (one ω ratio), M1a (nearly neutral), M2a (positive selection), M3 (discrete), M7 (beta), and M8 (beta & ω).

We further applied additional codon models to assess the impact of selection along the mtDNA phylogeny of fruit flies using five different methods on the DATAMONKEY web server [http://www.datamonkey.org/; last accessed 10 March 2016] (Pond and Frost, 2005):

- Single Likelihood Ancestral Counting (SLAC)
- Fixed Effects Likelihood (FEL)

Table 1. GenBank accession numbers, species designations, and references for mtDNA genomes of *Bactrocera* species downloaded from GenBank and used in the current study

GenBank accession #	Species	References
1. KT881556	*B. latifrons*	Yong *et al.* (2016)
2. KT881557	*B. melastomatos*	Yong *et al.* (2016)
3. KT881558	*B. umbrosa*	Yong *et al.* (2016)
4. NC_005333	*B. oleae*	Nardi *et al.* (2003)
5. NC_008748	*B. dorsalis*	Unpublished results
6. NC_009770	*B. papayae*	Unpublished results
7. NC_009771	*B. philippinensis*	Unpublished results
8. NC_009772	*B. carambolae*	Unpublished results
9. NC_014402	*B. minax*	Zhang *et al.* (2014)
10. NC_014611	*B. tryoni*	Nardi *et al.* (2010)
11. NC_016056	*B. cucurbitae*	Wu *et al.* (2013)
12. NC_018787	*B. correcta*	Unpublished results
13. NC_027254	*B. scutellata*	Unpublished results
14. NC_027290	*B. tau*	Tan *et al.* (2016)
15. NC_027725	*B. zonata*	Choudhary *et al.* (2015)
16. NC_028327	*B. arecae*	Yong *et al.* (2015)
17. NC_028347	*B. diaphora*	Zhang *et al.* (2015)

- Random Effects Likelihood (REL)
- Fast Unconstrained Bayesian AppRoximation (FUBAR)
- Mixed Effects Model of Evolution (MEME) (Murrell *et al.*, 2012, 2013).

Levels of significance of $P < 0.25$ in SLAC and FEL and $P < 0.05$ in MEME and Bayes factors >50 in REL were used to indicate positively selected sites. In all tests performed in DATAMONKEY, the appropriate evolutionary model for each data set was used after direct estimation on this web server.

We also applied the branch-site model (Yang *et al.*, 2005; Zhang *et al.*, 2005), implemented in CODEML, and the branch-site REL (Pond *et al.*, 2011), implemented in DATAMONKEY. For the CODEML method, all individual branches were tested as possible foreground lineages. The Bayes empirical method was used to identify sites under positive selection in the case of rejection of the null model (Yang *et al.*, 2005). We used the branch-site REL (Pond *et al.*, 2011) to perform a series of likelihood-ratio tests to identify all lineages in a phylogeny where a proportion of sites evolves with dN/dS > 1. This model is a generalization of existing branch-site methods.

We also applied TreeSAAP (Woolley *et al.*, 2003) to the current data set. TreeSAAP takes into account the magnitude of the impact of the amino acid replacements on local physicochemical properties. We interpreted a radical magnitude of change ≥6 (with $P \leq 0.001$) to indicate positive directional selection for a given physicochemical property. In addition, we only considered amino acid properties that had an accuracy of detection of selection greater than 85% (McClellan and Ellison, 2010). This meant we excluded 11 of the 31 properties analysed by TreeSAAP.

Relationships between amino acid changes and climatic variation

Bioclimatic data were obtained from the WORLDCLIM data set for 2.5-minute intervals [v.1.4, http://www.worldclim.org/bioclim.htm]. Nineteen bioclimatic variables were automatically extracted using DIVA-GIS v.7.5. Since a correlation between climatic variables can lead to incorrect results and that the use of all 19 variables can lead to overparameterization, we first applied a Spearman rank correlation test to each pair of variables using SPSS v.18.0 statistical software (SPSS Inc., 2009). Of all the variables tested, only four were selected and these did not correlate with each other:

- mean diurnal range [mean of monthly (maximum – minimum) temperature] (BIO2)
- maximum temperature of warmest month (BIO5)
- annual precipitation (BIO12)
- seasonality of precipitation (coefficient of variation) (BIO15) (see http://www.evolutionary-ecology.com/data/3082Appendix.pdf).

Accumulation of synonymous and non-synonymous substitutions through the evolutionary history of the *Bactrocera* genus has led to the current taxonomy of the species in question since divergence from their common ancestor. Comparing substitution changes with the ancestral state can be a useful estimator of genetic variation. Therefore, we reconstructed the ancestral sequence of all *Bactrocera* species studied using the maximum likelihood ancestral sequence reconstruction method implemented in FASTML [http://fastml.tau.ac.il/; last assessed 15 June 2016] (see 3082Appendix.pdf). We then calculated

non-synonymous and synonymous changes (dN, dS) and dN/dS ratios between the putative ancestral sequence and each sequence using DnaSP v.5.1 (Librado and Rozas, 2009). We also calculated the overall number of positively selected sites in each sequence compared with its ancestor. Non-synonymous and synonymous changes and number of positively selected sites (when considering all 13 genes) for each species (designated in the text by genetic variables) were tested for relationships with bioclimatic variables. Moreover, host range for each species (monophagous, oligophagous, polyphagous) as obtained from Vargas *et al.* (2015) was included as a factor that may influence substitutions in the mitochondrial genome.

Before analysis, we checked the data for normality using the Shapiro-Wilk *H*-test implemented in SPSS. Among all variables, only BIO2, BIO15, and number of positively selected sites were normally distributed (see 3082Appendix.pdf). Since all transformations applied to the non-normally distributed variables (log normal, log decimal, square root, inverse) failed to meet the assumptions of normality, we compared the original data sets with the variables after ln-transformation. First, associations between genotypes and bioclimatic data were assessed using a Pearson correlation test or a Spearman rank test (when normality of the data was not met). Second, variation in the genetic variables between the three host ranges was assessed using a Kruskal-Wallis *H*-test except for number of positively selected sites, for which we used an analysis of variance (ANOVA). Finally, we applied a multiple regression analysis in which the genetic parameters were used as the dependent (outcome) variables and the bioclimatic factors and host range as the independent variables. Thus we assessed the individual and combined contributions of these factors to influencing the genetic profiles.

Relative rates of evolution and co-evolution analyses

Comparing the total number of substitutions in a group of related species with their ancestor offers the chance to explore the relative evolutionary rates of each mitochondrial gene. Such determination might be useful when comparing inferred systematic or evolutionary history based on different gene sequences. Therefore, the evolutionary rates of the 13 mtDNA coding genes of *Bactrocera* species were compared with their corresponding ancestral sequences. Such direct estimation of total numbers of mutations compared with the ancestral sequences allows differential evolution rates to be traced. We used graphic representations of variation in the 13 coding genes for each species whereby levels of variation and unexpected variation could be observed. Moreover, a high positive correlation is expected to occur between all 13 co-evolving mitochondrial coding genes from the same linkage system. Thus we used Spearman rank correlation to test for pairwise correlations among the 13 mitochondrial genes.

RESULTS

Evidence of natural selection

Comparisons of model results obtained with PAML indicate the absence of sites under positive selection and suggest that the mitochondrial genome in *Bactrocera* has evolved in a neutral fashion (data not shown). However, the other methods did suggest evidence of positive selection affecting mtDNA protein-coding genes of *Bactrocera* fruit flies. Only sites suggested to be under positive selection by more than one method are shown in Table 2.

Table 2. Sites under positive selection according to seven different methods

		DATAMONKEY selection tests					PAML	TreeSAAP	
	Site	SLAC	REL	FEL	MEME	FUBAR	Branch-sites	Property	Direction
ATP8	45	×	—	×	—	—	—	—	—
ATP6	137	—	—	—	—	—	B (2)	—	—
COX1	475	—	×	×	—	—	—	*Ra* *pK′*	pos neg
	508	—	—	×	—	—	—	*pK′*	pos
COX3	209	—	—	—	—	—	B (9)	*p*	pos
	222	—	—	—	—	—	B (9)	*Br* *p* *Hp*	pos neg pos
	244	—	—	—	—	—	B (9)	*p*	neg
ND1	175	—	—	—	×	—	—	*pK′*	neg
	254	×	—	—	—	—	—	*pK′*	neg
	312	—	—	—	×	—	B (9)	—	—
ND2	4	—	×	—	—	×	—	*p*	neg
	47	×	×	×	—	×	—	—	—
	55	×	—	×	—	—	—	—	—
	310	—	—	—	—	—	—	*Ra* *pHi* E_t	pos pos neg
ND3	94	—	×	—	—	—	—	*pK′*	pos
	102	—	—	—	—	—	B (14)	*pK′*	pos
ND4	9	×	—	—	×	—	—	*h* *Hp* *pK′*	pos pos neg
	23	—	×	×	—	—	—	—	—
	37	—	—	—	—	—	—	*pK′*	pos
	54	—	—	—	—	—	B (9)	*h*	neg
	55	—	—	—	—	—	B (9)	*Hp*	pos
	334	—	×	×	—	—	—	—	—
	368	—	—	—	×	×	—	—	—
	406	—	—	—	×	×	—	—	—
ND4L	18	×	—	—	×	—	—	—	—
ND5	11	—	×	—	—		—	*Ra*	pos
	18	×	×	—	—		—	*pK′*	neg
	30	—	×	—	×		—	*pK′*	neg
	80	×	×	×	—		—	*Ra*	pos
	175	—	×	—	—		—	*pK′*	neg
	517	—	—	—	—		B (9)	*Ra*	pos
	537	—	—	—	—		B (9)	*Hp*	neg
	545	×	×	—	—		—	—	—
	546	—	—	—	—		B (9)	*pK′*	neg

		DATAMONKEY selection tests					PAML	TreeSAAP	
	Site	SLAC	REL	FEL	MEME	FUBAR	Branch-sites	Property	Direction
ND6	91	×	—	×	—		—	pK'	neg
	124	—	—	—	×		—	R_F pK'	neg pos
	135	—	—	—	—		B (9)	*Br*	neg

Note: only sites suggested to be under selection by two or more methods are shown. List of amino acid replacements fixed within mtDNA lineages of fruit fly species (for species names, see Table 1). TreeSAAP = physicochemical property identified by TreeSAAP as under radical positive evolution.

Abbreviations: *Br* = buriedness, *p* = polarity, R_F = chromatographic index, pK' = equilibrium constant (ionization of COOH), *Ra* = solvent accessible reduction ratio, *h* = hydropathy, *Hp* = surrounding hydrophobicity, *pHi* = isoelectric point, *P* = turn tendencies, E_t = total non-bonded energy.

The branch of the phylogenetic tree for which positive selection was suggested is indicated for the branch-site model of CODEML (Yang *et al.*, 2005; Zhang *et al.*, 2005). Fifty-one codons in 11 genes (no sites under positive selection in COX2 and CYTB genes) were shown to have evolved under recurrent positive selection by REL, FEL, SLAC, MEME, and FUBAR; 21 sites were concordant for two of the methods (Table 2). Twelve codons were shown to be under positive selection by the branch-site model analysis of CODEML. The TreeSAAP method identified numerous codons (335) as possibly evolving under positive selection. Of the properties for which radical changes were suggested, 'solvent accessible reduction ratio' and 'equilibrium constant (ionization of COOH)' were most represented. Thirty-seven codons (Table 2) were suggested to be under positive selection by more than one method. The branch-site REL method suggested that three branches were under episodic diversifying selection and that most branches have evolved neutrally (Fig. 1).

Association of genetic variability with climate and host range

Synonymous and non-synonymous changes, as well as the number of positively selected sites, were strongly and negatively correlated with mean diurnal range (BIO2) (Table 3). Moreover, Kruskal-Wallis and ANOVA analyses indicated a significant difference among the three host categories in all four genetic variables (Table 3). The ANOVA test between host and positively selected sites was supported by a Tukey *post-hoc* test, indicating significant differences between oligophagous and polyphagous groups ($P = 0.021$).

Multiple regression analysis showed that climatic variables and host type explained a large proportion of the variation recorded for the synonymous changes ($R^2 = 0.645$; $P = 0.039$) and the number of positively selected sites ($R^2 = 0.628$; $P = 0.048$) (Table 4). Of the climatic variables, only mean diurnal range (BIO2) was significantly associated with synonymous changes. However, host variation was significantly associated with synonymous changes and positively selected sites (Table 4). Notably, these results were confirmed by the multiple regression analysis of the transformed variables (see 3082Appendix.pdf). These analyses showed that climatic variables and host range are significantly associated with all four genetic variables. Similarly, mean diurnal range (BIO2) and host type are the only factors associated with genetic variables.

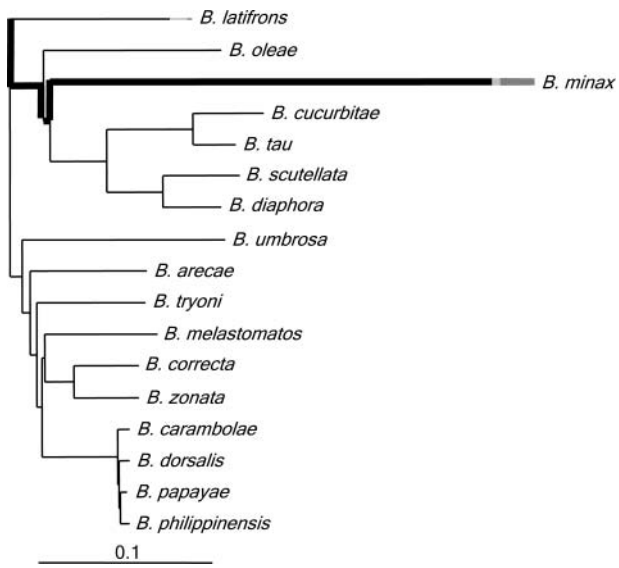

Fig. 1. Unrooted phylogeny of 17 *Bactrocera* fruit flies based on 13 protein-coding mtDNA sequences, estimated by the neighbour-joining method. Grey shading indicates strength of selection, with dark grey corresponding to $\omega > 5$, primary black to $\omega = 0$, and light grey to $\omega = 1$. The width of each colour component represents the proportion of sites in the corresponding class. Thicker branches have been classified as undergoing episodic diversifying selection by the sequential likelihood ratio test at corrected $P \leq 0.05$.

Table 3. Results of a Spearman correlation test between genotypes and bioclimatic variables

	Synonymous changes	Non-synonymous changes	dN/dS	Total number of positively selected sites
BIO2	−0.576 ($P = 0.016$)	−0.682 ($P = 0.003$)	−0.468 ($P = 0.058$)	**−0.589 ($P = 0.004$)**
BIO5	0.082 ($P = 0.756$)	0.002 ($P = 0.992$)	0.023 ($P = 0.929$)	0.103 ($P = 0.695$)
BIO12	−0.314 ($P = 0.220$)	−0.268 ($P = 0.298$)	−0.164 ($P = 0.529$)	−0.257 ($P = 0.319$)
BIO15	0.136 ($P = 0.603$)	0.077 ($P = 0.770$)	−0.167 ($P = 0.522$)	**−0.068 ($P = 0.578$)**
Host	$\chi^2 = 11.070$ ($P = 0.004$)	$\chi^2 = 11.588$ ($P = 0.003$)	$\chi^2 = 11.373$ ($P = 0.003$)	$F = 5.454$ ($P = 0.019$)

Note: Relationships between host and genotypes were assessed with a Kruskal-Wallis test except for the positively selected sites for which ANOVA was applied. **Bold** values: Person correlation test. For the correlations tests, the correlation coefficient is indicated first followed by the *P*-value. Significant values are shaded grey.

Relative rates of evolution and co-evolution of mtDNA coding genes

Figure 2, showing variation in the total number of substitutions from the ancestral sequences, reveals that some genes (ATP8, ND4L) are evolving more slowly than others (ND2, ND5, COX1). Most genes are evolving in a similar manner, i.e. a species that is evolving faster in one gene appears to be evolving faster in the other genes even if these genes do not evolve at the same rate. This tendency is confirmed by the significant pairwise positive correlation of the total substitutions in the studied mitochondrial genes (see 3082Appendix.pdf). However, the ND6 gene appears to evolve at a different rate from all

Table 4. Results of multiple regression analysis of each of the four genetic parameters and the five considered factors

	Synonymous changes $R^2 = 0.645$; $F = 0.628$; $P = 0.039$			Non-synonymous changes $R^2 = 0.499$; $F = 1.992$; $P = 0.165$			dN/dS $R^2 = 0.479$; $F = 1.837$; $P = 0.193$			Positively selected sites $R^2 = 0.628$; $F = 3.374$; $P = 0.048$		
Variable	β	t-value	P-value	β	t-value	P-value	β	t-value	P-value	β	t-value	P-value
BIO2	−0.505	−2.488	0.032	−0.444	−1.842	0.095	−0.383	−1.558	0.150	−0.462	−2.220	0.051
BIO5	−0.064	−0.287	0.780	0.181	0.687	0.508	0.241	0.898	0.390	−0.027	−0.121	0.906
BIO12	−0.222	−1.049	0.319	−0.406	−1.618	0.137	−0.349	−1.362	0.203	−0.286	−1.322	0.216
BIO15	0.138	0.580	0.574	−0.126	−0.448	0.664	−0.257	−0.890	0.394	−0.099	−0.406	0.694
Host	−0.474	−2.286	0.045	−0.303	−1.233	0.246	−0.368	−1.468	0.173	−0.493	−2.323	0.043

Note: Significant associations shaded grey.

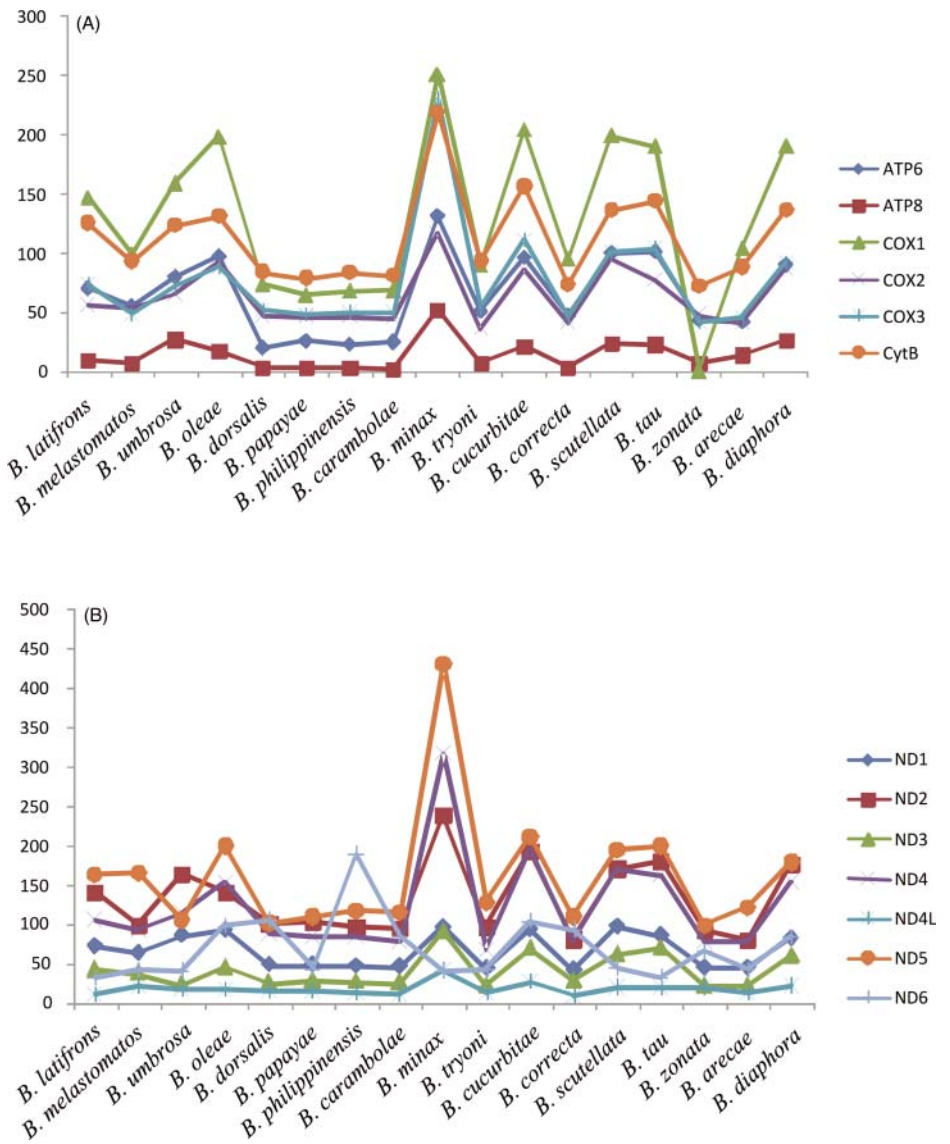

Fig. 2. Variation in the total number of substitutions in (A) genes of the cytochrome cb1 complex III, cytochrome c oxidase complex (COX1, COX2, and COX3), and ATPase complex V (ATP6 and ATP8) and (B) in genes of the NADH dehydrogenase complex I (ND1, ND2, ND3, ND4, ND4L, ND5, and ND6).

the others (as indicated by the absence of any correlation between this gene and the others). For example, in all mtDNA gene sequences, *B. philippinensis* is ancestral to all the other species (Fig. 2a, b), whereas in ND6 it has the most derived sequence. The opposite is true for *B. minax*, which is ancestral as inferred from substitutions in all genes except ND6, where this species is more derived.

DISCUSSION

Mitochondrial OXPHOS genes may play an important role in the adaptation of fruit flies to different habitats and different host plants. Here, we report the results of a mitochondrial genome analysis for positive selection in 17 fruit fly (*Bactrocera*) mitochondrial genomes using site- and branch-site based tests. We also assessed the relationships of substitution changes and number of positively selected sites with climate variation and host type. We found strong evidence of positive selection on 37 fixed amino acid replacements in genes involved in OXPHOS complexes – NADH dehydrogenase complex I (ND1, ND2, ND3, ND4, ND4L, ND5, ND6), cytochrome c oxidase complex IV (COX1 and COX3), and ATP synthase complex V (ATP6 and ATP8) – with most sites under strong purifying selection. Several of the amino acid substitutions were suggested by the TreeSAAP analysis to have an impact on the biochemical properties of the residues in question (Table 2). The functional impact of amino acid changes, specifically those under positive selection, is often assessed using a homology detection method to build 3D models and to predict ligand binding sites (see Melo-Ferreira *et al.*, 2014; Wang *et al.*, 2015) – this method failed to detect a functional effect. An accurate portrayal of the functional impact of amino acid changes appears only to be possible with integrated modelling of the cytonuclear protein–protein interactions.

The molecular signatures of positive selection in the mtDNA coding genes could be adaptively linked to environmental variation, as suggested by various studies using different organisms (da Fonseca *et al.*, 2008; Foote *et al.*, 2011; Garvin *et al.*, 2011; Melo-Ferreira *et al.*, 2014; Silva *et al.*, 2014). In our attempt to understand positive (diversifying) selection on mitochondrial (mt) OXPHOS genes in the context of geographical and environmental variables, we correlated genetic variables with climatic data. We found a significant effect of mean diurnal range in most of the genetic variables studied. Mean diurnal range is a complex variable that incorporates both maximum and minimum temperatures. Association of this factor with genetic variables confirms an effect of ambient temperature on mtDNA variability. Although we found a negative correlation between mean diurnal range and genetic changes, the results suggest the more homogeneous the climatic conditions (minimal difference between maximum and minimum temperature), the greater the genetic changes. Given the prominent role of mitochondria in cellular energy production, selective effects of ambient temperature on the mitochondrial genes encoding for the OXPHOS process are likely. In fruit flies, available data show that environmental conditions strongly affect these insects' metabolism. Being poikilothermic, the fruit flies are especially sensitive to heat (Jaworski and Hilszczański, 2013). For most insect pests, rising temperatures cause increasing energy levels, which, in turn, lead to greater mobility, higher feeding rates, higher reproduction rates, and lower mortality. In contrast, falling temperatures affect insect metabolism, reducing movement, feeding, and reproduction. Moreover, studies on the effects of heat on insect metabolism demonstrate some adaptability to thermally challenging environments where changes in behaviour and development have been documented as resulting from heat treatments (Hendrichs and Hendrichs, 1990; Cayol, 1996).

Host range also had a significant effect on the various genetic variables of the *Bactrocera* fruit flies, indicating that adaptation to large numbers of hosts includes not only morphological and physiological features but also genetic changes. Notably, the branch-site methods implemented in PAML and DATAMONKEY (Table 1, Fig. 1) suggested that only branches with oligophagous species are under positive selection. That result could indicate that amino acid changes in mtDNA are shaped by the specialization process to limit the

number of host plants of fruit flies. Indeed, Nosil (2002) suggested that several factors might promote specialization, including genetically based trade-offs in performance among different habitats, competition for resources, resistance to predators, high costs of information processing, mate-finding, and energy costs related to searching for suitable habitat. These factors most likely depend on cell signalling and energy production, functions often assumed by mitochondria. Moreover, Drew and Hancock (2001) suggested that host specialization is the driving force behind speciation in different insect species, including fruit flies. Finally, the composition and dynamics of the insect community that interacts with a plant are affected by several plant characteristics, including chemistry, physiology, and morphology (Stam *et al.*, 2014), all of which have a genetic basis. Therefore, the genotype of a plant and thus its expressed phenotype might influence insect populations that interact with the plant and shape their composition (Whitham *et al.*, 2012). Such co-evolution between insects and their host plants allows proliferation of better-adapted insects both morphologically and genetically. Such adaptation requires high species diversity, as in the case of fruit flies. Thus, with their role in energy production, which is essential for a successful invasion of host plants, the OXPHOS genes are expected also to play a central role in the adaptation of fruit fly species.

Surprisingly, a negative correlation between host range and the genetic variables was observed, indicating that an increase of substitutions from the ancestor is concomitant with a process of specialization (from monophagous to polyphagous species). In such a case, the ancestral state would be a polyphagous species that became more and more specialized to a limited number of species during its evolution.

Finally, in the non-recombining mitochondrial molecule (Ballard and Whitlock, 2004), different rates of evolution have been suggested to occur (e.g. Marshall *et al.*, 2009; Whitehead, 2009; Jacobsen *et al.*, 2016). We observed such a situation in the mitochondrial genomes we analysed in this study. We also demonstrated that evolutionary rates vary not only between different mitochondrial genes but also within the same gene in different species. Such findings might indicate selective pressure in co-evolving proteins promoting a specific relationship between mitochondrial proteins (i.e. an association between ancestral and derived proteins or vice versa). This is observed mainly for ND6, where the correlation test indicated that this gene is evolving independently of the others. Such results could explain discrepancies observed in different phylogenetic studies of a group of species using different mitochondrial genes.

REFERENCES

Ballard, J.W. and Melvin, R.G. 2010. Linking the mitochondrial genotype to the organismal phenotype. *Mol. Ecol.*, **8**: 1523–1539.

Ballard, J.W.O. and Whitlock, M.C. 2004. The incomplete natural history of mitochondria. *Mol. Ecol.*, **13**: 729–744.

Balloux, F., Handley, L.L., Jombart, T., Liu, H. and Manica, A. 2009. Climate shaped the worldwide distribution of human mitochondrial DNA sequence variation. *Proc. R. Soc. Lond. B: Biol. Sci.*, **276**: 3447–3455.

Cayol, J.P. 1996. Box thorn, key early season host of the Mediterranean fruit fly. *Int. J. Pest Manage.*, **42**: 325–329.

Choudhary, J.S., Naaz, N., Prabhakar, C.S., Rao, M.S. and Das, B. 2015. The mitochondrial genome of the peach fruit fly, *Bactrocera zonata* (Saunders) (Diptera: Tephritidae): complete DNA sequence, genome organization, and phylogenetic analysis with other tephritids using next generation DNA sequencing. *Gene*, **569**: 191–202.

da Fonseca, R.R., Johnson, W.E., O'Brien, S.J., Ramos, M.J. and Antunes, A. 2008. The adaptive evolution of the mammalian mitochondrial genome. *BMC Genomics*, **9**: 119 [doi: 10.1186/1471-2164-9-119].

Drew, R.A.I. and Hancock, D.L. 2001. Phylogeny of the tribe Dacini (Dacinae) based on morphological, distributional and biological data. In *Fruit Flies (Tephritidae): Phylogeny and Evolution of Behaviour* (M. Aluja and A. Norrbom, eds.), pp. 491–504. Boca Raton, FL: CRC Press.

Foote, A.D., Morin, P.A., Durban, J.W., Pitman, R.L., Wade, P., Willerslev, E. *et al*. 2011. Positive selection on the killer whale mitogenome. *Biol. Lett.*, **7**: 116–118.

Funk, D.J. 2010. Does strong selection promote host specialisation and ecological speciation in insect herbivores? Evidence from *Neochlamisus* leaf beetles. *Ecol. Entomol.*, **35**: 41–53.

Garvin, M.R., Bielawski, J.P. and Gharrett, A.J. 2011. Positive Darwinian selection in the piston that powers proton pumps in complex I of the mitochondria of Pacific salmon. *PLoS One*, **6**: e24127 [doi: 10.1371/journal.pone.0024127].

Garvin, M.R., Bielawski, J.P., Sazanov, L.A. and Gharrett, A.J. 2014. Review and meta-analysis of natural selection in mitochondrial complex I in metazoans. *J. Zool. Syst. Evol. Res.*, **53**: 1–17.

Hendrichs, J. and Hendrichs, M.A. 1990. Mediterranean fruit fly (Diptera: Tephritidae) in nature: location and diel pattern of feeding and other activities on fruiting and nonfruiting host and nonhosts. *Ann. Entomol. Soc. Am.*, **83**: 632–641.

Ingman, M. and Gyllensten, U. 2007. Rate variation between mitochondrial domains and adaptive evolution in humans. *Human Mol. Genet.*, **16**: 2281–2287.

Jacobsen, M.W., da Fonseca, R.R., Bernatchez, L. and Hansen, M.M. 2016. Comparative analysis of complete mitochondrial genomes suggests that relaxed purifying selection is driving high nonsynonymous evolutionary rate of the NADH2 gene in whitefish (*Coregonus* ssp.). *Mol. Phylogenet. Evol.*, **95**: 161–170.

Jaworski, T. and Hilszczański, J. 2013. The effect of temperature and humidity changes on insects' development and their impact on forest ecosystems in the context of expected climate change. *Forest Res. Papers*, **74**: 345–355.

Kelley, S.T., Farrell, B.D. and Mitton, J.B. 2000. Effects of specialization on genetic differentiation in sister species of bark beetles. *Heredity*, **84**: 218–227.

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

Marshall, H.D., Coulson, M.W. and Carr, S.M. 2009. Near neutrality, rate heterogeneity, and linkage govern mitochondrial genome evolution in Atlantic cod (*Gadus morhua*) and other gadine fish. *Mol. Biol. Evol.*, **26**: 579–589.

McClellan, D.A. and Ellison, D.D. 2010. Assessing and improving the accuracy of detecting protein adaptation with the TreeSAAP analytical software. *Int. J. Bioinform. Res. Appl.*, **6**: 120–133.

Melo-Ferreira, J., Vilela, J., Fonseca, M.M., da Fonseca, R.R., Boursot, P. and Alves, P.C. 2014. The elusive nature of adaptive mitochondrial DNA evolution of an arctic lineage prone to frequent introgression. *Genome Biol. Evol.*, **6**: 886–896.

Mishmar D., Ruiz-Pesini, E., Golik, P., Macaulay, V., Clark, A.G., Hosseini, S. *et al*. 2003. Natural selection shaped regional mtDNA variation in humans. *Proc. Natl. Acad. Sci. USA*, **100**: 171–176.

Morales, H.E., Pavlova, A., Joseph, L. and Sunnucks, P. 2015. Positive and purifying selection in mitochondrial genomes of a bird with mitonuclear discordance. *Mol. Ecol.*, **24**: 2820–2837.

Murrell, B., Wertheim, J.O., Moola, S., Weighill, T., Scheffle, K. and Kosakovsky, P.S.L. 2012. Detecting individual sites subject to episodic diversifying selection. *PLoS Genet.*, **8**: e1002764 [doi: 10.1371/journal.pgen.1002764].

Murrell, B., Moola, S., Mabona, A.A., Weighill, T., Sheward, D., Kosakovsky, P.S.L. *et al*. 2013. FUBAR: a Fast, Unconstrained Bayesian AppRoximation for inferring selection. *Mol. Biol. Evol.*, **30**: 1196–1205.

Nachman, M.W., Brown, W.M., Stoneking, M. and Aquadro, C.F. 1996. Nonneutral mitochondrial DNA variation in humans and chimpanzees. *Genetics*, **142**: 953–963.

Nardi, F., Carapelli, A., Dallai, R. and Frati, F. 2003. The mitochondrial genome of the olive fly *Bactrocera oleae*: two haplotypes from distant geographical locations. *Insect Mol. Biol.*, **12**: 605–611.

Nardi, F., Carapelli, A., Boore, J.L, Roderick, G.K, Dallai, R. and Frati, F. 2010. Domestication of the olive fly through a multiregional host shift to cultivated olives: comparative dating using complete mitochondrial genome. *Mol. Phylogenet. Evol.*, **57**: 678–686.

Nosil, P. 2002. Transition rates between specialization and generalization in phytophagous insects. *Evolution*, **56**: 1701–1706.

Pond, S.L.K. and Frost, S.D.W. 2005. Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Mol. Biol. Evol.*, **22**: 1208–1222.

Pond, S.L.K., Murell, B., Fourment, M., Frost, S.D., Delport, W. and Scheffler, K. 2011. A random effects branch-site model for detecting episodic diversifying selection. *Mol. Biol. Evol.*, **28**: 3033–3043.

Ruiz-Pesini, E., Mishmar, D., Brandon, M., Procaccio, V. and Wallace, D.C. 2004. Effects of purifying and adaptive selection on regional variation in human mtDNA. *Science*, **303**: 223–226.

Silva, G., Lima, F.P., Martel, P. and Castilho, R. 2014. Thermal adaptation and clinal mitochondrial DNA variation of European anchovy. *Proc. R. Soc. Lond. B: Biol. Sci.*, **281**: 20141093 [doi: 10.1098/rspb.2014.21093].

SPSS Inc. 2009. *PASW Statistics for Windows*, v.18.0. Chicago, IL: SPSS Inc.

Stam, J.M., Kroes, A., Li, Y., Gols, R., van Loon, J.J., Poelman, E.H. *et al*. 2014. Plant interactions with multiple insect herbivores: from community to genes. *Annu. Rev. Plant. Biol.*, **65**: 689–713.

Stewart, J.B., Freyer, C., Elson, J.L., Wredenberg, A., Cansu, Z., Trifunovic, A. *et al*. 2008. Strong purifying selection in transmission of mammalian mitochondrial DNA. *PLoS Biol.*, **6**: e10 [doi:10.1371/journal.pbio.0060010].

Tan, M., Zhang, R., Xiang, C. and Zhou, X. 2016. The complete mitochondrial genome of the pumpkin fruit fly, *Bactrocera tau* (Diptera: Tephritidae). *Mitochondrial DNA*, **27**: 2502–2503.

Taylor, R.W. and Turnbull, D.W. 2005. Mitochondrial DNA mutations in human disease. *Nat. Rev. Genet.*, **6**: 389–402.

Tuppen, H.A., Blakely, E.L., Turnbull, D.M. and Taylor, R.W. 2010. Mitochondrial DNA mutations and human disease. *Biochim. Biophys. Acta*, **1797**: 113–128.

Vargas, R.I., Piñero, J.C. and Leblanc, L. 2015. An overview of pest species of *Bactrocera* fruit flies (Diptera: Tephritidae) and the integration of biopesticides with other biological approaches for their management with a focus on the Pacific region. *Insects*, **6**: 297–318.

Wang, J., Yu, X., Hu, B., Zheng, J., Xiao, W., Hao, Y. *et al*. 2015. Physicochemical evolution and molecular adaptation of the Cetacean osmoregulation-related gene UT-A2 and implications for functional studies. *Sci. Rep.*, **5**: 8795 [doi: 10.1038/srep08795].

Whitehead, A. 2009. Comparative mitochondrial genomics within and among species of killifish. *BMC Evol. Biol.*, **9**: 11 [doi: 10.1186/1471–2148–9–11].

Whitham, T.G., Gehring, C.A., Lamit, L.J., Wojtowicz, T., Evans, L.M., Keith, A.R. *et al*. 2012. Community specificity: life and afterlife effects of genes. *Trends Plant Sci.*, **17**: 271–281.

Woolley, S., Johnson, J., Smith, M.J., Crandall, K.A. and McClellan, D.A. 2003. TreeSAAP: selection on amino acid properties using phylogenetic trees. *Bioinformatics*, **19**: 671–672.

Wu, P.-F., Roques, A., Xiong, Z.-P., Xu, L., Huang, Y.-Y. and Pan, Y.-Z. 2013. The complete mitochondrial genome of the melon fly *Bactrocera cucurbitae* (Diptera: Tephritidae). *Mitochondrial DNA*, **24**: 6–7.

Yang, Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.*, **24**: 1586–1591.

Yang, Z., Nielsen, R. and Goldman, N. 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics*, **155**: 431–449.

Yang, Z., Wong, W.S. and Nielsen, R. 2005. Bayes empirical Bayes inference of amino acid sites under positive selection. *Mol. Biol. Evol.*, **22**: 1107–1118.

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.

Yong, H.-S., Song, S.-L., Lim, P.-E., Eamsobhana, P. and Suana, I.W. 2016. Complete mitochondrial genome of three *Bactrocera* fruit flies of subgenus *Bactrocera* (Diptera: Tephritidae) and their phylogenetic implications. *PLoS One*, **11**: e0148201 [doi: 10.1371/journal.pone.0148201].

Zhang, B., Nardi, F., Hull-Sanders, H., Wan, X. and Liu, Y. 2014. The complete nucleotide sequence of the mitochondrial genome of *Bactrocera minax* (Diptera: Tephritidae). *PLoS One*, **9**: e100558 [doi: 10.1371/journal.pone.0100558].

Zhang, J., Nielsen, R. and Yang, Z. 2005. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Mol. Biol. Evol.*, **22**: 2472–2479.

Zhang, K.J., Liu, L., Rong, X., Zhang, G.-H., Liu, H. and Liu, Y.-H. 2015. The complete mitochondrial genome of *Bactrocera diaphora* (Diptera: Tephritidae). *Mitochondrial DNA*, **27**: 4314–4315.