Niche evolution and thermal adaptation in the temperate species *Drosophila americana*

N. SILLERO*, M. REIS†, C. P. VIEIRA‡, J. VIEIRA† & R. MORALES-HOJAS‡§

*Centro de Investigação em Ciências Geo-Espaciais (CICGE), Observatório Astronómico Prof. Manuel de Barros, Porto, Portugal
†Laboratório de Evolução Molecular, Instituto de Biología Molecular e Celular (IBMC), Universidade do Porto, Porto, Portugal
‡Laboratório de Evolução de Sistemas Biológicos, Instituto de Biologia Molecular e Celular (IBMC), Universidade do Porto, Porto, Portugal
§Genetics and Genomics group, The Pirbright Institute, Compton, Berkshire, UK

**Keywords:**
demographic history;
developmental time;
*Drosophila*;
ecological niche modelling;
locomotor activity;
niche conservatism;
phylogeography;
thermal adaptation.

**Abstract**

The study of ecological niche evolution is fundamental for understanding how the environment influences species’ geographical distributions and their adaptation to divergent environments. Here, we present a study of the ecological niche, demographic history and thermal performance (locomotor activity, developmental time and fertility/viability) of the temperate species *Drosophila americana* and its two chromosomal forms. Temperature is the environmental factor that contributes most to the species’ and chromosomal forms’ ecological niches, although precipitation is also important in the model of the southern populations. The past distribution model of the species predicts a drastic reduction in the suitable area for the distribution of the species during the last glacial maximum (LGM), suggesting a strong bottleneck. However, DNA analyses did not detect a bottleneck signature during the LGM. These contrasting results could indicate that *D. americana* niche preference evolves with environmental change, and thus, there is no evidence to support niche conservatism in this species. Thermal performance experiments show no difference in the locomotor activity across a temperature range of 15 to 38 °C between flies from the north and the south of its distribution. However, we found significant differences in developmental time and fertility/viability between the two chromosomal forms at the model’s optimal temperatures for the two forms. However, results do not indicate that they perform better for the traits studied here in their respective optimal niche temperatures. This suggests that behaviour plays an important role in thermoregulation, supporting the capacity of this species to adapt to different climatic conditions across its latitudinal distribution.

**Introduction**

The geographical distributions of species, in particular those of ectotherms, are partially determined by environmental variables such as temperature (Cossins & Bowler, 1987). Thus, the geographical range of any given species will be suggestive of their physiological tolerance (Hutchinson, 1957), and organisms respond to different environmental conditions by adapting genetically or by adjusting their phenotype (phenotypic plasticity) (Angilletta, 2009). In drosophilids, species’ distributions reflect their upper thermal tolerance and there is a strong phylogenetic inertia in heat tolerance (Kellermann *et al.*, 2012b). This suggests that the distribution of species is more limited by the lineage to which they belong than by local adaptation capacity. This may be true at the macroevolutionary level if climatic niche preference was constant over time; however, many *Drosophila* species have been able to invade different climatic regions (Parsons, 1983), demonstrating a capacity to adapt to new environments. For example, van Heerwaarden & Sgro (2013) have shown...
that *D. simulans* has the ability to respond to selection pressure towards increased upper thermal limits. This species also shows a plastic response to the diverse local conditions it encounters across its distributional range, performing optimally at a wide range of temperatures (Austin & Moehring, 2013). Contrary to this plasticity, individuals of *D. melanogaster* show adaptive response to the local conditions and perform better when subject to their local environment (Trotta et al., 2006). Thus, closely related species exhibit different responses to environmental change. In any case, there is evidence for a complex interplay between plastic, behavioural responses and the adaptive capacity of species (Calabria et al., 2012; Castaño et al., 2013).

The study of ecological niches (Hutchinson, 1957; Sobreron & Nakamura, 2009; Sillero, 2011) can be used to understand the effect of environmental change on species’ distributions. These methods have been successfully used to accurately model the present, past and future distributions of many species (e.g. Waltari et al., 2007; Cordellier & Penninger, 2009; Jakob et al., 2010; Sillero & Carretero, 2013). To predict the future and past species’ distributions, the current niche model of the species is extrapolated into the past and future climate reconstructions. However, this method relies on the niche conservatism assumption that species retain their ancestral ecological niche preferences. Combining ecological niche modelling (ENM) with inferences on the demographic history and phylogeography of the species provides a test of the assumption of niche conservatism. Recent phylogenetic analyses of temperature and desiccation tolerance in *Drosophila* suggest that this group of organisms show a conservative response to climate change (Kellermann et al., 2012a,b), which could be taken as evidence for phylogenetic niche conservatism. On the other hand, these studies did not specifically test this hypothesis using ENM and phylogeographical analyses for any of the *Drosophila* species included.

*Drosophila americana* belongs to the *D. virilis* group and has an estimated age of 1–1.6 million years (my) (Morales-Hojas et al., 2008, 2011). It is native to the USA, inhabiting the Central and Eastern regions from Texas-Florida to Montana-Maine. It comprises two chromosomal forms, the northern form (*D. a. americana*) being characterized by a fusion of the chromosomes X and 4, whereas the southern form (*D. a. texana*) has these chromosomes unfused. Despite this difference in karyotype between populations, there is no evidence for significant genetic differentiation or population structure within the species, except at the base of the X chromosome (Vieira et al., 2006; Morales-Hojas et al., 2008; Fonseca et al., 2013). The frequency of this rearrangement declines towards the south, where it is almost absent (McAllister, 2002; McAllister et al., 2008). This clinal distribution of the frequency of the chromosomal fusion indicates that natural selection is favouring one rearrangement over the other along a latitudinal transect. However, the shallow gradient of the cline, indicating a gradual change in frequency of the fusion, also shows that the selective force is not very strong. Despite this evidence, the factor responsible for the maintenance of the rearrangement gradient is unknown. The frequency of the fusion has been correlated with minimum January temperatures (McAllister et al., 2008), but this does not necessarily imply a causative relationship.

The main goal of the work presented here was to understand how a temperate species of *Drosophila* responds to different environmental conditions across its distribution. For this, we have first established what environmental factors constitute the niche and determined the distribution of *D. americana* using ecological niche modelling. We have also investigated the differences in niche preference between strains from the south and the north of the distribution and their performances under different thermal conditions to understand whether they are differentially adapted to different local temperature regimes. We have also analysed the evolution of niche preference of the species by extrapolating the present species model to the last glacial maximum (LGM: ~21 000 years ago) and the last interglacial (LIG: ~120 000–140 000 years ago) conditions in combination with a demographic analysis using DNA sequences to test the niche conservatism hypothesis.

**Materials and methods**

### Environmental niche modelling

We used correlative methods of ecological niche modelling (Guisan & Zimmermann, 2000) requiring georeferenced presence-only data and GIS-based environmental data. The location data for *D. americana* were obtained from our collection sites (available at [http://evolution.ibmc.up.pt](http://evolution.ibmc.up.pt); see also Reis et al., 2008) and those made available by Bryant McAllister ([http://wwwbiology.uiowa.edu/mcallister/bfm_flies.html](http://wwwbiology.uiowa.edu/mcallister/bfm_flies.html)) (Appendix S1; Fig. 1a,b). We used 19 climatic variables from the WorldClim series (Hijmans et al., 2005; www.worldclim.org/): BIO1 = Annual Mean Temperature; BIO2 = Mean Diurnal Range [Mean of monthly (max temp–min temp)]; BIO3 = Isothermality (BIO2/BIO7); BIO4 = Temperature Seasonality (standard deviation ×100); BIO5 = Max Temperature of Warmest Month; BIO6 = Min Temperature of Coldest Month; BIO7 = Temperature Annual Range (BIO5–BIO6); BIO8 = Mean Temperature of Wettest Quarter; BIO9 = Mean Temperature of Driest Quarter; BIO10 = Mean Temperature of Warmest Quarter; BIO11 = Mean Temperature of Coldest Quarter; BIO12 = Annual Precipitation; BIO13 = Precipitation of Wettest Month; BIO14 = Precipitation of Driest Month; BIO15 = Precipitation Seasonality (Coefficient of Variation); BIO16 = Precipitation of Wettest Quarter; BIO17 = Precipitation of Driest Quarter; BIO18 = Precipitation of Warmest Quarter, and;
BIO19 = Precipitation of Coldest Quarter. The spatial resolution of the variables was 30 arc-seconds (approximately 1 km²), except for the LGM variables, with 2.5 arc-minutes (approximately 5 km²). We calculated a Pearson correlation test to remove the bioclimatic variables that were most correlated (Appendix S2). Seven showed correlation values equal or lower than 0.75 (after rounding to two decimal places) and were included in the analyses. Models using alternative variables related to minimum temperatures were also estimated. We used three past climate scenarios in the analyses: one for the last interglacial (LIG: ~120 000–140 000 years BP; Otto-Bliesner et al., 2006) and two (CCSM and MIROC) for the last glacial maximum (LGM: ~21 000 years BP).

We applied the maximum entropy (Maxent) method to estimate the following models of the realized ecological niche (sensu Sillero, 2011):

1 The model of the species in current climatic conditions of North America and projected to past climatic conditions.

2 The models of the two chromosomal forms in current climatic conditions. Populations were included in the americana or texana models when they showed ≥ 10% homozygous individuals for the fused or unfused X and 4 chromosomes, respectively (Appendix S1).

3 The models of D. americana and its chromosomal forms using current climatic conditions, with an alternative set of variables (removed from model 1, above) and related to minimum temperatures. Models were estimated using the software Maxent 3.3.3k (Phillips et al., 2004, 2006). This method is well suited to noisy or sparse information and is capable of dealing with continuous and categorical variables at the same time. Furthermore, it has been shown to outperform other methods, especially when samples sizes are low (Elith et al., 2006; Hernandez et al., 2006). We ran Maxent with autofeatures (for establishing automatically the type of relationships among variables), randomly selecting 75% of the presence records as training data and 25% as test data. Ten model replicates were run.
MAXENT then estimated the arithmetic mean with standard deviation of their projections through an iterative process (for a review on consensus models see Marmion et al., 2009). We tested the performance of MAXENT with receiver operated characteristics (ROC) plots. The area under the curve (AUC) of the ROC plot is a measure of the overall fit of the models (Liu et al., 2005). We selected AUC because it is independent of prevalence (the proportion of presence in relation to the total dataset size; VanDerWal et al., 2009), although recent critiques question the reliability of this measure (Lobo et al., 2008).

We used these models to identify the importance of each environmental variable for explaining the species’ distribution by estimating: (i) the permutation importance determined by randomly changing the variable values among the training points and measuring the corresponding decrease in training AUC; thus, a model will depend strongly on those variables with a high decrease in permutation importance; (ii) jack-knife analysis of the average AUC with training and test data; and (iii) average percentage contribution of each ecogeographical variable (EGV) to the models.

Additionally, we applied Ecological Niche Factor Analysis (ENFA, Hirzel et al., 2002) to the chromosomal forms of D. americana using the complete set of variables. ENFA models provide valuable ecological information about the realized niche of the species (Hirzel et al., 2002; Sillero et al., 2009; see below). ENFA was performed with BIOMAPPER 4.0 software (Division of Conservation Biology, University of Bern, Bern, http://www.unil.ch/biomapper). It compares the distributions of the environmental variables between the presence dataset (the species niche) and the whole study area, summarizing all the environmental variables into new uncorrelated factors with ecological meaning. Individual models, ranging from 0 to 100, were derived using the distance geometric mean algorithm (Hirzel & Arlettaz, 2003; Brotons et al., 2004). This algorithm makes no assumption on the shape of the species’ distribution and takes into account the density of observation points in an environmental context by computing the geometric mean of all observation points. We evaluated ENFA models with an area-adjusted frequency cross-validation process (Boyce et al., 2002). The species locations were randomly partitioned into k different sets of equal size, k times. Each time, k-1 partitions were used to compute the model, and the excluded partition was used for validation. Each model was reclassified into i bins, obtaining the area-adjusted frequency for each bin as the proportion between the validation points (Ni) and the total area map (Ai): Fi = Ni/Ai. The expected Fi is 1.0 for all bins if the model is completely random. Low values of habitat suitability index should have a low F (below 1.0) and high values a high F (above 1.0). The predictive power of the model, measured by a Spearman rank correlation, was considered greatest when all Fi had a similar value (Boyce et al., 2002).

Samples and gene sequencing

The demographic history of the species was estimated using partial sequences from 16 genes in chromosomes 2 and 3 (Appendix S3). Polymorphic chromosomal inversions can confound the inference about the demographic history of the species, because the genetic signature reflects the evolutionary history of the rearrangement rather than that of the species. These chromosomes show no genetic structure across the species range (Morales-Hojas et al., 2008) and are thus good candidates to study the demographic history of the species. Gene sequences for CG3223, CG9631, elony, ill, phCl and tra were available in GenBank (Table 1; Maside & Charlesworth, 2007; Morales-Hojas et al., 2008, 2009). Genes geko, CG7255, sprk79D, dpr10, nkd, mub, CG8398, CG9674, CG14820 and Pk61C were recently obtained by the authors. Details on the amplification conditions, primers and sequencing procedures can be seen in Fonseca et al. (2013). Sequences have been deposited in GenBank (Appendix S3). Sequences used in analyses are from the texana strains Lake Ashbaugh (LA) Arkansas, Lake Hurricane, Arkansas (H), Monroe in Louisiana (ML) and Lone Star in Texas, (LP); americana strains used were from Niobrara in Nebraska (NN) and Gary in Illinois (G).

Demography analysis

We tested for population expansion by estimating Tajima’s D and Fu’s F for each gene with DNASP v. 5.10

Table 1 Contribution and importance after permutating the data of the different climatic variables towards the distribution model of Drosophila americana (a) and when using cold-related alternative variables (b). In bold are the values of the variables with highest contributions to the model.

| Variables                        | Per cent contribution | Permutation importance |
|----------------------------------|-----------------------|------------------------|
| (a)                              |                       |                        |
| Mean temperature warmest quarter (BIO10) | 51.2 | 64.8 |
| Precipitation driest month (BIO14)   | 18.1 | 4.1 |
| Annual precipitation (BIO12)       | 13.1 | 5.3 |
| Temp. annual range (BIO7)          | 9.6  | 15.5 |
| Mean diurnal range (BIO2)          | 2.9  | 4.6 |
| Precipitation seasonality (BIO15)  | 2.9  | 3.7 |
| Mean temperature wettest quarter (BIO8) | 2.3 | 2.1 |
| (b)                              |                       |                        |
| Mean temperature coldest quarter (BIO11) | 49.3 | 56.9 |
| Mean diurnal range (BIO2)          | 22.3 | 29.6 |
| Precipitation seasonality (BIO15)  | 18.5 | 11.1 |
| Precipitation coldest quarter (BIO18) | 9.9 | 2.4 |
These tests for deviation from evolutionary neutrality, and a significant negative value is indicative of a population expansion, negative selection or recent positive selection. We also investigated the mismatch distribution of pairwise genetic differences for each locus using *DNASP* v. 5.10, comparing the observed distributions to a unimodal expectation under a model of recent population expansion. The fit of the data to an expansion model is determined by a non-significant value of the raggedness index \( r \), Harpending, 1994. Significance was evaluated with 1000 coalescent simulations assuming recombination.

We explored the population history of *D. americana* using Extended Bayesian Skyline Plots (EBSP) as implemented in *BEAST* 1.7.4 (Drummond et al., 2012). Multilocus analyses were run including all 16 genes with unlinked substitution models, unlinked clock models and unlinked trees settings. We first determined the models of evolution that best fitted our data for each gene with *IMODELTEST* (Darriba et al., 2012). We performed two independent runs of 500 million iterations each (sampling every 50000 iterations), implementing a strict clock with a lognormal prior for the rate; initial value of 0.0005, log(mean) of \(-5\), log(SD) of 2 and the offset set to 0. These settings resulted in a 95% confidence interval (CI) of the clock rate that ranged from 2.511E-4 to 0.181 substitutions per nucleotide. The age of the species has been estimated to be in the range of 1–1.6 mya (Morales-Hojas et al., 2008, 2011; Morales-Hojas & Vieira, 2012). A root height prior was set to lognormal with a log(mean) of 0.095 and a log(SD) of 0.35, which resulted in a mean root height of 1.082 my (0.681–1.649 95% CI). The two runs were combined using *LOGCOMBINER* v. 1.7.4 (Drummond et al., 2012), and the appropriateness of run lengths and burnin level were analysed with *TRACER* v. 1.5 (http://beast.bio.ed.ac.uk/software/tracer/).

**Thermal performance assay**

We used locomotor activity of individual adult flies at different temperatures (15, 18, 20, 22, 25, 28, 30, 35 and 38 °C) to approximate thermal performance. Locomotor activity was measured using a *Drosophila* Activity Monitor (DAM, TriKinetics, Waltham, MA, USA). DAM comprises 32 holding tubes (7 mm diameter) crossed by an infrared beam. Every time a fly crosses the beam, the activity was measured using a *DAM* holding tubes with food the afternoon before running the experiments. Assays were performed in batches of 32 individuals per day during 7 days; thus, four individuals per strain (16 individuals of each chromosomal form) were tested each day in the same block. Assays were run from 09.25 to 18.40 h to expose the individual flies to all temperatures within the same day. We measured the activity of flies for 30 min at each experimental temperature after exposing them for 5 min to that same temperature; flies were rested for 30 min at room temperature in between each treatment to reduce stress and acclimation. The testing order of temperature treatments was randomly chosen every day; measurement at 35 and 38 °C were left to last to avoid high levels of stress and possible death.

Statistical tests were run with pooled locomotor activity scores for individuals at each temperature according to the chromosomal form. Normality of the data was tested using the Shapiro–Wilk test, and the nonparametric Mann–Whitney U test was used to compare the locomotor activity of the chromosomal forms at different test temperatures. Statistical analyses were performed in \( r \).

**Developmental time and fertility/viability assays**

For these experiments, we used the same strains as for the thermal performance ones (see above). We established crosses between three virgin females and three males per vial at 25 °C under 12-h dark–12-h light cycles for each of the eight strains described above. After 7 days, the individuals were changed to new vials and females were allowed to lay eggs on the 8th day. All the individuals were then discarded, and the vials were transferred to a climate chamber set at 22.5 or 26.6 °C under 12-h dark–12-h light cycles. These temperatures were selected based on the *MAXENT* ecological niche models for *americana* and *texana* chromosomal forms. We recorded the developmental time (DT) as the time between egg laying and adult emergence of every male individual born in each flask. Data were obtained for at least 100 male progeny for each strain and temperature. We pooled all individual DT values at each temperature according to chromosomal form and used nonparametric tests as implemented in *SPSS* Statistics 19.0 (SPSS Inc., Chicago, IL, USA) to test for DT differences.

The fertility/viability for each strain was estimated as the ratio between the total number of male progeny born and the total number of crosses needed to obtain them, using each vial as a single data point. We per-
formed chi-square tests with one degree of freedom to compare fertility/viability estimates between and within chromosomal forms at the two experimental temperatures. The expected number of individuals, under the assumption of no differences between treatments and chromosomal forms, was obtained using information on the relative proportion of vials used and the total number of individuals involved in the comparison.

**Results**

**Environmental niche modelling of D. americana**

The **maxent** model of the species’ present distribution with the first set of variables (Table 1a) had an AUC value of 0.93, indicating a reliable model performance that closely corresponds to the known distribution of the species (Fig. 1a,b). The variable that contributed most to the niche model (51.2%) was BIO10, the mean temperature of the warmest quarter, followed by BIO14 (precipitation of the driest month 18.1%) and BIO12 (annual precipitation 13.1%) (Table 1a). Nevertheless, the model was mostly dependent on BIO10, which showed the highest drop in training AUC after reevaluation of the model using permuted data (64.8%). This variable was most highly correlated with the annual mean temperature (BIO1; Pearson’s correlation 0.92) and the maximum temperature of the warmest month (BIO5; Pearson’s correlation 0.96), although it also showed high correlations with the minimum temperature of the coldest month (BIO6; Pearson’s correlation 0.81) and the mean temperature of the coldest quarter (BIO11; Pearson’s correlation 0.83) (Appendix S2).

Given that temperature variables were highly correlated, the model was also run including alternative variables related to minimum temperatures (Table 1b). In this case, BIO11 (mean temperature of the coldest quarter) was the variable with the largest contribution to the model (49.3%) and with the largest drop in AUC when using permuted data (56.9%) (Table 1b). This result is not surprising given the correlation between BIO10 (included in the model above) and BIO11 (Appendix S2), but reinforces the result pointing to temperature as the most relevant variable for the species’ distribution model.

To further investigate the limiting climatic factors of the species’ distribution, we extracted the values that maximized the probability of occurrence from the response curves of variable BIO10 and BIO11 (Fig. 2). The probability of occurrence was maximized around 25 °C mean temperature during the warmest quarter, increasing abruptly from 15 °C and decreasing until 35 °C. In the case of BIO11, the probability of presence was maximized around 4 °C mean temperature during the coldest quarter, with an increase from −15 °C and a decrease to a minimum probability of presence around 25 °C.

**Environmental niche models, marginality and specialization of chromosomal forms**

The models using the first set of environmental variables had high AUC values (0.943 and 0.987 for americana and texana chromosomal forms, respectively; Table 2a,b). The predicted distribution of the X/4 fused
form was larger than that of the nonfused form (Fig. 3). Furthermore, MAXENT analyses also indicate that the niches of the two chromosomal forms differ in the relevance of precipitation. The mean temperature of the warmest quarter (BIO10) showed contribution values of 47.2% and 39.5% in the *americana* and *texana* chromosomal forms’ models, respectively (Table 2a,b). The permutation importance of this variable was high with drops of 63.9% and 47.7%, respectively (Table 2a,b). The other variables of relevance in the distribution models of the two forms were different. Thus, the variable that contributes most to the *texana* model (53.6%) was BIO14 (precipitation of driest month), with a permutation importance of 34%, whereas in *americana*, the annual precipitation (BIO12) had a contribution of 17.9% but with a low permutation importance (1.3%) (Table 2a,b). In general, most precipitation-related variables are correlated, making it difficult to unravel the relative importance of each. BIO12 and BIO14 have a Pearson correlation value of 0.71 (Appendix S2). However, the fact that their permutation test values were different for the models of each chromosomal form suggests that precipitation is more important to the *texana* than the *americana* model.

Results also showed differences between the two chromosomal forms’ models when run with the alternative cold-related climatic variables for (c) *americana* and (d) *texana* forms. In Table 2, the variables with the highest contribution for each chromosomal form.

**Table 2** Contribution and permutation importance of the different climatic variables towards the distribution model of (a) *americana* and (b) *texana* forms, and results with the alternative cold-related climatic variables for (c) *americana* and (d) *texana* forms. In bold, the variables with the highest contribution for each chromosomal form.

| Variables                                | Per cent contribution | Permutation importance | Per cent contribution | Permutation importance |
|-------------------------------------------|-----------------------|------------------------|-----------------------|------------------------|
| Mean temperature warmest quarter (BIO10) | 47.2                  | 63.9                   | 39.5                  | 47.7                   |
| Temperature annual range (BIO7)          | 20.6                  | 17.4                   | 1.2                   | 2.9                    |
| Annual precipitation (BIO12)             | 17.9                  | 1.3                    | 0.6                   | 2.3                    |
| Precipitation driest month (BIO14)       | 6.5                   | 6                      | 0.6                   | 2.3                    |
| Mean diurnal range (BIO2)                | 3.6                   | 7.2                    | 2.1                   | 1.8                    |
| Precipitation seasonality (BIO15)        | 2.8                   | 2.2                    | 1.4                   | 10.7                   |
| Mean temperature wettest quarter (BIO8)  | 1.6                   | 2                      | 1.6                   | 2                      |

| Variables                                | Per cent contribution | Permutation importance | Per cent contribution | Permutation importance |
|-------------------------------------------|-----------------------|------------------------|-----------------------|------------------------|
| Mean temperature coldest quarter (BIO11) | 69.6                  | 79.3                   | 36.5                  | 76.3                   |
| Mean diurnal range (BIO2)                | 21.7                  | 15.2                   | 4.4                   | 2.2                    |
| Precipitation seasonality (BIO15)        | 4.5                   | 2.4                    | 12                    | 14.3                   |
| Precipitation coldest quarter (BIO19)    | 4.2                   | 3.1                    | 47.1                  | 7.2                    |

Fig. 3 Prediction of present distribution range of (a) *americana* (X/4 fused form) and (b) *texana* (unfused form) estimated with MAXENT using the WorldClim dataset.
LIG was estimated from the distribution maps (no. shown). The area occupied by the species during the life of the available habitat for the species (map not shown). The southernmost regions of the USA during the LGM (~120 000 ybp) to have been 8441765.99 km². This area was reduced to 3463.24 km² during the LGM (MIROC) after which it has increased to the present estimated area of 4125639.86 km².

Demography of *D. americana*

Tajima’s D and Fu’s F values were all negative suggesting a population expansion (Table 3). Deviation from the neutral model was significant for eight genes in the case of Tajima’s D (sprk79D, pHCl, CG8398, CG9674, dpr10, nkd, pk61C and tra) and four for Fu’s F (pHCl, CG9674, dpr10 and mub). The interpretation of a population expansion is supported by the nonsignificant values of the raggedness index (r) observed in all genes except mub (Table 3).

Two runs of 500 million generations each were combined to obtain good ESS values (> 200) as estimated with traceml. Multilocus EBSP results show highest posterior probability for two changes in population size. The plot indicates that the population of *D. americana* has been steadily increasing since its origin, with two expansions occurring around the time of origin (~1 million years before present (mybp) to ~0.7 mybp and at ~300 000 ybp (Fig. 4).

**Thermal performance of the chromosomal forms**

Locomotor activity of the flies increased with temperature as shown in the performance plot (Fig. 5; raw values in Appendix S4). The activity remained low between 15 and 25 °C and started to increase at 28 °C reaching a maximum at 38 °C. In the present range of temperatures, the expected rapid decrease in activity after the maximum was not reached, suggesting that the thermotolerance of this species is higher. The two

| Gene    | D    | F    | r    |
|---------|------|------|------|
| CG3223  | -0.218 | -0.737 | 0.048 |
| CG9631  | -1.238 | -1.257 | 0.069 |
| Eb      | 0.049  | -1.457 | 0.049 |
| Mub     | -0.521 | -1.778 | 0.045 |
| nkd     | -1.562 | -6.364 | 0.047 |
| pHCl    | -1.615 | -4.982 | 0.056 |
| CG14820 | -0.304 | -3.741 | 0.028 |
| CG7255  | -0.785 | -3.331 | 0.066 |
| CG8398  | -0.854 | -3.361 | 0.076 |
| CG9674  | -1.569 | -5.607 | 0.041 |
| dpr10   | -1.322 | -6.277 | 0.045 |
| Geko    | -0.632 | -4.452 | 0.026 |
| Nkd     | -0.606 | -5.553 | 0.011 |
| PK61C   | -0.476 | -1.390 | 0.026 |
| Tra     | -0.736 | -21.717 | 0.014 |

© 2014 THE AUTHORS. J. EVOL. BIO. 27 (2014) 1549–1561

JOURNAL OF EVOLUTIONARY BIOLOGY PUBLISHED BY JOHN WILEY & SONS LTD ON BEHALF OF EUROPEAN SOCIETY FOR EVOLUTIONARY BIOLOGY
chromosomal forms showed no difference in locomotor activity at the various test temperatures [nonparametric Mann–Whitney U tests with correction for multiple testing (Bonferroni, Holm, Hochberg, Hommel, BH, BY and SGoF); Table 4].

Developmental time and fertility/viability of the chromosomal forms

There were significant differences in males’ DT within chromosomal forms at different temperatures and between chromosomal forms at the same temperature (Mann–Whitney U test \( P < 0.001 \); Table 5; number of individuals and number of vials needed to obtain it are shown in Appendix S5). DT was higher at 22.5 \(^\circ\)C and decreased on average 3.1 and 4.0 days at 26.6 \(^\circ\)C for \textit{americana} and \textit{texana} chromosomal forms, respectively (Table 5). Males from the \textit{americana} chromosomal form developed faster on average than \textit{texana} males. The differences in DT between chromosomal forms reached 1.2 days at 22.5 \(^\circ\)C and 0.4 at 26.6 \(^\circ\)C.

Statistically significant differences in fertility/viability were found within chromosomal forms at both experimental temperatures (22.5 \(^\circ\)C vs. 26.6 \(^\circ\)C; \( \chi^2 = 7.64, P < 0.01 \) and \( \chi^2 = 99.82, P < 0.001 \) for \textit{americana} and \textit{texana}, respectively; Table 6a). On average, a decrease in one degree corresponded to an increase of 4.2% and 19.5% in viability for the \textit{americana} and \textit{texana} chromosomal forms, respectively. A statistically significant difference in fertility/viability was also found when comparing \textit{americana} and \textit{texana} at 22.5 \(^\circ\)C (\( \chi^2 = 100.36, P < 0.001 \); Table 6b), but not at 26.6 \(^\circ\)C (\( \chi^2 = 1.69, P = 0.193 \); Table 6b).

**Discussion**

Temperature is commonly considered to be the main variable shaping the geographical distribution of
Table 6 Comparison of the fertility/viability estimates within chromosomal forms at different temperatures (a) and between chromosomal forms at the same temperature (b).

(a)  

| Chrom. | 22.5 °C | 26.6 °C |  |
|--------|--------|--------|---|
|        | N      | E      | T  | R  | N      | E      | T  | R  | Chi | P     |
| americana | 651   | 599.5  | 121 | 5.4 | 771   | 822.5  | 166 | 4.6 | 7.64 | < 0.01 |
| texana   | 872   | 692.3  | 97  | 9.0 | 427   | 606.7  | 85  | 5.0 | 99.82 | < 0.001 |

(b)  

| Temp. (°C) | americana | texana |  |
|------------|-----------|--------|---|
| 22.5       | N         | E      | T  | R  | Chi | P     |
|            | 651       | 845.3  | 121 | 5.4 | 872  | 677.7  | 97  | 9.0 | 100.36 | < 0.001 |
| 26.6       | 771       | 792.3  | 166 | 4.6 | 427  | 405.7  | 85  | 5.0 | 1.69 | > 0.193 |

Chrom., chromosomal form; Temp., temperature; Chi, value of \( \chi^2 \); P, Probability; N, number of males; E, expected number of males under the assumption of no differences in fertility/viability between the cells being compared; T, number of vials needed to get N flies; R is N/T. All significant values remain significant after Bonferroni correction for multiple testing.

*Drosophila* species. Although previous work has focused on populations from different latitudes assuming that temperature was the main factor affecting the individual fitness, there are no studies addressing the relevant environmental variables that drive the ecological niche of *Drosophila* in natural populations. In this study, we showed that temperature is the main limiting climatic variable of those included in the ecological niche model (ENM) of *D. americana*. This is also true when considering the two chromosomal forms (*X*/4 fusion, *americana* form and *X* free, *texana* form) separately, although precipitation also plays a major role in the ecological niche model of the southern, *texana* form. These results are in agreement with the previous study by McAllister et al. (2008), who found a correlation between the frequency of the *X*/4 chromosomal rearrangement and the minimum temperature of January. The ENMs revealed a difference of 4 °C between the *texana* and *americana* optimal model temperatures in the case of the warmest quarter (22.5 and 26.6 °C for *americana* and *texana*, respectively) and of 10 °C in the coldest quarter (−2 and 8 °C for *americana* and *texana*, respectively). The model temperature response curves largely overlap for the two forms, which is in agreement with the north–south clinal distribution of the *X*/4 chromosomal rearrangement. Overall, results indicate that the northern and southern populations of *D. americana* are subject to slightly different but overlapping environmental ranges. Nevertheless, the climatic factors that drive their ecological niches differ in their relevance. The southern form is more dependent on precipitation than the northern form, whose model is preferentially driven by temperature, suggesting a difference in their niche preference.

Populations can respond differently to the varying climatic conditions occurring in their ranges. For instance, there is evidence for the adaptation of *Drosophila* populations to different temperatures along latitudinal gradients and between tropical and temperate populations (Balanyà et al., 2003; Trotta et al., 2006; McAllister et al., 2008; Rezende et al., 2010; van Heerwaarden & Sgro, 2013). In addition, there is evidence that behavioural responses are of relevance when coping with temperature differences (Austin & Moehring, 2013; Castañeda et al., 2013). In the case of *D. americana*, it has been previously suggested that the alternative chromosomal rearrangements influence divergent overwintering strategies and may represent the early stages of ecological speciation (McAllister et al., 2008). However, the question of whether populations of this species are differentially adapting to their respective niches has, so far, not been investigated. Although ENMs showed different variables of varying importance, we observed no significant differentiation between forms in locomotor activity at the various temperatures tested. This suggests that optimal body temperature for locomotor performance does not differ between adult males of the two chromosomal forms, although longer exposure to the different temperatures could uncover as yet undetected differences. However, there is no evident tendency towards different preferred temperatures for each form after 35 min of exposure while there is a clear increase in activity with temperature. If 35 min were not sufficient to elicit a locomotor response, we would not expect to observe a tendency in the temperature preference gradient either. Results, thus, suggest that there is no evident differential adaptation in the locomotor activity optimal body temperature between the chromosomal forms, which could suggest a role for behavioural thermoregulation in adult flies for this trait.

In contrast, we found significant differences in the developmental time and fertility/viability between and within chromosomal forms at the optimum mean temperatures of the warmest quarter as inferred from their
Niche evolution and thermal adaptation in Drosophila americana 1559

d respectively. These results indicate that performance of D. americana related to these traits is clearly dependent on temperature, as in the case of other species, such as D. melanogaster and D. simulans (James et al., 1997; Petavy et al., 2001; Trotta et al., 2006). However, our results show that the populations are not optimally adapted in these traits to the average environmental temperatures they encounter in their distribution area. Thus, developmental time is significantly lower in both chromosomal forms at 26.6 °C than at 22.5 °C, and viability/fertility is significantly smaller at 26.6 °C than at 22.5 °C. If chromosomal forms were differentially adapted in these traits, the direction of change from one temperature to the other should be opposite in each chromosomal form. This is not observed, showing no evidence for differential adaptation to the local environment in the traits studied here. However, these results do not eliminate the possibility of thermal adaptation in other traits that we have not measured. In Drosophila, there is a trade-off between fecundity and survival that is mediated through body size and developmental time (Roff, 1981). Thus, given that the americana form develops faster and still has a similar fecundity at 26.6 °C when compared to the southern texana form, the americana form would be expected at higher frequencies in the southern populations than it is observed. This points towards a disadvantage with respect to the local texana form in some other traits not studied in the present work.

In the case of D. americana, the width of the locomotor activity curve is small, and developmental time (DT) as well as fertility/viability is significantly reduced within a narrow range of temperatures. This indicates a narrow thermal range in comparison with that observed in other species such as D. melanogaster and D. simulans (Petavy et al., 2001). The thermal range of D. americana is also in agreement with the observed values of marginality and specialization obtained with the ENFA analyses, which suggest a narrow niche breadth and marginal habitat in contrast to the distribution model’s predicted range of temperatures. This observation could be explained if phenotypic plasticity and/or behavioural thermoregulation at the microhabitat level were relevant in the thermal adaptation process of D. americana, and therefore in its latitudinal geographical distribution.

According to the realized niche models, the suitable distribution for D. americana during the last interglacial period [LIG; ~140 000 to ~120 000 years before present (ybp)] is predicted to have been twice the size of the current one. The models also predicted a drastic reduction in the suitable area for D. americana during the last glacial maximum (LGM) ~21 000 ybp, which suggest a strong bottleneck during that period followed by a recent (~21 000 ybp) expansion to the present day distribution, approximately 1190 times larger. However, there is no evidence from the DNA analyses for a bottleneck during the LGM, but rather demographic expansions of the population of D. americana around 1–0.7 mya and 300 000 ybp. The lack of support for a LGM bottleneck with post-glacial expansion is in agreement with previous studies that concluded that the species had a stable recent historical population size (Schäfer et al., 2006; Morales-Hojas et al., 2008). Other studies suggest a fairly large recent increase in population size within a time span of approximately 0.11 × 2Ne generations, corresponding to approximately 61 600 ybp (de Proce et al., 2012). This estimate relies, however, on the assumption of an effective population size that is 38% of that recently estimated for this species based on a much larger polymorphism dataset and the sequence of two genomes (Fonseca et al., 2013). Therefore, the expansion may have occurred as long as 79 200 ybp, under the assumption of five generations a year (as in de Proce et al., 2012), but it could be as old as 99 000 ybp under the assumption of eight generations a year, based on developmental time data for D. americana obtained in this study and that estimated by Fonseca et al. (2013). In any case, the demographic history of this species as estimated from molecular data is not detecting a bottleneck during the last glaciation (~21 000 ya). Contradictory results between species’ past distribution inferred from the ENMs and the demographic history inferred from the DNA-based analyses could be an indication of the absence of niche conservatism in this species. Under this scenario, the populations of D. americana would have been able to respond to different climatic conditions from those that comprise their present niches. This explanation would support the idea that niche preference of D. americana evolves as a response to environmental change through time, and behavioural thermoregulation probably plays a key role in this capacity to respond to the diverse climatic conditions it encounters. However, there are other possibilities that would explain the contradictory results between the niche models and the molecular data. If the genetic diversity of the species had recovered rapidly after the bottleneck, its molecular signature would be lost. Two possible scenarios would meet this condition: (i) there were more than one refuge for the species during the LGM, and after the ice retreated the genetically differentiated populations mixed rapidly, increasing the genetic diversity of the species; (ii) there was one single population, and its effective population size was sufficiently high to prevent a drastic reduction in the genetic diversity of the species. Results from the ENM using the MIROC model show at least three main isolated areas suitable for the species (Fig. 1c), which would support the first scenario over the latter. A third possible explanation, compatible with the previous two, relates to the assumption that the past and current land surfaces of North America are equal. As the glaciers increased in volume, the sea level dropped exposing new land surface. Thus, the real reduction in habitable area for the
species may have not been as drastic as predicted by the past projection of the model, and the species would have occupied areas now under water, maintaining a large effective population size throughout the LGM.

In conclusion, there is no evidence for local divergent adaptation in the traits studied here to the different environmental conditions observed along the species’ geographical range. However, we cannot reject the possibility of local adaptation along the latitudinal gradient occurring in other traits not included in the present study. The inconsistent results between the demographic molecular analyses and the predicted past species’ distribution models suggest that this species has a capacity to respond to different environmental conditions, supporting a role for behavioural thermal regulation. Nevertheless, there are some caveats in the molecular analyses and further sampling will be fundamental to fully comprehend the phylogeographical history of the species.

Acknowledgments

This work was funded by FEDER Funds through the Operational Competitiveness Programme – COMPETE and by national funds through FCT – Fundação para a Ciência e a Tecnologia under the projects FCOMP-01-0124-FEDER-037277 (PEst-C/SAU/LA0002/2013) and a PhD grant attributed by FCT with the reference SFRH/BD/61142/2009 to M.R. We would like to thank Robert Austin, C.J. & Moehring, A.J. 2013. Optimal temperature for the freshwater gastropod Radix balthica (Pulmonata, Basommatophora). Mol. Ecol. 18: 534–544.

Cossins, A.R. & Bowler, K. 1987. Temperature Biology of Animals. Chapman and Hall, London.

Darrá, D., Taboada, G.L., Doallo, R. & Posada, D. 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat. Methods 9: 772.

Drummond, A.J., Suchard, M.A., Xie, D. & Rambaut, A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29: 1969–1973.

Elith, J., Graham, C.H., Anderson, R.P., Dudík, M., Ferrier, S., Guisan, A. et al. 2006. Novel methods improve prediction of species’ distributions from occurrence data. Ecoigraphy 29: 129–151.

Fonseca, N.A., Morales-Hojas, R., Reis, M., Rocha, H., Vieira, C.P., Nolte, V. et al. 2013. Drosophila americana as a model species for comparative studies on the molecular basis of phenotypic variation. Genome Biol. Evol. 5: 661–679.

Guisan, A. & Zimmermann, N.E. 2000. Predictive habitat distribution models in ecology. Ecol. Model. 135: 147–186.

Harpending, H.C. 1994. Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. Hum. Biol. 66: 591–600.

van Heerwaarden, B. & Sgro, C.M. 2013. Multivariate analysis of adaptive capacity for upper thermal limits in Drosophila simulans. J. Evol. Biol. 26: 800–809.

Hernandez, P.A., Graham, C.H., Master, L.L. & Albert, D.L. 2006. The effect of sample size and species characteristics on performance of different species distribution modeling methods. Ecoigraphy 29: 773–785.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. 2005. Very high resolution interpolated climate surfaces for global land areas. Int. J. Climatol. 25: 1965–1978.

Hirzel, A.H. & Arlettaz, R. 2003. Modeling habitat suitability for complex species distributions by environmental-distance geometric mean. Environ. Manage. 32: 614–623.

Hirzel, A.H., Haußer, J., Chessel, D. & Perrin, N. 2002. Ecological-niche factor analysis: how to compute habitat-suitability maps without absence data? Ecology 83: 2027–2036.

Hutchinson, G.E. 1957. Population studies - animal ecology and demography - concluding remarks. Cold Spring Harb. Symp. Quant. Biol. 22: 415–427.

Jakob, S.S., Heibl, C., Rodder, D. & Blattner, F.R. 2010. Population demography influences climatic niche evolution: evidence from diploid American Hordeum species (Poaceae). Mol. Ecol. 19: 1423–1438.

James, A.C., Azevedo, R.B. & Partridge, L. 1997. Genetic and environmental responses to temperature of Drosophila melanogaster from a latitudinal cline. Genetics 146: 881–890.

Kellermann, V., Loe schcke, V., Hoffmann, A.A., Kristensen, T.N., Flogaard, C., David, J.R. et al. 2012a. Phylogenetic constraints in key functional traits behind species’ climatic niches: patterns of desiccation and cold resistance across 95 Drosophila species. Evolution 66: 3377–3389.

Kellermann, V., Overgaard, J., Hoffmann, A.A., Flogaard, C., Svenning, J.C. & Loeschcke, V. 2012b. Upper thermal limits of Drosophila are linked to species distributions and strongly constrained phylogenetically. Proc. Natl. Acad. Sci. USA 109: 16228–16233.

Librado, P. & Rozas, J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25: 1451–1452.
Liu, C.R., Berry, P.M., Dawson, T.P. & Pearson, R.G. 2005. Selecting thresholds of occurrence in the prediction of species distributions. *Ecography* **28**: 385–393.

Lobo, J.M., Jiménez-Valverde, A. & Real, R. 2008. AUC: a misleading measure of the performance of predictive distribution models. *Glob. Ecol. Biogeogr.* **17**: 145–151.

Marmion, M., Luoto, M., Heikkinnen, R.K. & Thuiller, W. 2009. The performance of state-of-the-art modelling techniques depends on geographical distribution of species. *Ecol. Model.* **220**: 3512–3520.

Maside, X. & Charlesworth, B. 2007. Patterns of molecular variation and evolution in *Drosophila americana* and its relatives. *Genetics* **176**: 2293–2305.

McAllister, B.F. 2002. Chromosomal and allelic variation in *Drosophila americana*: selective maintenance of a chromosomal cline. *Genome* **45**: 13–21.

McAllister, B.F., Sheeley, S.L., Mena, P.A., Evans, A.L. & Schlottner, C. 2008. Clinal distribution of a chromosomal rearrangement: a precursor to chromosomal speciation? *Evolution* **62**: 1852–1865.

Morales-Hojas, R. & Vieira, C.P. 2012. Phylogenetic patterns of geographical and ecological diversification in the subgenus *Drosophila*. *PLoS ONE* **7**: e99552.

Morales-Hojas, R., Vieira, C.P. & Vieira, J. 2008. Inferring the evolutionary history of *Drosophila americana* and *Drosophila novamexicana* using a multilocus approach and the influence of chromosomal rearrangements in single gene analyses. *Mol. Ecol.* **17**: 2910–2926.

Morales-Hojas, R., Vieira, C.P., Reis, M. & Vieira, J. 2009. Comparative analysis of five immunity-related genes reveals different levels of adaptive evolution in the *viridis* and *melanogaster* groups of *Drosophila*. *Heredity* **102**: 573–578.

Morales-Hojas, R., Reis, M., Vieira, C.P. & Vieira, J. 2011. Resolving the phylogenetic relationships and evolutionary history of the *Drosophila viridis* group using multilocus data. *Mol. Phylogen. Evol.* **60**: 249–258.

Otto-Bliesner, B.L., Marsha, S.J., Overpeck, J.T., Miller, G.H., Morales-Hojas, R., Vieira, C.P., Reis, M. & Vieira, J. 2008. Inferring the warmth and icefield retreat in the last interglaciation. *Science* **311**: 1751–1753.

Parsons, P.A. 1983. The *Evolutionary Biology of Colonizing Species*. Cambridge University Press, Cambridge.

Petavy, G., David, J.R., Gibert, P. & Moreteau, B. 2001. Viability and rate of development at different temperatures in *Drosophila*: a comparison of constant and alternating thermal regimes. *J. Therm. Biol.* **26**: 29–39.

Phillips, S.J., Dudik, M. & Schapire, R.E. 2004. A maximum entropy approach to species distribution modeling. In: Twenty-first Century Internations Conference on Machine Learning, pp. 655–662 Proceedings of the Twenty-first Century Internations Conference on Machine Learning.

Phillips, S.J., Anderson, R.P. & Schapire, R.E. 2006. Maximum entropy modeling of species geographic distributions. *Ecol. Model.* **190**: 231–259.

de Proce, S.M., Zeng, K., Betancourt, A.J. & Charlesworth, B. 2012. Selection on codon usage and base composition in *Drosophila americana*. *Biol. Lett.* **8**: 82–85.

Reis, M., Vieira, C.P., Morales-Hojas, R. & Vieira, J. 2008. An old *bilbo*-like non-LTR retroelement insertion provides insight into the relationship of species of the *viridis* group. *Gene* **425**: 48–55.

Rezende, E.L., Balanyà, J., Rodríguez-Trelles, F., Rego, C., Fragata, I., Matos, M. et al. 2010. Climate change and chromosomal inversions in *Drosophila subobscura*. *Clim. Res.* **43**: 103–114.

Roff, D.A. 1981. On being the right size. *Am. Nat.* **118**: 405–422.

Schäfer, M.A., Orsini, L., McAllister, B.F. & Schlottterer, C. 2006. Patterns of microsatellite variation through a transition zone of a chromosomal cline in *Drosophila americana*. *Heredity* **97**: 291–295.

Siliero, N. 2011. What does ecological modelling model? A proposed classification of ecological niche models based on their underlying methods. *Ecol. Model.* **222**: 1343–1346.

Siliero, N. & Carretero, M.A. 2013. Modelling the past and future distribution of contracting species. The Iberian lizard *Podarcis carbonellii* as a case study. *Zool. Antar.* **252**: 289–298.

Siliero, N., Brito, J.C., Skidmore, A.K. & Toxopeus, A.G. 2009. Biogeographical patterns derived from remote sensing variables: the amphibians and reptiles of the Iberian Peninsula. *Amphibia-Reptilia* **30**: 185–206.

Soberon, J. & Nakamura, M. 2009. Niches and distributional areas: concepts, methods, and assumptions. *Proc. Natl. Acad. Sci. USA* **106**: 19644–19650.

Trotta, V., Calboli, F.C., Ziosi, M., Guerra, D., Pazzoli, M.C., David, J.R. et al. 2006. Thermal plasticity in *Drosophila melanogaster*: a comparison of geographic populations. *BMC Evol. Biol.* **6**: 67.

VanDerWal, J., Shoo, L.P., Graham, C. & Williams, S.E. 2009. Selecting pseudo-absence data for presence-only distribution modeling: how far should you stray from what you know? *Ecol. Model.* **220**: 589–594.

Vieira, C.P., Almeida, A., Dias, J.D. & Vieira, J. 2006. On the location of the gene(s) harbouring the advantageous variant that maintains the X/Y fusion of *Drosophila americana*. *Genet. Res.* **87**: 163–174.

Wallari, E., Hijmans, R.J., Peterson, A.T., Nyári, Á.S., Perkins, S.L. & Guralnick, R.P. 2007. Locating Pleistocene refugia: comparing phylogeographic and ecological niche model predictions. *PLoS ONE* **2**: e563.

Supporting information

Additional Supporting Information may be found in the online version of this article:

**Appendix S1** Georeferenced locations for the species and chromosomal forms, and the chromosome frequencies for each population.

**Appendix S2** Pearson’s correlation results between WorldClim’s Bioclimatic variables.

**Appendix S3** Genes, chromosomal location and accession numbers of sequences.

**Appendix S4** Activity counts for individual flies.

**Appendix S5** Total number of males, developmental time details and details on the number of crosses.

Data deposited at Dryad: doi:10.5061/dryad.5ng71

Received 18 November 2013; revised 1 April 2014; accepted 2 April 2014