Temperature, rainfall and wind variables underlie environmental adaptation in natural populations of Drosophila melanogaster

María Bogaerts-márquez, Sara Guirao-rico, Mathieu Gautier, Josefa González

Received: 7 May 2020 | Revised: 16 December 2020 | Accepted: 18 December 2020

DOI: 10.1111/mec.15783

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Temperature, rainfall and wind variables underlie environmental adaptation in natural populations of *Drosophila melanogaster*

María Bogaerts-Márquez[1,2] | Sara Guirao-Rico[1,2] | Mathieu Gautier[3] | Josefa González[1,2]

[1]Institute of Evolutionary Biology (CSIC-Universitat Pompeu Fabra), Barcelona, Spain

[2]The European Drosophila Population Genomics Consortium (DrosEU), Université de Montpellier, Montpellier, France

[3]CBGP, INRA, CIRAD, IRD, Montpellier SupAgro, Université de Montpellier, Montpellier, France

Correspondence
Josefa González, Institute of Evolutionary Biology (CSIC-UPF) Pso Maritimo Barceloneta, 37-49. 08003 Barcelona, Spain.
Email: josefa.gonzalez@ibe.upf-csic.es

Funding information
European Commission, Grant/Award Number: H2020-ERC-2014-CoG-647900

Abstract

While several studies in a diverse set of species have shed light on the genes underlying adaptation, our knowledge on the selective pressures that explain the observed patterns lags behind. *Drosophila melanogaster* is a valuable organism to study environmental adaptation because this species originated in Southern Africa and has recently expanded worldwide, and also because it has a functionally well-annotated genome. In this study, we aimed to decipher which environmental variables are relevant for adaptation of *D. melanogaster* natural populations in Europe and North America. We analysed 36 whole-genome pool-seq samples of *D. melanogaster* natural populations collected in 20 European and 11 North American locations. We used the BayPass software to identify single nucleotide polymorphisms (SNPs) and transposable elements (TEs) showing signature of adaptive differentiation across populations, as well as significant associations with 59 environmental variables related to temperature, rainfall, evaporation, solar radiation, wind, daylight hours, and soil type. We found that in addition to temperature and rainfall, wind related variables are also relevant for *D. melanogaster* environmental adaptation. Interestingly, 23%–51% of the genes that showed significant associations with environmental variables were not found overly differentiated across populations. In addition to SNPs, we also identified 10 reference transposable element insertions associated with environmental variables. Our results showed that genome-environment association analysis can identify adaptive genetic variants that are undetected by population differentiation analysis while also allowing the identification of candidate environmental drivers of adaptation.

KEYWORDS
allele frequency, *Drosophila melanogaster*, genetic adaptation, genome-environment, transposable elements

1 | INTRODUCTION

Understanding how organisms adapt to different environments is a major goal in evolutionary biology (Hoban et al., 2016; Nelson et al., 2019). The genetic basis of adaptative traits has been studied in several organisms, such as lactase persistence (Tishkoff et al., 2007) and skin colour in humans (Norton et al., 2007), and dark colour in the peppered moth *Biston Betularia* (Van't Hof et al., 2016) among many others. Genome-wide studies aimed at elucidating the genetic basis of environmental adaptation have also been conducted in several species such as plants (Flood & Hancock, 2017), bacteria (Gorter et al., 2016) and *Drosophila* (Rech et al., 2019). However, knowledge on the specific environmental variables driving these adaptations lags behind.

In the past few years, the availability of whole genome sequences as well as the development of different analytical tools, have facilitated the performance of genome-environment association (GEA) analyses. GEA analyses are useful approaches to identify the genetic variants and the environmental factors that are involved in the adaptive processes. Combining outcomes from GEA analysis with classical population genome-wide selection scans, such as those based on differentiation statistics, may help to link the genetic variants underlying local adaptation with their environmental drivers (Ahrens et al., 2018; Hoban et al., 2016). These analyses have already been applied to several species including Plant, Chordata, and to a lesser extent Arthropoda, Mollusca, Cnidaria, Echinodermata, and Nematoda (Ahrens et al., 2018). However, there are still important limitations behind GEA analyses. One of the main drawbacks is the difficulty in distinguishing the patterns associated with demographic processes from those that are the consequence of selection (reviewed in Rellstab et al., 2015). A second limitation is related to the choice of environmental variables to include in the analysis. Prior selection of the most relevant variables for any particular GEA analysis is complicated since some previous knowledge about which variables may be relevant in the adaptation process is needed. Indeed, most environmental variables used in GEA studies are related to temperature and precipitation, while other variables such as solar radiation, daylight hours, evaporation, and wind, that could also play a role in adaptation are not widely used. Solar radiation, and more specifically UV-B radiation, could be relevant as DNA damage responses are known to play a role in adaptation of several species such as birds, insects or fungi (Körner, 2007; Svetec et al., 2016; Wu et al., 2019; Zhou et al., 2020). Daylight hours is related to the circadian rhythm, which for example is known to play a role in *Drosophila* behavioural adaptation to high latitudes (Helfrich-Förster et al., 2020). Evaporation is involved in organism thermoregulation and response to desiccation stress (Ferveur et al., 2018; Rajpurohit et al., 2018; Smit et al., 2018). Finally, wind direction is involved in plant adaptation by modifying pollen flow and therefore changing the spatial genetic structure (Balkenhol et al., 2017; Gardiner et al., 2016; Wang et al., 2016) and in the case of insects, antennae and specifically the Johnston's organ, are directly involved in neuron response to wind (Fuller et al., 2014; Patella & Wilson, 2018).

In addition, climate variables such as temperature and precipitation can be highly correlated (Lotterhos et al., 2018). Relationship between explanatory variables, i.e., multicollinearity, compromises the results of multivariate regression analysis (Kim, 2019). Multicollinearity could yield unreliable regression parameter estimation, magnitude and sign of regression, which impedes the assessment of the relative importance of the explanatory variables (Sokal & Rohlf, 2013). This problem may be overcome by using synthetic variables obtained via principal component analysis (PCA) of the environmental variables of interest. However, using PCs based on climate variables may lead to a limited interpretation of the environment drivers of selection. The PCs will represent the environmental variables that covary the most, but this may not coincide with the combination of variables that drive divergent selection and local adaptation (Houle et al., 2002; Lotterhos et al., 2018).

Recently developed software such as the BayPass package (Gautier, 2015), have overcome some of the limitations of the GEA analyses mentioned above. On one hand, this software identifies those genetic variants with statistically different allele frequencies between populations and those associated with environmental variables, while taking into account the covariance between population allele frequencies due to, for instance, the joint demographic history of the samples analysed (Gautier, 2015). On the other hand, this software includes different modules based on different models: a single-covariate regression model where the association is estimated for each covariate, and a multiple-covariate regression model where the association is estimated for several covariates assumed to be orthogonal (Gautier, 2015).

Drosophila melanogaster is a valuable model organism to study environmental adaptation. This species originated from Southern Africa and has recently expanded worldwide colonizing a wide range of environmental conditions (12,000–19,000 years ago; Arguello et al., 2019; Pool et al., 2012; Sprengelmeyer et al., 2020). In addition, this species offers many key advantages as it has a small and well-annotated genome which facilitates the identification of putatively adaptive loci (Mohr et al., 2014), as well as a short lifecycle implying many generations in short periods of time (15 generations per year in nature; Pool, 2015). Past studies carried out with North American and Australian *D. melanogaster* populations have already shown clinal and seasonal genetic patterns suggesting that this species could be a good model to study environmental adaptation (Bergland et al., 2014, 2016; Fabian et al., 2012; Hoffmann & Weeks, 2007; Kolaczkowski et al., 2011; Machado et al., 2019). Indeed, this species has already been studied in other continents such as Europe, where clinal patterns and correlations between genetic variants and environmental variables have also been identified (Kapun et al., 2020; Lerat et al., 2019).

In this study, we combined genome scans for adaptive differentiation and whole-genome GEA analysis using pool-seq data available for 36 samples of *D. melanogaster*, representative of the genetic diversity across the European continent (n = 20 locations) and across a latitudinal cline in eastern North America (n = 11 locations). We focused on these two continents because they have an

approximately similar range of climatic conditions (mostly temperate climates) and they were both recently colonized, which allow us to focus on short-term evolutionary events. Our threefold aims were to characterize: (i) to which extent environmental variables contributed to adaptive differentiation in *D. melanogaster*; (ii) which climatic variables, namely temperature, rainfall, evaporation, solar radiation, wind, daylight hours and soil type, may be contributing to this environmental adaptation; and (iii) to which extent the observed signals were parallel across two different geographic areas: Europe and North America.

2 | MATERIALS AND METHODS

2.1 | Data sets

European pool-sequencing samples were obtained from the 2014 *DrosEU* data set (Kapun et al., 2020). We discarded 18 out of the 48 samples available, for which Tajima's *D* was very low (Tajima's $D < -0.2$; Kapun et al., 2020). For some locations, samples were collected several times across 2014. When several samples were available for the same season, summer or fall, we only included the earliest collected sample in the analysis. Thus, overall, we analysed 25 samples from 20 different locations (Figure 1). To perform the analysis, we created three data sets including only one sample per location (Table S1): Europe (20 samples), Europe Summer (14 samples), and Europe Fall (10 samples). Note that for the Europe data set, when samples of both seasons were available for the same location, we only included the summer sample. Average sequencing coverage among samples ranged from 25 to 190X (Table S1). VCFs are available at http://hdl.handle.net/10261/180630.

Eleven North American pool-sequencing samples collected from 2003 to 2014 were obtained from Machado et al., (2019) sampled in eleven different locations in the North American East coast (Figure 1; Table S1). We focused in these samples because clinality has been detected in previous studies (Bergland et al., 2014; Fabian et al., 2012; Figure 1; Table S1). VCFs are available at https://datadryad.org/stash/share/rHMqJSiXuGX12eBYyPvKE_Ng1b-FMTrLLnmegosbQ74.

In addition to SNPs, we also included in our analysis transposable element (TE) insertions. We estimated population frequencies for 1,630 euchromatic reference TE insertions (Rech et al., 2019) using *T-lex3* (Bogaerts-Márquez et al., 2019). For each data set, we only analysed those insertions that were polymorphic in at least one of the populations (403 TE insertions; Table S2).

2.2 | Environmental variables

We downloaded environmental data from four different sources: WorldClim (Fick & Hijmans, 2017), Copernicus (Hersbach et al., 2020), the US Naval Observatory (https://www.usno.navy.mil/

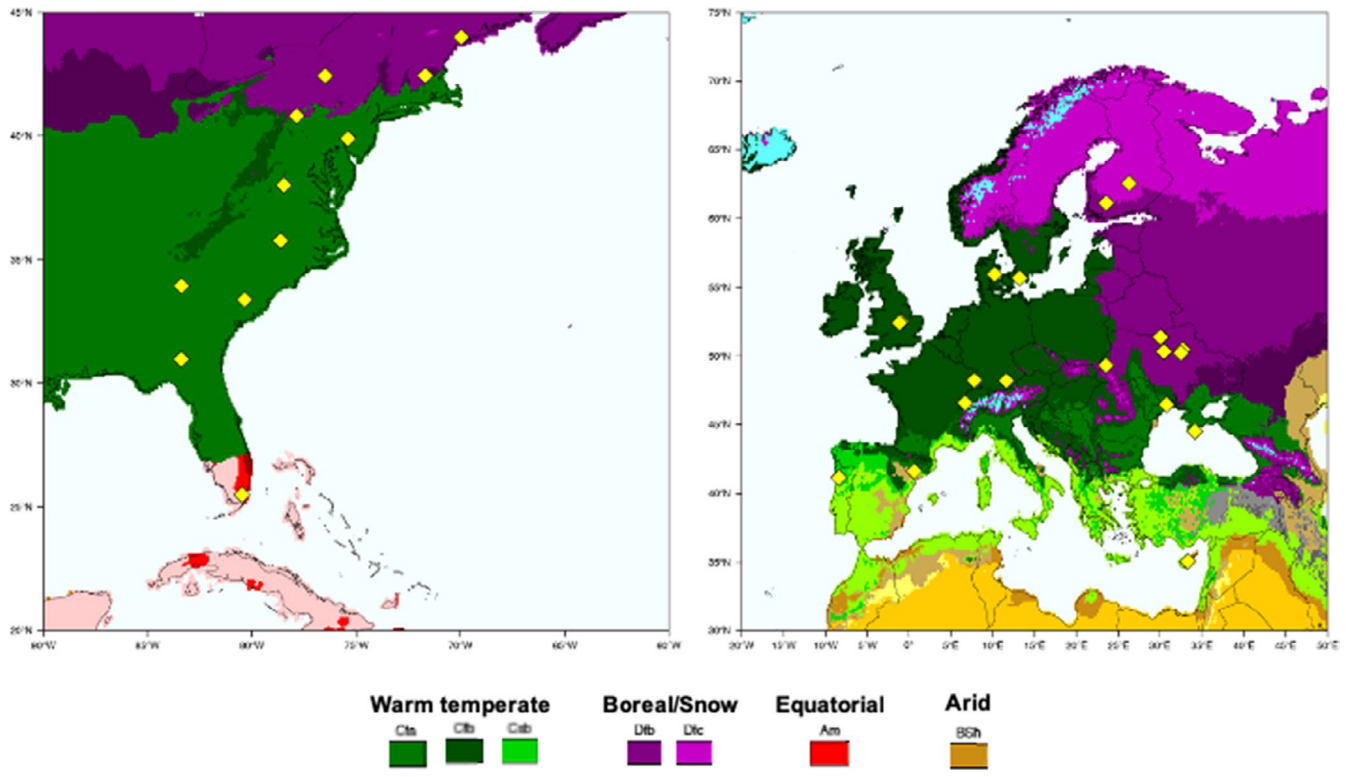

FIGURE 1 *Drosophila melanogaster* samples used in this study. Samples were collected across Europe and North America east coast (see Table S1) in four main climate zones and seven subclimate zones (depending on precipitation and level of heat), according to the Köppen-Geiger climate distribution (Kottek et al., 2006) [Colour figure can be viewed at wileyonlinelibrary.com]

USNO/astronomical-applications/data-services/data-services) and The Astronomical Data Portal UK Hydrographic Office (http://astro.ukho.gov.uk/). From WorldClim, we used the 19 Bioclimatic variables, which are derived from the monthly temperature and rainfall values from the 1970–2000 time range. We used the R package raster (v.2.6–7) for downloading this data. In addition, we also used year-specific environmental variables from the year previous to the collection date of each sample. For the year-specific variables, we used ERA5 "hourly data on single levels from 1979 to present" database from Copernicus, to obtain data on temperature (2 m temperature), rainfall (Total precipitation), evaporation (Evaporation), solar radiation (Clear-sky direct solar radiation at surface), wind (10 m u-component of wind and 10 m v-component of wind), and soil type. This data was downloaded as GRIB files and parsed using ecCodes package (v.2.8.2). Finally, daylight hours were obtained from the US Naval Observatory for the European data sets, and from The Astronomical Data Portal UK Hydrographic Office for the North American data set. For the year-specific variables, similar variables as the ones in WorldClim were constructed with the data from the year previous to the collection date using different python scripts (v.2.7.12). In total, we analysed 59 environmental variables related to temperature (11), rainfall (8), evaporation (11), solar radiation (7), wind (14), daylight hours (7) and soil type (1) (Table S3). We tested whether those variables that are common between WorldClim and Copernicus (i.e., temperature and rainfall) but that were obtained from different time ranges (average data from 1970 to 2000 and year-specific data from the year before the collection date) were significantly correlated using a Spearman correlation test ($\rho > 0.8$) (Table S3A). Most of the eight rainfall related variables were not correlated (six in Europe, six in Europe Summer, five in Europe Fall, and eight in North America), while most of the 11 variables related to temperature in Europe and NA data sets were correlated (seven in Europe, eight in Europe Summer, nine in Europe Fall, and seven in NA; Tables S3A–B). For variables that were not correlated, we included the two values of the same environmental variable in the analysis. When the correlation was significant, we performed the analysis using the variable corresponding to the WorldClim database.

In order to study the correlation among environmental variables, we performed Spearman rank correlation test for each pair of variables in each data set using the R function cor.test() and R package corrplot (v.0.84) (Table S3C). We considered and reported as strong correlations those with Spearman's $\rho > 0.8$. We found that solar radiation and daylight hours in Europe and in North America were highly correlated. In North America, but not in the European data sets, we also found that temperature, solar radiation, and daylight hours variables were highly correlated (Figure S1A–D). We also performed a PCA of the environmental variables for each of the four data sets analysed using stats (v.3.5.2) and plotnScree function in nFactors (v.2.4.1) R packages (Table S3D1–4). We found that temperature and solar radiation explained most of the variation of the PC1 both in Europe and Europe Summer data sets. In the Europe Fall data set solar radiation and daylight hours explain most of the PC1 variation while temperature and rainfall explain most of the PC1 variation in North America (Table S3D1–4).

2.3 | Whole-genome scans for adaptive differentiation

Whole genome-scans for adaptive differentiation were performed using BayPass (v.2.1) (Gautier, 2015). The model underlying BayPass accounts for the correlation structure among allele frequencies and allows identifying putative genetic variants subjected to adaptive differentiation based on the XtX statistic (Günther & Coop, 2013). This method, as demonstrated in Gautier (2015), has several critical advantages: (i) it explicitly and efficiently accounts for the confounding factors of the shared demographic history (via the covariance matrix -Ω); (ii) it makes no simplifying assumptions about the underlying demographic model; and (iii) it can explicitly model pool-seq data (via a binomial likelihood under the so-called poolseqmode) to account for the extra-variance introduced when sequencing pools of DNAs that are not individually identifiable (which basically prevents from distinguishing reads that are identical because they were obtained from the same sequence or from two distinct but identical sequences) (see Gautier, 2015 for a more detailed explanation). The genotyping input files contain the read count data (reference and alternative) per site and per sample. For SNPs, this information was obtained from the VCF files, while for the transposable elements (TEs) the information was obtained using T-lex: the absent read count information was used as the alternative read count, and the present read count information as the reference read count (Bogaerts-Márquez et al., 2019; Fiston-Lavier et al., 2015). To generate the input files for the North America samples, VCFs were parsed using the poolfstat (v.1.1.1) R package (Hivert et al., 2018). For the three European data sets, VCFs were parsed using python scripts, and bash and awk command lines. TE frequencies were added to the data using python scripts. Invariant and nonbiallelic positions were removed from each data set. We ran BayPass for autosomes (2L, 2R, 3L and 3R) and X chromosome separately because the autosome and X-linked variants have different haploid sample sizes as samples were obtained from male flies, and more importantly autosomes and X-chromosome have different demographic histories (e.g., Clemente et al., 2018).

We ran BayPass under the core model for the computation of the XtX genetic differentiation statistics for each data set separately (Europe, North America, Europe Summer and Europe Fall). As the number of SNPs in the autosomes in the four data sets was large (Table 1), we used a subsampling strategy as in Gautier et al. (2018), dividing each data set into 50 subdata sets containing only one SNP every 50 SNPs. We run the 50 subdata sets in parallel, thus all the SNPs available were used for the analysis. This strategy allows a more efficient analysis as it requires less computational time because each of the 50 pseudo-independent files are run in parallel. Note that the SNPs and TEs in each pseudo-independent file have low level of background linkage disequilibrium

TABLE 1 Summary of the four data sets used in this analysis: three European and one North American data set

		Autosomes		X chromosome		Total	
Data set	No of populations	SNPs	TEs	SNPs	TEs	SNPs	TEs
Europe	20	2,846,701	249	119,228	53	2,965,929	302
North America	11	2,147,276	280	291,632	64	2,438,908	344
Europe Fall	10	2,663,700	227	115,147	49	2,778,847	276
Europe Summer	14	2,725,176	222	117,332	42	2,842,508	264

(LD). Three independent runs (using the option -seed) were performed for each data set. To check for consistency the results of the three independent runs were evaluated using the Forstner and Moonen Distance (FMD) (Förstner & Moonen, 2003) between pairs of covariance matrices (Ω) with the R function fmd.dist() (provided within the BayPass package). We compared the covariance matrices among the 50 different subdata sets, and among the three different seeds runs. We found that FMD was low for all comparisons (Table S4A). Consistency was additionally tested comparing the posterior means of the two parameters α and β of the *Beta* distribution of the estimated population allele frequencies (Table S4B).

Prior to further analysis, we removed SNPs and TEs with very low allele frequency (MAF <0.01) based on the mean of the posterior distribution of the frequency of the reference allele across populations for each site (included in column *M_P* of the BayPass output summary_pi_xtx.out).

To obtain a calibrated estimator of the XtX statistic, we relied on the XtX* estimator recently described in (Olazcuaga et al., 2020). We further derived bilateral *p*-values assuming that XtX* follows a Chi-square distribution with npop degrees of freedom under the null hypothesis of neutral differentiation (Olazcuaga et al., 2020). To control for multiple testing, we estimated the associated q-value with the R package qvalue (v 3.9) (Storey & Tibshirani, 2003). For further analysis, we focused on the SNPs and TEs with the most highly significant XtX* scores (top 0.05% and q-value ≤0.05).

2.4 | Genome-environment association analysis

The so-called BayPass STD model extends the previous analysis to evaluate association of the genetic variant allele frequencies with population-specific covariates. We ran this model with the environmental variables previously described for each data set, and default options except for the -scalecov option that was used to scale each covariate. As we did with the previous model, we run the four data sets independently, as well as autosomes and X chromosome separately. For the autosome data sets, we also used the subsampling strategy mentioned above. Three independent runs were performed for each data set (using the option -seed).

The support for association between the genetic variants and the environmental variables was assessed using a Bayes factor (BF) measured in deciban (dB) units and estimated with an importance sampling algorithm from the MCMC samples (Coop et al., 2010; Gautier, 2015). More specifically, we used as an estimate the median BF among the three independent MCMC runs. We discarded SNPs and TEs present at very low allele frequency (MAF <0.01) in both observed and simulated data (see below). We considered a BF threshold of 20 dB (i.e., "decisive evidence" according to the Jeffreys' rule [Jeffreys, 1961]) as evidence for association between an environmental variable and a TE and an even more stringent threshold of 30 dB for SNPs to limit the number of false positives (SNPs being far more numerous than TEs). We evaluated the false positive rates (FPR) associated with these thresholds based on the analysis of pseudo-observed data sets (PODs) generated using the same parameters as for the observed data sets according to the approach described in Gautier (2015). Briefly, the rationale of this approach is to provide an empirical null distribution of the BF statistic, i.e., neutrally evolving SNPs are simulated under the generative model parameterized with the matrix Ω, which is estimated on the real data to summarize the joint demographic history of the populations (and to capture its effect on the neutral covariance structure on population allele frequencies). The estimated FPR for the 20 BF and 30 BF thresholds ranged from 0% to 2.40% and 0% to 0.33%, respectively for the association tests with the different covariates (Table S5). When a SNP or a TE was significantly associated with more than one environmental variable, we considered as the primary variable the one with the relative highest BF compared to the 99.9% of the POD distribution (Table S5), which usually coincides with the absolute highest BF value.

2.5 | Analysis of candidate genes

We identified the genes where the significant SNPs were located using bedtools intersect (v.2.27.1) and the *D. melanogaster* FlyBase reference genome annotations v6.12 and v5.50 for the European and the North American data sets, respectively (Thurmond et al., 2019). We also identified significant SNPs located in gene regulatory region (<1 kb upstream of genes) (Hoskins et al., 2011) using SnpEff (v.4.3) (BDGP5.75 data base for North American and BDGP6.86 for European data sets). For TEs, we also used FlyBase annotations to check whether they were located <1 kb upstream or downstream of a gene, inside a gene, or in intergenic regions.

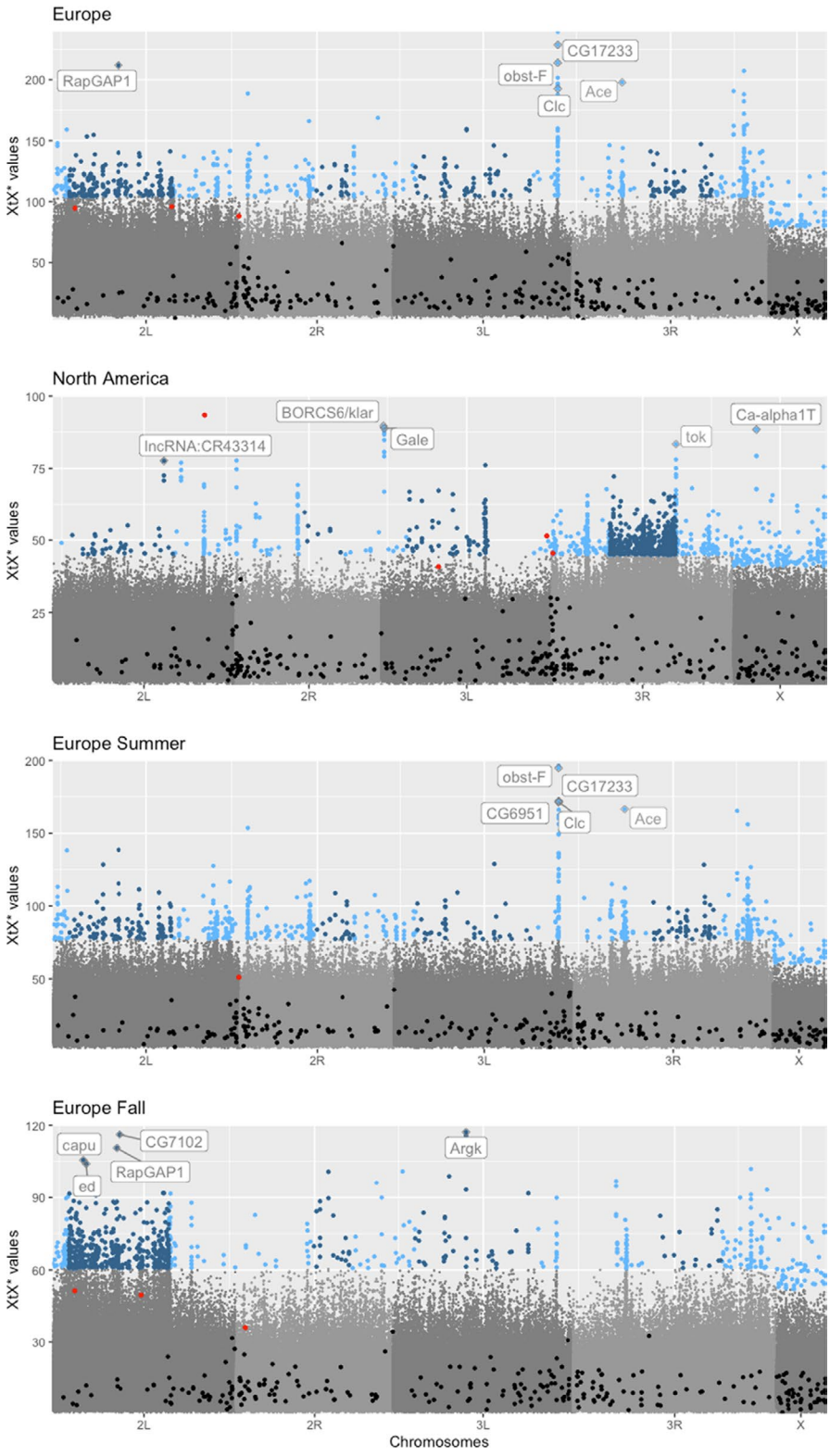

FIGURE 2 Genome-wide distribution of the XtX* values (population differentiation) associated with single nucleotide polymorphisms (SNPs) and transposable elements (TEs) in the four data sets. Nonsignificant SNPs and TEs are plotted in grey and black, respectively. Significant SNPs located inside and outside of inversions are plotted in dark blue and blue, respectively. Significant TEs are plotted in red. Genes for the five most significant SNPs for each data set are highlighted [Colour figure can be viewed at wileyonlinelibrary.com]

We performed a gene ontology (GO) enrichment analysis of the candidate genes using DAVID (v.6.8). A cluster was considered to be significant when the enrichment score was higher than 1.3 (Huang et al., 2009).

To test if there was a significant overlap of candidate genes between Europe and North America (624,069 shared SNPs corresponding to 15,944 genes), we use the SuperExactTest R package (v.1.0.7).

3 | RESULTS

3.1 | Development and signalling underlie population differentiation in Europe and North America

To characterize the patterns of genetic differentiation in European and in North American *D. melanogaster* natural populations, we ran the BayPass core model in two data sets containing 20 and 11 populations, respectively (Figure 1, Table 1 and Table S1). Samples were collected from seven climate types distributed in four climate zones (Figure 1, Table S1; Kottek et al., 2006). For some European populations, we had samples collected in summer and fall. Thus, in addition to the whole European data set, we also analysed the summer (Europe Summer data set) and the fall (Europe Fall data set) samples separately (Kapun et al., 2020).

We first analysed the distribution of the SNPs that showed significant population differentiation patterns across chromosomes (Figure 2). We tested whether any of the four main cosmopolitan inversions described in *D. melanogaster* (*In(2L)t*, *In(2R)NS*, *In(3L)P*, and *In(3R)P*) were enriched for significantly differentiated SNPs by comparing the SNPs located inside each insertion with the rest of the chromosome (Corbett-Detig & Hartl, 2012; Figure 2, Table S6A). In the European and European Fall data sets, inversion *In(2L)t* was enriched for significantly differentiated SNPs (Fisher exact test *p*-value <0.001). This inversion shows a strong frequency gradient in European populations ranging from 2% to 50% (Table S6E–F; Kapun et al., 2020). On the other hand, in the North American data set inversion *In(3R)P* was enriched for significantly differentiated SNPs (Fisher's exact test *p*-value <0.001), which also shows a strong frequency gradient in North American populations ranging from 0% to 41% (Tables S6E–F; Kapun et al., 2016; Kapun & Flatt, 2019). Our results are consistent with previous analyses that found that these two inversions show latitudinal and/or seasonal clinal patterns mainly in Australia and North America (Kapun et al., 2016; Kapun & Flatt, 2019).

We considered genes with at least one significant SNP located in the gene body region or in their 1 kb upstream regions to be candidates for adaptive differentiation, i.e., to be subjected to selection (Table S7; see Materials and Methods; Hoskins et al., 2011). Overall, we identified 1,300 candidate genes. Among our candidates, we found genes previously known to play a role in adaptation, such as *cpo*, involved in reproductive dormancy (Cogni et al., 2014; Schmidt et al., 2008), *sgg* involved in circadian rhythm (Rand et al., 2010; Wolf et al., 2007), *mth* involved in longevity and stress response (Schmidt et al., 2000), and *Ace*, involved mainly in insecticide resistance (Fournier et al., 1992; Menozzi et al., 2004). We also found other interesting genes, which have been previously reported as candidates in North America but not in Europe, such as *obst-F*, which has been suggested to be involved in longitudinal clinality (Table 2; Campo et al., 2013). *obst-F* is involved in the cuticle formation, which acts as a barrier between the fly and the environment protecting it from insecticides, solar radiation, and desiccation (Balabanidou et al., 2018; Behr & Hoch, 2005; Ferveur et al., 2018; Rajpurohit et al., 2018). Indeed, most of the SNPs with the highest differentiation score were located in genes that are candidate for several stress responses, such as oxidative and starvation stress response, and behavioural phenotypes (Table 2).

To identify which biological processes underlay the population differentiation in the four data sets analysed, we performed a GO term enrichment analysis (Table 3 and Table S8). Both in Europe and in North America, the most significant clusters were related to development and signalling, suggesting that similar biological processes have been involved in adaptation in the two continents. Signalling was the most enriched cluster in the Europe Fall data set, while development and morphogenesis were enriched both in the Europe Summer and Europe Fall data sets (Table 3). These results are similar to previous analysis performed in *D. melanogaster*: development and morphogenesis have been reported in population differentiation studies in different continents such as Europe, Australia, and North America (Fabian et al., 2012; Kolaczkowski et al., 2011; Mateo et al., 2018; Reinhardt et al., 2014). Note that excluding the SNPs that are located inside inversions led to very similar enriched biological processes GO terms (Table S8B).

Finally, we also tested whether there was a significant overlap between the candidate genes for local adaptation found in Europe and in North America. We found 55 significant genes overlapping in the two continents (SuperExactTest *p*-value <0.05; Table S9A, see Materials and Methods). Among these 55 genes, we found already known genes such as *cpo* and *Ace*, as well as other genes previously identified in North American clinal studies such as *Cow*, involved in neuromuscular junction development (Kopke et al., 2020) or *dpy*, involved in wing and trachea development (Wilkin et al., 2000) (Table S7; Table S9B). We performed a GO enrichment analysis with these 55 overlapping genes among continents, and the main clusters were related to regulation, signalling and response to stimulus, and development (Table S9C).

3.2 | Temperature, rainfall, and wind are the most contributing variables in the genome-environment association analyses

To identify the environmental variables that are relevant for adaptation in *D. melanogaster* natural populations, we looked for significant associations between SNPs frequencies and several environmental variables using the BayPass standard model (see Materials and Methods). We analysed 59 environmental variables related to temperature, rainfall, evaporation, solar radiation, wind, daylight hours, and soil type (Table S3 and S5, see also Materials and Methods), and we found significant associations between at least one of these variables and 748 genes (Table S10).

For all data sets, temperature was the variable associated with the highest number of genes, followed by wind in Europe and Europe Fall, and rainfall in the North America and Europe Summer data sets (Table 4). Significant SNPs located in some of these genes were

TABLE 2 Candidate genes showing the most significant population differentiation patterns

Gene	SNP location	Data set	XtX*	Phenotype
BORCS6/klar	Gene body/Upstream	NA	89.63	–/Alcohol, Starvation
Gale	Upstream	NA	88.85	Aggressiveness; Diapause; Immunity; Starvation
Ca-alpha1T	Gene body	NA	88.40	Olfactory
tok	Gene body	NA	83.40	Circadian; Starvation
lncRNA:CR43314	Upstream	NA	77.60	–
CG6951	Upstream	EuS	171.38	Alcohol; Oxidative
Argk	Upstream	EuF	117.21	Immunity; Starvation
capu	Gene body	EuF	105.60	Alcohol, Circadian behavior, Oxidative, Xenobiotic
ed	Gene body	EuF	104.07	Olfactory, Oxidative
CG7102	Gene body	EuF	116.21	–
Ace	Gene body	Eu	197.83	Diapause, Insecticide resistance, Olfactory, Starvation
	Gene body	EuS	166.55	
obst-F	Gene body	Eu	213.90	–
	Gene body	EuS	194.73	
CG17233	Gene body	Eu	228.63	–
	Gene body	EuS	172.06	
Clc	Upstream/ Gene body	Eu	192.62	Starvation
		EuS	171.38	
RapGAP1	Gene body	Eu	211.90	Aggressiveness
	Gene body	EuF	110.64	

For each data set, top 5 genes with SNPs located in the gene body or upstream region (< 1kb) with the highest significant XtX* values and their associated phenotype (see Table S12).

Abbreviations: Eu, Europe; EuF, Europe Fall; EuS, Europe Summer; NA, North America.

associated with more than one variable as expected from the correlation found between some of the covariates (Figure S1). For instance, in North America most of the SNPs associated with solar radiation variables were also associated with Temperature variables (84/103), which is consistent with the correlation found between these variables (Figure 3b). Note that, the majority of SNPs that were associated with wind were not associated with any other environmental variables (Figure 3a,d), which is consistent with the lack of significant correlation between wind and other environmental variables (Figure S1). On the other hand, in Europe Summer the majority of SNPs associated with evaporation were also associated with temperature, although we did not find a strong correlation between evaporation and temperature variables (Figure 3c, Figure S1D). However, there are studies reporting similar responses to cold and desiccation stress (Sinclair et al., 2007). Among the top five genes with the highest BF scores in the four data sets, we found several associations with Annual mean temperature and Annual mean solar radiation (Table 5).

We also tested whether candidate genes with SNPs associated with environmental variables were enriched inside cosmopolitan inversions (Figure 2; Corbett-Detig & Hartl, 2012). We found an enrichment of significant SNPs in the *In(2L)t* inversion in the Europe Fall data set and in the *In(3R)P* inversion in the North America data set (Fisher's exact test *p*-value <0.001; Table S6B). The *In(2L)t* inversion in the Europe Fall data set was enriched for SNPs associated with temperature variables while the *In(3R)P* in the North America data set was enriched for SNPs associated with rainfall, solar radiation, and wind variables (Table S6C–D). In previous studies *In(2 l) t* and *In(3R)P* were correlated with climatic factors varying latitudinally in North America, specifically with temperature and rainfall (Kapun et al., 2016). Our analysis suggests that other climatic factors such as wind may also be correlated with inversions.

Finally, we also found a significant overlap between the genes with SNPs significantly associated with environmental variables identified in the North American and the European data sets (SuperExactTest *p*-value <0.05; see Materials and Methods; Table S9 and S10). Among the 32 significantly overlapping genes, *fipi* is involved in the *Drosophila* courtship song (Fedotov et al., 2018) and was associated with variables related to wind in Europe and North America. Courtship song, as well as wind have been shown to activate neurons which are related to the antennal and mechanosensory motor center in the central brain in *D. melanogaster* (Matsuo & Kamikouchi, 2013; Yorozu et al., 2009).

TABLE 3 GO enrichment analysis of candidate genes for local adaptation

Data set	Significant SNPs	Significant genes	GO enrichment terms
Europe	719	410	Neuron development; eye development; signalling; organ morphogenesis; growth
North America	1,164	583	Response to stimulus; organ development; regulation of growth; nervous system development; localization and transport
Europe Summer	752	396	Learning/memory; eye development; neuron development; sensory perception of pain; organ morphogenesis
Europe Fall	821	412	Signalling; localization/transport; organ morphogenesis; neuron development; heart morphogenesis

For each data set, the number of genes and significant single nucleotide polymorphisms (SNPs), located in the gene body and upstream region (<1 kb), and the top five most enriched GO terms (significance >1.3). The significance of the SNPs was determined based on the empirical distribution of the calibrated XtX* values (top 0.05%), which corresponds to q-value thresholds of 7.56e-10 in Europe, 3.70e-07 in Europe Summer, 1.10e-05 in Europe Fall, 9.90e-06 in North America for autosomes; and to *q*-value thresholds of 1.49e-05 in Europe, 0.0003 in Europe Summer, 0.000441 in Europe Fall and 0.03 North America for X chromosome.

TABLE 4 Summary of results obtained for environmental association

	Europe	NA	Europe Summer	Europe Fall
Temperature	**143**	**217**	**87**	**33**
Wind	**118**	83	17	**20**
Rainfall	29	**152**	**62**	**13**
Evaporation	36	79	**51**	7
Solar radiation	**45**	**103**	19	3
Soil	–	4	–	–
Daylight hours	18	52	8	1
Total	296	382	155	64

Number of genes with significant SNPs (BF ≥ 30) located in the gene body or upstream region (< 1 kb) for each type of environmental variable. In bold, three top type of environmental variables with more genes for each data set.

3.3 | 23% to 51% of the genes significantly associated with environmental variables did not show adaptive differentiation across populations

We found that, across data sets, only 12% to 37% of the genes that showed patterns of population differentiation (XtX*) were associated with at least one environmental variable (Table S7). Indeed, most of the genes that showed the highest association with environmental variables, such as *Ace* and *obst-F*, were also among the top candidates for significant population differentiation (Tables 2 and 5). Another example is *Gale*, which in North America was associated with a wind variable, and has been related with aggressiveness and diapause responses as well as with immunity and starvation stresses (Tables 2 and 5; Clark et al., 2013; Edwards et al., 2006; Fukuyama et al., 2013; Grönke et al., 2005; Harbison et al., 2005; Kučerová et al., 2016; Shorter et al., 2015; Zhao et al., 2016).

On the other hand, we found that 23% to 51% of the genes that showed a significant association with at least one environmental variable, did not show population differentiation patterns (Table S10). Among these genes, *RFeSP* has one of the highest association scores with Wind seasonality in the Europe Fall data set (Table 5 and Table S10). This gene encodes *Rieske iron sulphur* proteins which are highly conserved functional constituents of energy-transducing respiratory complexes (Gontijo et al., 2011).

3.4 | Ten transposable elements insertions are associated with environmental variables

In addition to SNPs, we also analysed the population differentiation patterns and correlations with environmental variables for TE insertions (Table S11A). We found that nine out of the 403 TE insertions showed patterns of population differentiation (XtX*) in at least one of the data sets analysed; however, we did not find overlap between continents (Figure 2, Table 6, Table S11A–B). Four of these TEs have previously been identified as candidate adaptive TEs (Table 6).

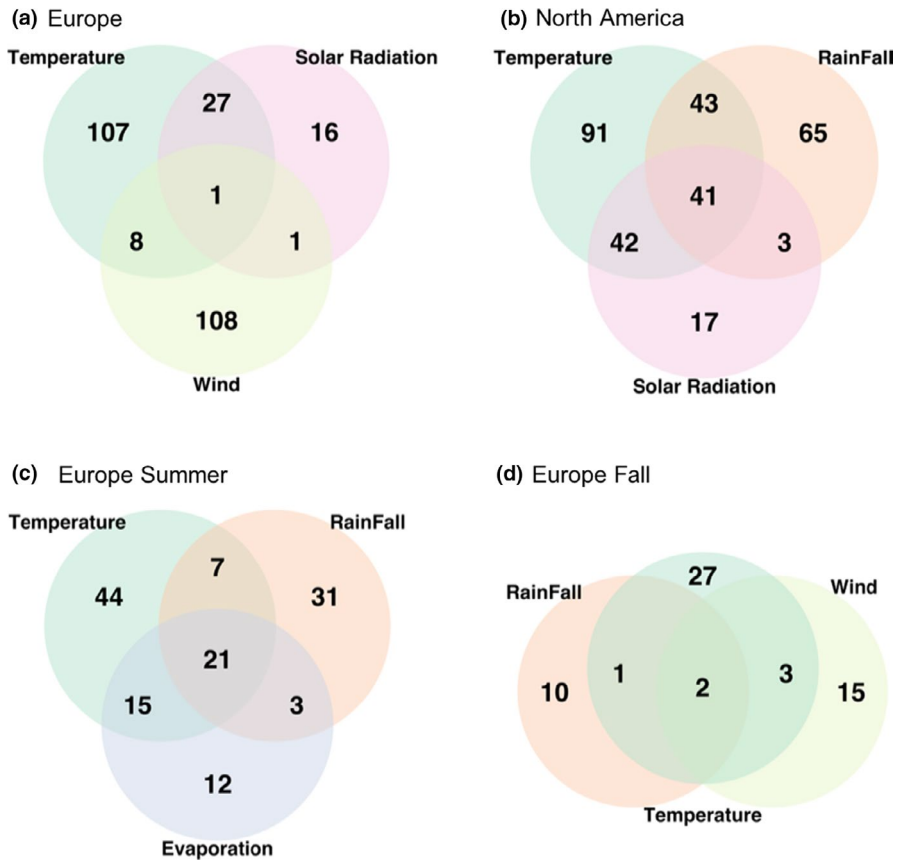

FIGURE 3 Overlap between genes with single nucleotide polymorphisms (SNPs) significantly associated with environmental variables. For each data set, genes with SNPs in the gene body and upstream region (<1 kb) significantly associated with the three groups of environmental variables with more genes associated with them are depicted [Colour figure can be viewed at wileyonlinelibrary.com]

In addition, we identified six TE insertions in the Europe data set that showed significant associations with different environmental variables: four of them showing the highest association with temperature variables, one with evaporation, and one with rainfall (Table 7 and Table S11C). Three of these insertions also showed significant patterns of population differentiation (*FBti0019112*, *FBti0019164* and *FBti0019862*; Table 6). *FBti0019112* showed the highest BF value and was associated with the Minimum temperature of the coldest month variable (Table 7). This insertion is located in an intron of the *lilli* gene, which is mainly involved in cell identity and growth, and plays a role in retinal development (Distefano et al., 2012; Wittwer et al., 2001). This result is interesting given that other studies in *Drosophila* showed the impact of temperature in eye development genes (Del Bel et al., 2018). In addition, *lilli* has been suggested to have a role in local adaptation, as it was recently reported to be part of a strong outlier region in a study comparing *D. melanogaster* collected in wilderness areas and collected in the nearby of towns in southern-central Africa (Sprengelmeyer et al., 2020). *FBti0018880* showed the second strongest association also with temperature, Isothermality (temperature variability index), and has been reported to play a role in oxidative stress response (Guio et al., 2014). Other studies performed in *D. melanogaster* showed that metabolites involved in oxidative stress are altered by selection in cold tolerance (Koštál et al., 2016; Williams et al., 2014).

We also identified four TE insertions in the North America data that showed significant associations with wind, solar radiation, rainfall and evaporation (Table 7; Table S11A,C). Two of these insertions also showed patterns of population differentiation (*FBti0061428* and *FBti0020306*; Table 6).

4 | DISCUSSION

In this work, we aimed to identify the main environmental drivers of adaptation of *D. melanogaster* natural populations in a large continental geographical scale. To accomplish this, we used GEA analysis on a large set of population samples (up to 20 in Europe and 11 in North America) representative of different environments and considering a wide-range of environmental covariates to capture this variability. We found that in addition to temperature and rainfall, wind related variables appear to be also relevant for *D. melanogaster* environmental adaptation. Temperature and rainfall are the most widely used variables in GEA analysis in several species including *D. melanogaster* (Božičević et al., 2016; Cavedon et al., 2019; Gao et al., 2017; Hopley & Byrne, 2019; Kapun et al., 2020; Leroy et al., 2020; Mayol et al., 2020; Pina-Martins et al., 2019; Todesco et al., 2019). Our results show that the majority of genes associated with environmental variables were associated

TABLE 5 Candidate genes associated with environmental variables

Gene Name	SNP location	Data set	Strongest association variable	BF	Phenotype
Ace	Gene body	Eu	Annual mean temperature	84.89	Diapause, insecticide resistance, olfactory, starvation
Sap-r	Gene body	Eu	Annual mean solar radiation	62.74	Starvation
obst-F	Gene body	Eu	Annual mean temperature	72.05	-
apn	Gene body	Eu	Mean temperature of warmest quarter	60.59	-
Ptr	Gene body	Eu	Mean evaporation of warmest quarter	70.75	-
tok	Gene body	NA	Annual mean solar radiation/ Solar rad mean diurnal range	64.34	Circadian, starvation
Mhc	Gene body	NA	Solar rad mean diurnal range	61.76	-
CG13705	Gene body	NA	Temperature seasonality	58.18	-
Abd-B	Gene body	NA	Annual mean solar radiation	51.11	Alcohol, dessication, pigmentation
Gale	Upstream	NA	Wind mean diurnal range	50.57	Aggressiveness, diapause, immunity, starvation
Ace	Gene body	EuS	Annual mean temperature	74.38	Diapause, olfactory, starvation
obst-F	Gene body	EuS	Max temperature of warmest month	68.08	-
CG7290/ CG7298	Gene body/ Upstream	EuS	Max temperature of warmest month	58.37	- /Hypoxia, immunity, xenobiotic
CG10257	Upstream	EuS	Precipitation of driest quarter	58.48	Xenobiotic
Ptr	Gene body	EuS	Mean evaporation of warmest quarter	62.11	-
capu	Gene body	EuF	Wind variability index	49.59	Alcohol, circadian behaviour, oxidative, xenobiotic
Kek2	Gene body	EuF	Wind variability index	44.49	-
Argk	Upstream	EuF	Temperature seasonality	65.35	Immunity, starvation
RFeSP	Gene body	EuF	Wind seasonality	48.51	Hypoxia
CG43750	Gene body	EuF	Wind seasonality	42,35	-

For each data set, the top five genes with significant single nucleotide polymorphisms (SNPs) located in the gene body and upstream region (<1 kb) with the highest significant Bayes factor (BF) scores, the environmental variable with the strongest association and their associated phenotype (see Table S12). All significant genes can be found in Table S10.

Abbreviations: EuA, Europe; EuF, Europe Fall; EuS, Europe Summer; NA, North America.

with a temperature-related one (400/748), while the number of genes associated with rainfall was smaller (241/748) (Table S10). These results are consistent with similar GEA analysis performed previously in *D. melanogaster* (Kapun et al., 2020). Moreover, among the 748 candidate genes associated with environmental variables in our study, 226 were associated with a wind-related variable being the third variable group with most associations, and far from the following ones (evaporation and solar radiation with 153/748 genes each) (Table S10). Wind-related variables have been studied mainly in plant adaptation (Balkenhol et al., 2017), and are often included as part of PCs where individual wind effect cannot be properly measured (Gao et al., 2017). In other species, wind has also been reported to be involved in desiccation stress and thermoregulation (Baig & Tranquillini, 1980; Ortega et al., 2017). In *Drosophila* species, including in *D. melanogaster*, it has already been suggested that wind might be relevant for adaptation (Fuller et al., 2014; Patella & Wilson, 2018). The effect of wind variables in *Drosophila* could be related to the Johnston's organ, which is the largest mechanosensory organ in *Drosophila*. This organ is involved in a variety of behaviours such as hearing, touch, vestibular sensing, proprioception and wind sensing (Patella & Wilson, 2018). In addition, Fuller et al. (2014) reported how flies regulate flight speed according to the information from their visual system and their antennae, and how they can overcome the effect of sudden wind disturbances. To the best of our knowledge, our analysis is the first that identifies genome-wide variants associated with wind-related variables in *D. melanogaster*. Although temperature, rainfall, and wind seem to be important drivers of adaptation, we still lack information about other variables directly related to them and that may be actually underlying adaptive processes. Further analysis testing the direct effect of these three variables on the genetic variation should be performed to obtain conclusive evidence.

TABLE 6 Significant transposable element insertions found in the population differentiation analysis

Transposable element	Family	Location	Gene	Data set	Evidence of selection
FBti0019112	*pogo*	First intron	*lilli*	Eu, EuF	iHS, H12, nSL (Rech et al., 2019)
FBti0019164	*X-element*	First intron	*CG9932*	Eu	Population differentiation (González et al., 2008)
FBti0019144	*Rt1b*	First intron	*CG44153*	EuF	Population differentiation (González et al., 2008)
FBti0019276	S-element	Second intron	*Adf1*	EuF	CSTV (Lerat et al., 2019)
FBti0019862	*G6*	432 bp downstream	*Tif-IA*	Eu, EuS	–
FBti0020056	BS	507 bp downstream	*bin*	NA	–
FBti0020306	hopper	Third intron/first intron	*atms/CG44098*	NA	–
FBti0060187	G2	First intron	*Syn1*	NA	–
FBti0061428	hobo	52 bp upstream/529 bp downstream	*CG31809/CG6012*	NA	–

TE insertions were considered significant if their associated XtX* values were above the top 1% of the empirical distribution of XtX* values, and *q*-value <0.05. When the TE insertion is located in intergenic regions, genes located nearby are reported (Table S11A).

Abbreviations: CSTV, correlation with spatiotemporal variables; Eu, Europe; EuF, Europe Fall; EuS, Europe Summer; NA, North America.

TABLE 7 Significant candidate TE insertions associated with environmental variables (BF ≥ 20)

Transposable element	Environmental variable	Significant XtX*	BF	Data set
FBti0018880	Isothermality	No	30.53	Eu
FBti0019112	Min temperature of coldest month	Yes	43.38	Eu
FBti0019164	Temperature Annual range	Yes	24.79	Eu
FBti0019165	Evaporation Mean diurnal range	No	20.52	Eu
FBti0020057	Precipitation Seasonality	No	22.12	Eu
FBti0019862	Isothermality	Yes	23.69	Eu
FBti0061428	Annual mean wind	Yes	43.95	NA
FBti0020086	Solar radiation variability index	No	26.83	NA
FBti0020306	Precipitation of wettest quarter	Yes	26.02	NA
FBti0019318	Mean evaporation of coldest quarter	No	28.57	NA

The environmental variable with highest score is reported (Table S11).

Abbreviations: Eu, Europe; NA, North America.

We found that an important proportion of the genes showing signals of adaptive differentiation did not show associations with any of the environmental variables studied (>60%). As we addressed previously, it is difficult to know a priori the variables that may be relevant for adaptation, so, it could be that we are not including in our analysis the environmental variables responsible for the adaptive processes in which these genes are involved. For example, 185 of the 1,300 genes showing population differentiation patterns are candidates for xenobiotic stress response (Rech et al., 2019). For the majority of these genes, we did not find any association with environmental variables. Thus, including variables related to pollution might help explain the population differentiation in some of these genes. It could also be that although the relevant environmental variables are included in the analysis, our samples do not allow us to capture the whole range of the variation of these environmental variables making the GEA analyses less powerful. Alternatively, population differentiation patterns in some of these genes might be due to selective pressures not related to the environment. We also found that between 23% and 51% of

the candidate genes showed association with an environmental variable but did not show significant population differentiation. Thus, GEA analyses not only identifies the relevant environmental variables, but also allows to identify genetic variants involved in environmental adaptation that cannot be detected through population differentiation analysis, as expected if they result in too subtle changes in allele frequencies across populations (Gautier, 2015).

Our work also aimed at evaluating to which extent our observed signal of adaptation were consistent across the European and North American continents. We found that 55 genes showing patterns of population differentiation, and 32 genes showing association with at least one environmental variable, overlapped in these two continents. Out of these 32 genes, 14 were associated with a different environmental variable group in each continent, and 12 were associated with different variables in the same group. This suggests that although there is a pattern of parallel adaptation, there may be different environmental pressures which may drive adaptation for the same genes.

We also assessed whether samples collected in European populations in summer and fall differed in their association with environmental variables. Recent studies have shown the role of temperature in seasonal variation (Machado et al., 2019). We found that in the summer and in the fall data sets the majority of genes were associated with temperature and rainfall (Figure 3c, d, Table 4). However, there was a substantial proportion of candidate genes associated with evaporation in Europe Summer but not in Europe Fall (33% vs. 11%; Table 4). On the other hand, there were more candidate genes associated with wind variables in the Europe Fall than in the Europe Summer data set (31% vs. 11%; Table 4). These results suggest that different environmental variables, evaporation and wind, might play a role across seasons. However, temporal data series from several years should be analysed to confirm these results.

Finally, we identified four TE insertions showing significant population differentiation patterns, five TE insertions associated with an environmental variable, and five insertions showing both. We described as candidates for the first time three of these TE insertions, *FBti0019862*, *FBti0020306* and *FBti0061428*, which are associated with environmental variables and showed significant population differentiation patterns, as well as *FBti0019164* only reported in González et al. (2008) and *FBti0019112* which has shown previous evidence of selection (Rech et al., 2019). This analysis is, however, limited as we could only investigate those TE insertions present in the reference genome and that were polymorphic in our samples. We suggest that both reference and non-reference insertions should be included in future analysis in order to get a comprehensive picture of the role of TE insertions in environmental adaptation.

Overall, we identified temperature, rainfall and wind as environmental variables which may play a critical role in environmental adaptive processes in *D. melanogaster*. In addition to increasing the number of populations and of TE insertions analysed, we further suggest that performing GEA analysis in populations collected across time should inform us about how the role of environmental variables changes through time and contributes to the dynamics of genetic variation across populations and to the maintenance of adaptive variants. Extending this analysis to other continents should also further enhance our understanding of the role of environmental variables in adaptive evolution.

ACKNOWLEDGEMENTS

We acknowledge the support from the European Drosophila Population Genomics Consortium (DrosEU), which is funded by a Special Topics Network (STN) grant from the European Society of Evolutionary Biology (ESEB). We also thank Thomas Flatt for comments on the manuscript. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (H2020-ERC-2014-CoG-647900).

AUTHOR CONTRIBUTIONS

Designed research: Josefa González. *Performed research*: María Bogaerts-Márquez, Sara Guirao-Rico and Mathieu Gautier. *Contributed new reagents or analytical tools*: María Bogaerts-Márquez. Analysed data: María Bogaerts-Márquez and Josefa González. *Drafted the manuscript*: María Bogaerts-Márquez and Josefa González. *Edited and approved the manuscript*: María Bogaerts-Márquez, Sara Guirao-Rico, Mathieu Gautier and Josefa González.

DATA AVAILABILITY STATEMENT

SNP genotyping data for European Samples are available at the public repository DIGITAL.CSIC (http://hdl.handle.net/10261/180630) and North American samples are available in the Dryad database https://datadryad.org/stash/share/rHMqJSiXuGX12eBYyPvKE_Ng1b-FMTrLLnmegosbQ74. SRA accession numbers are specified in Table S1, and environmental variable data is available in Table S3B.

Scripts are available at https://github.com/GonzalezLab/environmental_variables

ORCID

María Bogaerts-Márquez https://orcid.org/0000-0001-9107-984X
Sara Guirao-Rico https://orcid.org/0000-0001-9896-4665
Mathieu Gautier https://orcid.org/0000-0001-7257-5880
Josefa González https://orcid.org/0000-0001-9824-027X

REFERENCES

Ahrens, C. W., Rymer, P. D., Stow, A., Bragg, J., Dillon, S., Umbers, K. D. L., & Dudaniec, R. Y. (2018). The search for loci under selection: trends, biases and progress. *Molecular Ecology*, *27*(6), 1342–1356. https://doi.org/10.1111/mec.14549.

Arguello, J. R., Laurent, S., & Clark, A. G. (2019). Demographic history of the human commensal *Drosophila* melanogaster. *Genome Biology Evolution*, *11*(3), 844–854. https://doi.org/10.1093/gbe/evz022.

Baig, M. N., & Tranquillini, W. (1980). The effects of wind and temperature on cuticular transpiration of Picea abies and Pinus cembra and their significance in dessication damage at the alpine treeline. *Oecologia*, *47*(2), 252–256. https://doi.org/10.1007/BF00346828.

Balabanidou, V., Grigoraki, L., & Vontas, J. (2018). Insect cuticle: a critical determinant of insecticide resistance. *Current Opinion in Insect Science*, *27*, https://doi.org/10.1016/j.cois.2018.03.001.

Balkenhol, N., Dudaniec, R. Y., Krutovsky, K. V., Johnson, J. S., Cairns, D. M., Segelbacher, G., Selkoe, K. A., von der Heyden, S., Wang, I. J., Selmoni, O., & Joost, S. (2017). Landscape Genomics: Understanding Relationships Between Environmental Heterogeneity and Genomic Characteristics of Populations. In O. P. Rajora (Ed.), *Population Genomics: Concepts, Approaches and Applications* (pp. 261–322). Springer International Publishing. https://doi.org/10.1007/13836_2017_2.

Behr, M., & Hoch, M. (2005). Identification of the novel evolutionary conserved obstructor multigene family in invertebrates. *FEBS Letters*, *579*(30), 6827–6833. https://doi.org/10.1016/j.febslet.2005.11.021.

Bergland, A. O., Behrman, E. L., O'Brien, K. R., Schmidt, P. S., & Petrov, D. A. (2014). Genomic evidence of rapid and stable adaptive oscillations over seasonal time scales in *Drosophila*. *PLoS Genetics*, *10*(11), e1004775. https://doi.org/10.1371/journal.pgen.1004775.

Bergland, A. O., Tobler, R., González, J., Schmidt, P., & Petrov, D. (2016). Secondary contact and local adaptation contribute to genome-wide patterns of clinal variation in Drosophila melanogaster. *Molecular Ecology*, *25*(5), 1157–1174. https://doi.org/10.1111/mec.13455.

Bogaerts-Márquez, M., Barrón, M. G., Fiston-Lavier, A. S., Vendrell-Mir, P., Castanera, R., Casacuberta, J. M., & González, J. (2019). T-lex3: an accurate tool to genotype and estimate population frequencies of transposable elements using the latest short-read whole genome sequencing data. *Bioinformatics*, *36*: 1191–1197. https://doi.org/10.1093/bioinformatics/btz727.

Božičević, V., Hutter, S., Stephan, W., & Wollstein, A. (2016). Population genetic evidence for cold adaptation in European Drosophila melanogaster populations. *Molecular Ecology*, *25*(5), 1175–1191. https://doi.org/10.1111/mec.13464.

Campo, D., Lehmann, K., Fjeldsted, C., Souaiaia, T., Kao, J., & Nuzhdin, S. V. (2013). Whole-genome sequencing of two North American Drosophila melanogaster populations reveals genetic differentiation and positive selection. *Molecular Ecology*, *22*(20), 5084–5097. https://doi.org/10.1111/mec.12468.

Cavedon, M., Gubili, C., Heppenheimer, E., vonHoldt, B., Mariani, S., Hebblewhite, M., Hegel, T., Hervieux, D., Serrouya, R., Steenweg, R., Weckworth, B. V., & Musiani, M. (2019). Genomics, environment and balancing selection in behaviourally bimodal populations: The caribou case. *Molecular Ecology*, *28*(8), 1946–1963. https://doi.org/10.1111/mec.15039.

Clark, R. I., Tan, S. W., Péan, C. B., Roostalu, U., Vivancos, V., Bronda, K., Pilátová, M., Fu, J., Walker, D. W., Berdeaux, R., Geissmann, F., & Dionne, M. S. (2013). MEF2 is an in vivo immune-metabolic switch. *Cell*, *155*(2), 435–447. https://doi.org/10.1016/j.cell.2013.09.007.

Clemente, F., Gautier, M., & Vitalis, R. (2018). Inferring sex-specific demographic history from SNP data. *PLoS Genetics*, *14*(1), e1007191. https://doi.org/10.1371/journal.pgen.1007191.

Cogni, R., Kuczynski, C., Koury, S., Lavington, E., Behrman, E. L., O'Brien, K. R., Schmidt, P. S., & Eanes, W. F. (2014). The intensity of selection acting on the couch potato gene–spatial-temporal variation in a diapause cline. *Evolution*, *68*(2), 538–548. https://doi.org/10.1111/evo.12291.

Coop, G., Witonsky, D., Di Rienzo, A., & Pritchard, J. K. (2010). Using environmental correlations to identify loci underlying local adaptation. *Genetics*, *185*(4), 1411–1423. https://doi.org/10.1534/genetics.110.114819.

Corbett-Detig, R. B., & Hartl, D. L. (2012). Population genomics of inversion polymorphisms in Drosophila melanogaster. *PLoS Genetics*, *8*(12), e1003056. https://doi.org/10.1371/journal.pgen.1003056.

Del Bel, L. M., Griffiths, N., Wilk, R., Wei, H. C., Blagoveshchenskaya, A., Burgess, J., Polevoy, G., Price, J. V., Mayinger, P., & Brill, J. A. (2018). The phosphoinositide phosphatase Sac1 regulates cell shape and microtubule stability in the developing. *Development*, *145*(11), dev151571. https://doi.org/10.1242/dev.151571.

Distefano, G. M., Gangemi, A. J., Khandelwal, P. J., Saunders, A. J., & Marenda, D. R. (2012). *Drosophila* lilliputian is required for proneural gene expression in retinal development. *Developmental Dynamics*, *241*(3), 553–562. https://doi.org/10.1002/dvdy.23738.

Edwards, A. C., Rollmann, S. M., Morgan, T. J., & Mackay, T. F. (2006). Quantitative genomics of aggressive behavior in *Drosophila* melanogaster. *PLoS Genetics*, *2*(9), e154. https://doi.org/10.1371/journal.pgen.0020154.

Fabian, D. K., Kapun, M., Nolte, V., Kofler, R., Schmidt, P. S., Schlötterer, C., & Flatt, T. (2012). Genome-wide patterns of latitudinal differentiation among populations of *Drosophila* melanogaster from North America. *Molecular Ecology*, *21*(19), 4748–4769. https://doi.org/10.1111/j.1365-294X.2012.05731.x.

Fedotov, S. A., Bragina, J. V., Besedina, N. G., Danilenkova, L. V., Kamysheva, E. A., & Kamyshev, N. G. (2018). Gene CG15630 (fipi) is involved in regulation of the interpulse interval in *Drosophila* courtship song. *Journal of Neurogenetics*, *32*(1), 15–26. https://doi.org/10.1080/01677063.2017.1405000.

Ferveur, J. F., Cortot, J., Rihani, K., Cobb, M., & Everaerts, C. (2018). Desiccation resistance: effect of cuticular hydrocarbons and water content in D*rosophila melanogaster* adults. *PeerJ*, *6*, e4318. https://doi.org/10.7717/peerj.4318.

Fick, S. E., & Hijmans, R. J. (2017). Worldclim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, *37*(12), 4302–4315. https://doi.org/10.1002/joc.5086

Fiston-Lavier, A. S., Barrón, M. G., Petrov, D. A., & González, J. (2015). T-lex2: genotyping, frequency estimation and re-annotation of transposable elements using single or pooled next-generation sequencing data. *Nucleic Acids Research*, *43*(4), e22. https://doi.org/10.1093/nar/gku1250.

Flood, P. J., & Hancock, A. M. (2017). The genomic basis of adaptation in plants. *Current Opinion in Plant Biology*, *36*, 88–94. https://doi.org/10.1016/j.pbi.2017.02.003.

Fournier, D., Bride, J. M., Hoffmann, F., & Karch, F. (1992). Acetylcholinesterase. Two types of modifications confer resistance to insecticide. *Journal of Biological Chemistry*, *267*(20), 14270–14274.

Fukuyama, H., Verdier, Y., Guan, Y., Makino-Okamura, C., Shilova, V., Liu, X., Maksoud, E., Matsubayashi, J., Haddad, I., Spirohn, K., Ono, K., Hetru, C., Rossier, J., Ideker, T., Boutros, M., Vinh, J., & Hoffmann, J. A. (2013). Landscape of protein-protein interactions in *Drosophila* immune deficiency signaling during bacterial challenge. *Proceedings of the Naional Academy of Sciences U S A*, *110*(26), 10717–10722. https://doi.org/10.1073/pnas.1304380110.

Fuller, S. B., Straw, A. D., Peek, M. Y., Murray, R. M., & Dickinson, M. H. (2014). Flying *Drosophila* stabilize their vision-based velocity controller by sensing wind with their antennae. *Proceedings of the Naional Academy of Sciences U S A*, *111*(13), E1182–1191. https://doi.org/10.1073/pnas.1323529111 [Record #101 is using a reference type undefined in this output style.]

Gao, Y., Gautier, M., Ding, X., Zhang, H., Wang, Y., Wang, X., Faruque, M. O., Li, J., Ye, S., Gou, X., Han, J., Lenstra, J. A., & Zhang, Y. (2017). Species composition and environmental adaptation of indigenous Chinese cattle. *Scientific Reports*, *7*(1), 16196. https://doi.org/10.1038/s41598-017-16438-7.

Gardiner, B., Berry, P., & Moulia, B. (2016). Review: Wind impacts on plant growth, mechanics and damage. *Plant Science*, *245*, 94–118. https://doi.org/10.1016/j.plantsci.2016.01.006.

Gautier, M. (2015). Genome-Wide Scan for Adaptive Divergence and Association with Population-Specific Covariates. *Genetics*, *201*(4), 1555–1579. https://doi.org/10.1534/genetics.115.181453.

Gautier, M., Yamaguchi, J., Foucaud, J., Loiseau, A., Ausset, A., Facon, B., Gschloessl, B., Lagnel, J., Loire, E., Parrinello, H., Severac, D., Lopez-Roques, C., Donnadieu, C., Manno, M., Berges, H., Gharbi, K., Lawson-Handley, L., Zang, L. S., Vogel, H., … Prud'homme, B.

(2018). The Genomic Basis of Color Pattern Polymorphism in the Harlequin Ladybird. *Current Biology*, *28*(20), 3296–3302.e3297. https://doi.org/10.1016/j.cub.2018.08.023.

Gontijo, A. M., Miguela, V., Whiting, M. F., Woodruff, R. C., & Dominguez, M. (2011). Intron retention in the *Drosophila* melanogaster Rieske Iron Sulphur Protein gene generated a new protein. *Nature Communications*, *2*, 323. https://doi.org/10.1038/ncomms1328.

González, J., Lenkov, K., Lipatov, M., Macpherson, J. M., & Petrov, D. A. (2008). High rate of recent transposable element-induced adaptation in *Drosophila* melanogaster. *PLoS Biology*, *6*(10), e251. https://doi.org/10.1371/journal.pbio.0060251.

Gorter, F. A., Scanlan, P. D., & Buckling, A. (2016). Adaptation to abiotic conditions drives local adaptation in bacteria and viruses co-evolving in heterogeneous environments. *Biology Letters*, *12*(2), 20150879. https://doi.org/10.1098/rsbl.2015.0879.

Grönke, S., Mildner, A., Fellert, S., Tennagels, N., Petry, S., Müller, G., Jäckle, H., & Kühnlein, R. P. (2005). Brummer lipase is an evolutionary conserved fat storage regulator in *Drosophila*. *Cell Metabolism*, *1*(5), 323–330. https://doi.org/10.1016/j.cmet.2005.04.003.

Guio, L., Barrón, M. G., & González, J. (2014). The transposable element Bari-Jheh mediates oxidative stress response in *Drosophila*. *Molecular Ecology*, *23*(8), 2020–2030. https://doi.org/10.1111/mec.12711.

Günther, T., & Coop, G. (2013). Robust identification of local adaptation from allele frequencies. *Genetics*, *195*(1), 205–220. https://doi.org/10.1534/genetics.113.152462.

Harbison, S. T., Chang, S., Kamdar, K. P., & Mackay, T. F. (2005). Quantitative genomics of starvation stress resistance in *Drosophila*. *Genome Biology*, *6*(4), R36. https://doi.org/10.1186/gb-2005-6-4-r36.

Helfrich-Förster, C., Bertolini, E., & Menegazzi, P. (2020). Flies as models for circadian clock adaptation to environmental challenges. *European Journal of Neuroscience*, *51*(1), 166–181. https://doi.org/10.1111/ejn.14180.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., … Thépaut, J.-N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049.https://doi.org/https://doi.org/10.1002/qj.3803 .

Hivert, V., Leblois, R., Petit, E. J., Gautier, M., & Vitalis, R. (2018). Measuring Genetic Differentiation from Pool-seq Data. *Genetics*, *210*(1), 315–330. https://doi.org/10.1534/genetics.118.300900.

Hoban, S., Kelley, J. L., Lotterhos, K. E., Antolin, M. F., Bradburd, G., Lowry, D. B., Poss, M. L., Reed, L. K., Storfer, A., & Whitlock, M. C. (2016). Finding the Genomic Basis of Local Adaptation: Pitfalls, Practical Solutions, and Future Directions. *The American Naturalist*, *188*(4), 379–397. https://doi.org/10.1086/688018.

Hoffmann, A. A., & Weeks, A. R. (2007). Climatic selection on genes and traits after a 100 year-old invasion: a critical look at the temperate-tropical clines in *Drosophila* melanogaster from eastern Australia. *Genetica*, *129*(2), 133–147. https://doi.org/10.1007/s10709-006-9010-z.

Hopley, T., & Byrne, M. (2019). Gene Flow and Genetic Variation Explain Signatures of Selection across a Climate Gradient in Two Riparian Species. *Genes (Basel)*, *10*(8), https://doi.org/10.3390/genes10080579.

Hoskins, R. A., Landolin, J. M., Brown, J. B., Sandler, J. E., Takahashi, H., Lassmann, T., Yu, C., Booth, B. W., Zhang, D., Wan, K. H., Yang, L., Boley, N., Andrews, J., Kaufman, T. C., Graveley, B. R., Bickel, P. J., Carninci, P., Carlson, J. W., & Celniker, S. E. (2011). Genome-wide analysis of promoter architecture in *Drosophila* melanogaster. *Genome Research*, *21*(2), 182–192. https://doi.org/10.1101/gr.112466.110.

Houle, D., Mezey, J., & Galpern, P. (2002). Interpretation of the results of common principal components analyses. *Evolution*, *56*(3), 433–440. https://doi.org/10.1111/j.0014-3820.2002.tb01356.x.

Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocol*, *4*(1), 44–57. https://doi.org/10.1038/nprot.2008.211.

Jeffreys, H. (1961). *Theory of Probability*, 3rd ed.. Oxford University Press.

Kapun, M., Barrón, M. G., Staubach, F., Obbard, D. J., Wiberg, R. A. W., Vieira, J., Goubert, C., Rota-Stabelli, O., Kankare, M., Bogaerts-Márquez, M., Haudry, A., Waidele, L., Kozeretska, I., Pasyukova, E. G., Loeschcke, V., Pascual, M., Vieira, C. P., Serga, S., Montchamp-Moreau, C., … González, J. (2020). Genomic analysis of European *Drosophila* melanogaster populations reveals longitudinal structure, continent-wide selection, and previously unknown DNA viruses. *Molecular Biology and Evolution*, *9*, 2661–2678. https://doi.org/10.1093/molbev/msaa120.

Kapun, M., Fabian, D. K., Goudet, J., & Flatt, T. (2016). Genomic evidence for adaptive inversion clines in *Drosophila* melanogaster. *Molecular Biology and Evolution*, *33*(5), 1317–1336. https://doi.org/10.1093/molbev/msw016.

Kapun, M., & Flatt, T. (2019). The adaptive significance of chromosomal inversion polymorphisms in *Drosophila* melanogaster. *Molecular Ecology*, *28*(6), 1263–1282. https://doi.org/10.1111/mec.14871.

Kim, J. H. (2019). Multicollinearity and misleading statistical results. *Korean J Anesthesiology*, *72*(6), 558–569. https://doi.org/10.4097/kja.19087.

Kolaczkowski, B., Kern, A. D., Holloway, A. K., & Begun, D. J. (2011). Genomic differentiation between temperate and tropical Australian populations of *Drosophila* melanogaster. *Genetics*, *187*(1), 245–260. https://doi.org/10.1534/genetics.110.123059.

Kopke, D. L., Leahy, S. N., Vita, D. J., Lima, S. C., Newman, Z. L., & Broadie, K. (2020). Carrier of Wingless (Cow). *Regulation of Drosophila Neuromuscular Junction Development*. *Eneuro*, *7*(2), https://doi.org/10.1523/ENEURO.0285-19.2020.

Körner, C. (2007). The use of 'altitude' in ecological research. *Trends in Ecology and Evolution*, *22*(11), 569–574. https://doi.org/10.1016/j.tree.2007.09.006.

Koštál, V., Korbelová, J., Štětina, T., Poupardin, R., Colinet, H., Zahradníčková, H., Opekarová, I., Moos, M., & Šimek, P. (2016). Physiological basis for low-temperature survival and storage of quiescent larvae of the fruit fly *Drosophila* melanogaster. *Scientific Reports*, *6*, 32346. https://doi.org/10.1038/srep32346.

Kottek, M., Grieser, J., Beck, C., Rudolf, B., & Rubel, F. (2006). World Map of the Köppen-Geiger Climate Classification Updated. *Meteorologische Zeitschrift*, *15*, 259–263. https://doi.org/10.1127/0941-2948/2006/0130.

Kučerová, L., Kubrak, O. I., Bengtsson, J. M., Strnad, H., Nylin, S., Theopold, U., & Nässel, D. R. (2016). Slowed aging during reproductive dormancy is reflected in genome-wide transcriptome changes in *Drosophila* melanogaster. *BMC Genomics*, *17*, 50. https://doi.org/10.1186/s12864-016-2383-1.

Lerat, E., Goubert, C., Guirao-Rico, S., Merenciano, M., Dufour, A.-B., Vieira, C., & González, J. (2019). Population-specific dynamics and selection patterns of transposable element insertions in European natural populations. *Molecular Ecology*, *28*(6), 1506–1522. https://doi.org/10.1111/mec.14963.

Leroy, T., Louvet, J. M., Lalanne, C., Le Provost, G., Labadie, K., Aury, J. M., Delzon, S., Plomion, C., & Kremer, A. (2020). Adaptive introgression as a driver of local adaptation to climate in European white oaks. *The New Phytologist*, *226*, 1171–1182. https://doi.org/10.1111/nph.16095.

Lotterhos, K. E., Yeaman, S., Degner, J., Aitken, S., & Hodgins, K. A. (2018). Modularity of genes involved in local adaptation to climate despite physical linkage. *Genome Biology*, *19*(1), 157. https://doi.org/10.1186/s13059-018-1545-7.

Machado, H. E., Bergland, A. O., Taylor, R., Tilk, S., Behrman, E., Dyer, K., Fabian, D. K., Flatt, T., González, J., Karasov, T. L., Kozeretska, I., Lazzaro, B. P., Merritt, T. J. S., Pool, J. E., O'Brien, K., Rajpurohit,

S., Roy, P. R., Schaeffer, S. W., Serga, S., … Petrov, D. A. (2019). *Broad Geographic Sampling Reveals Predictable, Pervasive, and Strong Seasonal Adaptation in Drosophila.*, *bioRxiv*, 337543. https://doi.org/10.1101/337543.

Mateo, L., Rech, G. E., & González, J. (2018). Genome-wide patterns of local adaptation in Western European *Drosophila* melanogaster natural populations. *Scientific Reports*, *8*(1), 16143. https://doi.org/10.1038/s41598-018-34267-0.

Matsuo, E., & Kamikouchi, A. (2013). Neuronal encoding of sound, gravity, and wind in the fruit fly. *Journal of Comparative Physiology A, Neuroethology Sensory, Neural, and Behavioral Physiology*, *199*(4), 253–262. https://doi.org/10.1007/s00359-013-0806-x.

Mayol, M., Riba, M., Cavers, S., Grivet, D., Vincenot, L., Cattonaro, F., Vendramin, G. G., & González-Martínez, S. C. (2020). A multiscale approach to detect selection in nonmodel tree species: Widespread adaptation despite population decline in *Taxus baccata L*. *Evolutionary Applications*, *13*(1), 143–160. https://doi.org/10.1111/eva.12838.

Menozzi, P., Shi, M. A., Lougarre, A., Tang, Z. H., & Fournier, D. (2004). Mutations of acetylcholinesterase which confer insecticide resistance in *Drosophila* melanogaster populations. *BMC Evolutionary Biology*, *4*, 4. https://doi.org/10.1186/1471-2148-4-4.

Mohr, S. E., Hu, Y., Kim, K., Housden, B. E., & Perrimon, N. (2014). Resources for functional genomics studies in *Drosophila* melanogaster. *Genetics*, *197*(1), 1–18. https://doi.org/10.1534/genetics.113.154344.

Nelson, T. C., Jones, M. R., Velotta, J. P., Dhawanjewar, A. S., & Schweizer, R. M. (2019). UNVEILing connections between genotype, phenotype, and fitness in natural populations. *Molecular Ecology*, *28*(8), 1866–1876. https://doi.org/10.1111/mec.15067.

Norton, H. L., Kittles, R. A., Parra, E., McKeigue, P., Mao, X., Cheng, K., Canfield, V. A., Bradley, D. G., McEvoy, B., & Shriver, M. D. (2007). Genetic evidence for the convergent evolution of light skin in Europeans and East Asians. *Molecular Biology and Evolution*, *24*(3), 710–722. https://doi.org/10.1093/molbev/msl203.

Olazcuaga, L., Loiseau, A., Parrinello, H., Paris, M., Fraimout, A., Guedot, C., Diepenbrock, L. M., Kenis, M., Zhang, J., Chen, X., Borowiec, N., Facon, B., Vogt, H., Price, D. K., Vogel, H., Prud'homme, B., Estoup, A., & Gautier, M. (2020). A whole-genome scan for association with invasion success in the fruit fly *Drosophila* suzukii using contrasts of allele frequencies corrected for population structure. *Molecular Biology and Evolution*, https://doi.org/10.1093/molbev/msaa098.

Ortega, Z., Mencía, A., & Pérez-Mellado, V. (2017). Wind constraints on the thermoregulation of high mountain lizards. *International Journal of Biometeorology*, *61*(3), 565–573. https://doi.org/10.1007/s00484-016-1233-9.

Patella, P., & Wilson, R. I. (2018). Functional Maps of Mechanosensory Features in the *Drosophila* Brain. *Current Biology*, *28*(8), 1189–1203. e1185. https://doi.org/10.1016/j.cub.2018.02.074.

Pina-Martins, F., Baptista, J., Pappas, G., & Paulo, O. S. (2019). New insights into adaptation and population structure of cork oak using genotyping by sequencing. *Global Change Biology*, *25*(1), 337–350. https://doi.org/10.1111/gcb.14497.

Pool, J. E. (2015). The Mosaic Ancestry of the *Drosophila* Genetic Reference Panel and the D. melanogaster Reference Genome Reveals a Network of Epistatic Fitness Interactions. *Molecular Biology and Evolution*, *32*(12), 3236–3251. https://doi.org/10.1093/molbev/msv194.

Pool, J. E., Corbett-Detig, R. B., Sugino, R. P., Stevens, K. A., Cardeno, C. M., Crepeau, M. W., Duchen, P., Emerson, J. J., Saelao, P., Begun, D. J., & Langley, C. H. (2012). Population Genomics of sub-saharan *Drosophila* melanogaster: African diversity and non-African admixture. *PLoS Genetics*, *8*(12), e1003080. https://doi.org/10.1371/journal.pgen.1003080.

Rajpurohit, S., Gefen, E., Bergland, A. O., Petrov, D. A., Gibbs, A. G., & Schmidt, P. S. (2018). Spatiotemporal dynamics and genome-wide association genome-wide association analysis of desiccation tolerance in *Drosophila* melanogaster. *Molecular Ecology*, *27*(17), 3525–3540. https://doi.org/10.1111/mec.14814.

Rand, D. M., Weinreich, D. M., Lerman, D., Folk, D., & Gilchrist, G. W. (2010). Three selections are better than one: clinal variation of thermal QTL from independent selection experiments in *Drosophila*. *Evolution*, *64*(10), 2921–2934. https://doi.org/10.1111/j.1558-5646.2010.01039.x.

Rech, G. E., Bogaerts-Márquez, M., Barrón, M. G., Merenciano, M., Villanueva-Cañas, J. L., Horváth, V., Fiston-Lavier, A. S., Luyten, I., Venkataram, S., Quesneville, H., Petrov, D. A., & González, J. (2019). Stress response, behavior, and development are shaped by transposable element-induced mutations in *Drosophila*. *PLoS Genetics*, *15*(2), e1007900. https://doi.org/10.1371/journal.pgen.1007900.

Reinhardt, J. A., Kolaczkowski, B., Jones, C. D., Begun, D. J., & Kern, A. D. (2014). Parallel geographic variation in *Drosophila* melanogaster. *Genetics*, *197*(1), 361–373. https://doi.org/10.1534/genetics.114.161463.

Rellstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M., & Holderegger, R. (2015). A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology*, *24*(17), 4348–4370. https://doi.org/10.1111/mec.13322.

Schmidt, P. S., Duvernell, D. D., & Eanes, W. F. (2000). Adaptive evolution of a candidate gene for aging in *Drosophila*. *Proceedings of the National Academy of Sciences U S A*, *97*(20), 10861–10865. https://doi.org/10.1073/pnas.190338897.

Schmidt, P. S., Zhu, C. T., Das, J., Batavia, M., Yang, L., & Eanes, W. F. (2008). An amino acid polymorphism in the couch potato gene forms the basis for climatic adaptation in *Drosophila* melanogaster. *Proceedings of the National Academy of Sciences U S A*, *105*(42), 16207–16211. https://doi.org/10.1073/pnas.0805485105.

Shorter, J., Couch, C., Huang, W., Carbone, M. A., Peiffer, J., Anholt, R. R., & Mackay, T. F. (2015). Genetic architecture of natural variation in *Drosophila* melanogaster aggressive behavior. *Proceedings of the National Academy of Sciences U S A*, *112*(27), E3555–3563. https://doi.org/10.1073/pnas.1510104112.

Sinclair, B. J., Gibbs, A. G., & Roberts, S. P. (2007). Gene transcription during exposure to, and recovery from, cold and desiccation stress in *Drosophila* melanogaster. *Insect Molecular Biology*, *16*(4), 435–443. https://doi.org/10.1111/j.1365-2583.2007.00739.x.

Smit, B., Whitfield, M. C., Talbot, W. A., Gerson, A. R., McKechnie, A. E., & Wolf, B. O. (2018). Avian thermoregulation in the heat: phylogenetic variation among avian orders in evaporative cooling capacity and heat tolerance. *Journal of Experimental Biology*, *221*(Pt, 6). https://doi.org/10.1242/jeb.174870.

Sokal, R., & Rohlf, F. (2013). Biometry : the principles and practice of statistics in biological research / Robert R. Sokal and F. James Rohlf. SERBIULA (sistema Librum 2.0).

Sprengelmeyer, Q. D., Mansourian, S., Lange, J. D., Matute, D. R., Cooper, B. S., Jirle, E. V., Stensmyr, M. C., & Pool, J. E. (2020). Recurrent Collection of *Drosophila* melanogaster from Wild African Environments and Genomic Insights into Species History. *Molecular Bioogy and Evolution*, *37*(3), 627–638. https://doi.org/10.1093/molbev/msz271.

Storey, J. D., & Tibshirani, R. (2003). Statistical significance for genome-wide studies. *Proceedings of the National Academy of Sciences U S A*, *100*(16), 9440–9445. https://doi.org/10.1073/pnas.1530509100.

Svetec, N., Cridland, J. M., Zhao, L., & Begun, D. J. (2016). The Adaptive Significance of Natural Genetic Variation in the DNA Damage Response of *Drosophila* melanogaster. *PLoS Genetics*, *12*(3), e1005869. https://doi.org/10.1371/journal.pgen.1005869.

Thurmond, J., Goodman, J. L., Strelets, V. B., Attrill, H., Gramates, L. S., Marygold, S. J., Matthews, B. B., Millburn, G., Antonazzo, G., Trovisco, V., Kaufman, T. C., Calvi, B. R., & Consortium, F (2019). FlyBase 2.0: the next generation. *Nucleic Acids Research*, *47*(D1), D759–D765. https://doi.org/10.1093/nar/gky1003.

Tishkoff, S. A., Reed, F. A., Ranciaro, A., Voight, B. F., Babbitt, C. C., Silverman, J. S., Powell, K., Mortensen, H. M., Hirbo, J. B., Osman, M., Ibrahim, M., Omar, S. A., Lema, G., Nyambo, T. B., Ghori, J., Bumpstead, S., Pritchard, J. K., Wray, G. A., & Deloukas, P. (2007). Convergent adaptation of human lactase persistence in Africa and Europe. *Nature Genetics*, *39*(1), 31–40. https://doi.org/10.1038/ng1946.

Todesco, M., Owens, G. L., Bercovich, N., Légaré, J.-S., Soudi, S., Burge, D. O., Huang, K., Ostevik, K. L., Drummond, E. B. M., Imerovski, I., Lande, K., Pascual, M. A., Cheung, W., Staton, S. E., Muños, S., Nielsen, R., Donovan, L. A., Burke, J. M., Yeaman, S., & Rieseberg, L. H. (2019). *Massive Haplotypes Underlie Ecotypic Differentiation in Sunflowers*., *bioRxiv*, 790279. https://doi.org/10.1101/790279.

Van't Hof, A. E., Campagne, P., Rigden, D. J., Yung, C. J., Lingley, J., Quail, M. A., Hall, N., Darby, A. C., & Saccheri, I. J. (2016). The industrial melanism mutation in British peppered moths is a transposable element. *Nature*, *534*(7605), 102–105. https://doi.org/10.1038/nature17951.

Wang, Z.-F., Lian, J.-Y., Ye, W.-H., Cao, H.-L., Zhang, Q.-M., & Wang, Z.-M. (2016). Pollen and seed flow under different predominant winds in wind-pollinated and wind-dispersed species Engelhardia roxburghiana. *Tree Genetics & Genomes*, *12*(2), 19. https://doi.org/10.1007/s11295-016-0973-3.

Wilkin, M. B., Becker, M. N., Mulvey, D., Phan, I., Chao, A., Cooper, K., Chung, H. J., Campbell, I. D., Baron, M., & MacIntyre, R. (2000). *Drosophila* dumpy is a gigantic extracellular protein required to maintain tension at epidermal-cuticle attachment sites. *Current Biology*, *10*(10), 559–567. https://doi.org/10.1016/s0960-9822(00)00482-6.

Williams, C. M., Watanabe, M., Guarracino, M. R., Ferraro, M. B., Edison, A. S., Morgan, T. J., Boroujerdi, A. F., & Hahn, D. A. (2014). Cold adaptation shapes the robustness of metabolic networks in *Drosophila* melanogaster. *Evolution*, *68*(12), 3505–3523. https://doi.org/10.1111/evo.12541.

Wittwer, F., van der Straten, A., Keleman, K., Dickson, B. J., & Hafen, E. (2001). Lilliputian: an AF4/FMR2-related protein that controls cell identity and cell growth. *Development*, *128*(5), 791–800.

Wolf, F. W., Eddison, M., Lee, S., Cho, W., & Heberlein, U. (2007). GSK-3/Shaggy regulates olfactory habituation in *Drosophila*. *Proceedings of the National Academy of Sciences U S A*, *104*(11), 4653–4657. https://doi.org/10.1073/pnas.0700493104.

Wu, E. J., Wang, Y. P., Shen, L. L., Yahuza, L., Tian, J. C., Yang, L. N., Shang, L. P., Zhu, W., & Zhan, J. (2019). Strategies of P*hytophthora infestans* adaptation to local UV radiation conditions. *Evolutionary Applications*, *12*(3), 415–424. https://doi.org/10.1111/eva.12722.

Yorozu, S., Wong, A., Fischer, B. J., Dankert, H., Kernan, M. J., Kamikouchi, A., Ito, K., & Anderson, D. J. (2009). Distinct sensory representations of wind and near-field sound in the *Drosophila* brain. *Nature*, *458*(7235), 201–205. https://doi.org/10.1038/nature07843.

Zhao, X., Bergland, A. O., Behrman, E. L., Gregory, B. D., Petrov, D. A., & Schmidt, P. S. (2016). Global Transcriptional Profiling of Diapause and Climatic Adaptation in *Drosophila* melanogaster. *Molecular Biology and Evolution*, *33*(3), 707–720. https://doi.org/10.1093/molbev/msv263.

Zhou, C., James, J. G., Xu, Y., Tu, H., He, X., Wen, Q., Price, M., Yang, N., Wu, Y., Ran, J., Meng, Y., & Yue, B. (2020). Genome-wide analysis sheds light on the high-altitude adaptation of the buff-throated partridge (Tetraophasis szechenyii). *Molecular Genetics and Genomics*, *295*(1), 31–46. https://doi.org/10.1007/s00438-019-01601-8.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Bogaerts-Márquez M, Guirao-Rico S, Gautier M, González J. Temperature, rainfall and wind variables underlie environmental adaptation in natural populations of *Drosophila melanogaster*. *Mol Ecol*. 2021;30:938–954. https://doi.org/10.1111/mec.15783