Analysis of a contact area between two distinct evolutionary honeybee units: an ecological perspective

Fernando Cánovas, Pilar De la Rúa, José Serrano & José Galián

Journal of Insect Conservation
An international journal devoted to
the conservation of insects and related
invertebrates

ISSN 1366-638X

J Insect Conserv
DOI 10.1007/s10841-014-9701-1

Springer

J Insect Conserv
DOI 10.1007/s10841-014-9701-1

ORIGINAL PAPER

Analysis of a contact area between two distinct evolutionary honeybee units: an ecological perspective

Fernando Cánovas · Pilar De la Rúa · José Serrano · José Galián

Received: 25 June 2014 / Accepted: 10 September 2014

Abstract Determining the relationships between environmental variables and genetic diversity is critical to understand the processes that drive evolution in species-contact areas. We employed a combination of modeling approaches and multivariate statistical analyses to analyze mtDNA diversity in a hybrid zone between two evolutionary lineages of honeybees in order to interpret the microevolutionary processes that led to the observed spatial pattern of diversity. The model located the west European honeybee lineage mainly across temperate areas characterized by mild winters and high water availability throughout the year, whereas the African lineage was mainly associated with warmer and drier areas. Selection could be playing an important role in shaping the life history evolution, particularly affecting the mitochondria, and also resulting in hitchhiking effects on particular regions of the mitochondrial genome.

Keywords *Apis mellifera* · Canonical correspondence analysis · Generalized additive models · Synthesis maps · LGM

Electronic supplementary material The online version of this article (doi:10.1007/s10841-014-9701-1) contains supplementary material, which is available to authorized users.

F. Cánovas (✉)
Centro de Ciências do Mar do Algarve, CCMAR, Campus de Gambelas, 8005-139 Faro, Portugal
e-mail: fcgarcia@ualg.pt

P. De la Rúa · J. Serrano · J. Galián
Department of Zoology and Physic Anthropology, University of Murcia, 30100 Espinardo, Murcia, Spain

Introduction

Given the large body of literature covering the behavior, physiology and evolution of the honeybee, this species represents a well-suited model organism to study interactions between environmental and micro-evolutionary events (reviewed in Elekonich and Roberts 2005; Lodesani and Costa 2005). *Apis mellifera* L. exhibits patterns of biological variation rarely seen in other animal groups, mainly due to the unusually widespread geographical distribution of this species, adapting to quite different environmental conditions (Ruttner 1988; Danforth 2007). This extensive distribution also promotes hybridization where subspecies overlap or come into contact (summarized in De la Rúa et al. 2005b), and it provides a useful means of studying climate-organism interactions (Le Conte and Navajas 2008), as well as temporal changes in distribution over time (natural range shifts: Thomas et al. 2001; Hickling et al. 2006).

The hybrid origin of Iberian honeybees has been identified using morphological and ethological characteristics (Ruttner et al. 1978; Ruttner 1988; Arias et al. 2006; Miguel et al. 2011), allozymes (Smith and Glenn 1995; Arias et al. 2006) and nuclear DNA markers (Franck et al. 1998; Garnery et al. 1998b; Franck et al. 2001; Cánovas et al. 2011; Miguel et al. 2011; Chávez-Galarza et al. 2013), supporting the view that hybridization by secondary contact occurred between north African and western European populations. Mitochondrial DNA diversity also reveals the presence of two evolutionary lineages among Iberian honeybees: African and west European (Smith et al. 1991; Franck et al. 1998, 2001; Garnery et al. 1998a; Miguel et al. 2007; Cánovas et al. 2008; Pinto et al. 2013). The frequencies of these lineages follow a latitudinal cline whereby African evolutionary haplotypes decrease from

Table 1 Variables included in the environmental data set and the individual ranges

Variables	Name	Min/max value	Comp. I	*P*
Annual mean temperature (°C)	*tann*	7.6/18.3	0.71	***
Maximum temperature of the hotest month (°C)	*mxtwm*	22.0/36.3	0.60	***
Minimum temperature of the coldest month (°C)	*mntcm*	−3.6/8.3	0.93	***
Mean temperature in the warmest quarter (°C)	*tmnwq*	14.3/26.9	0.57	***
Mean temperature in the coldest quarter (°C)	*mtcq*	1.2/12.3	0.84	***
Temperature seasonality	*tmseas*	33.42/68.96	−0.09	***
Isothermality (mean diurnal range/annual range)	*isotherm*	30/44	0.96	***
Annual rainfall (mm)	*rfann*	266/1,445	0.12	***
Precipitation in the wettest month (mm)	*rfwm*	37/233	0.29	***
Precipitation in the driest month (mm)	*rfdm*	0/55	−0.81	***
Precipitation in the warmest quarter (mm)	*rfwq*	102/581	0.40	***
Precipitation in the coldest quarter (mm)	*rfcq*	11/189	−0.76	***
Rainfall seasonality	*rfseas*	17/73	0.92	***

Ordination scores and levels of significance for the first component of the CCA analysis using the mitochondrial haplotype distribution are also shown. *P* values were based on 9,999 permutations

Significance codes: 0<"***" < 0.001<"**" < 0.01<"*" < 0.05<"." < 0.1<"NS" < 1

south-west to north-east of the Iberian Peninsula, while the west European ones increase proportionally. This peculiar cline pattern is the result of a new range expansion of the two lineages inhabiting the Iberian Peninsula although their underlying causes remain unclear (Cánovas et al. 2008, 2011). The spatial distribution of mitochondrial diversity appears to be stable despite the recent introduction of intensive management practices, including queen trade, transhumance and artificial selection (Cánovas et al. 2008). Usually, beekeepers from Southern and Eastern Iberia move their colonies to northern regions during spring and summer to take advantage of optimal flowering conditions, resulting in homogenization of the composition of the nuclear genome (Cánovas et al. 2011; Miguel et al. 2011). However, these colonies subsequently return to lower latitudes, conserving the original queens and ultimately preserving the mitochondrial DNA distribution.

Single nucleotide polymorphism (SNP) genome scans suggest selection as an underlying cause of diversity patterns in Iberian honey bees (Chávez-Galarza et al. 2013). It has been assumed that mtDNA evolves according to a neutral model of molecular evolution. However, increasing evidence indicates that this assumption could be often violated, and polymorphism patterns could indicate a recent history of positive selection (Ballard and Kreitman 1995; Dowling et al. 2008). Potential targets of positive selection can be then identified from polymorphism data for recent adaptive fixation events, studying'hitchhiking' by neutral regions at nearby sites (Barrett and Schluter 2007). This approach has been used previously in *A. mellifera*, which experienced novel selection pressures after their recent expansion out of Africa (Whitfield et al. 2006; Zayed and Whitfield 2008).

The considerable amount of genetic information related to Iberian honeybees, which shows a stable spatial distribution of the mitochondrial variation, provide a useful opportunity to test hypotheses about the origin of the latitudinal cline between the two distinct evolutionary units of *A. mellifera*. The aim of our analyses is to estimate the individual effects of each eco-geographical variable that defines the suitability of habitats, distinguishing the effects of the recent human-induced changes from those following natural dynamics. The main hypothesis is that the spatial pattern observed has arisen due to the effect of climatic factors in the process of range expansion of existing lineages in Iberia since the Last Glacial Maximum (LGM), that is, the pattern reflects the environmental tolerance of each lineage.

Materials and methods

Sampling and genetic data

Information of a total of 2,188 Iberian honeybees, belonging to 214 localities were used in this study. Data were reported previously from Spanish locations (for further information see Cánovas et al. 2002, 2008; De la Rúa et al. 1999, 2001, 2004, 2005a). Data provided by other authors were included to better cover the study area: Porto and Evora in Portugal (Garnery et al. 1998a; Franck et al. 2001). A detailed data set of these samples is given in Appendix I (Supplementary Electronic Material).

Genetic data were based on the *Dra*I COI-COII RFLP test, which discriminates between and within honeybee lineages from the polymorphisms of the tRNAleu-COII intergenic mitochondrial DNA region (Garnery et al. 1993). Due to the maternal inheritance of the mitochondrial DNA, which is shared by the entire colony and remains stable during the entire life, 1 specimen was genotyped as a representative of the colony, collecting different colonies at each location.

Table 2 Stages of the stepwise procedure in fitting the GAM model for Iberian populations of *A. mellifera* (percentage of deviance explained: *GCV* generalized cross validation score, *AIC* Akaike information criterion, *COR* correlation coefficient. *P* value of $\chi 2$)

	Desc. dev.	GCV	AIC	*COR*	*P*
Comp I CCA ~ .					
s(rfseas)	70.37	0.30	350.09	0.84	***
s(rfseas) + s(tann)	79.47	0.22	279.29	0.89	***
s(rfseas) + s(tann) + s(rfdm)	83.18	0.19	248.21	0.91	***
s(rfseas) + s(tann) + s(rfdm) + s(mtcq)	90.17	0.13	158.82	0.95	***
Lineage A ~ .					
s(rfseas)	42.87	0.14	187.01	0.66	***
s(rfseas) + s(tann)	43.8	0.14	186.04	0.66	***
s(rfseas) + s(tann) + s(mxtwm)	46.6	0.13	178.04	0.68	***
s(rfseas) + s(tann) + s(mxtwm) + s(mntcm)	48.9	0.13	173.15	0.70	***
Lineage M ~ .					
s(rfseas)	43.31	0.12	155.61	0.66	***
s(rfseas) + s(tann)	50.51	0.11	138.30	0.71	***

Significance codes: 0<"***" < 0.001<"**" < 0.01<"*" < 0.05<"." < 0.1<"NS" < 1

Dependent variables

The genetic data (mitochondrial haplotypes) represented the dependent variables in the analysis. To account for intra- and inter-mitochondrial variation in evolutionary lineages, data were entered into two separate matrices based on the diversity of the mitochondrial haplotypes and the evolutionary lineages. The first diversity matrix represented the variation of the mitochondrial genetic diversity, which was inferred from the abundance of each haplotype (mitochondrial variant), coded as a letter for each lineage followed by a number. This matrix was used as input for the ordination analysis and included all available data. By contrast, the second diversity matrix included the variation in the diversity of the mitochondrial lineage, but deduced from the presence-absence of each evolutionary lineage, coded as African or European.

Independent variables

Climatological variables expected to influence population structuring were represented by sets of estimators pooled into a single matrix. Environmental data were used to describe every 30 × 30 s ($\approx$1 km) unit grid of the study area (Table 1). Estimates of temperature and precipitation were obtained from the World Climatic Atlas (Hijmans et al. 2005), which ensured a single and reliable source of data. This data set also provided estimations of past climate history under two LGM climate scenarios: Community Climate System Model (CCSM; Collins et al. 2006) and Model for Interdisciplinary Research on Climate (MIROC; Hasumi and Emori 2004). For each temperature descriptor, the following seven variables were used: the annual mean or total (respectively); highest monthly maximum mean; lowest monthly minimum mean; mean or total for the driest, wettest, coldest and hottest quarter (abbreviations of these variables are indicated in Table 1). The degree of deviation from the expected temperature was calculated by an index of temperature seasonality (*tseas*), inversely related to the continentality gradient. Six rainfall variables were considered (abbreviations in Table 1). The seasonality of precipitation was defined as the tendency for a region to experience more rainfall in specific months or seasons (Markham 1970). Hypothesized location of glaciated mountains during the Pleistocene in the Iberian Peninsula were also used to create a mask for unsuitable areas, delimiting ice-sheets and tundra (Schütt 2005; González-Trueba et al. 2008).

Before constructing any models, all pairwise correlations between predictors were calculated to remove correlated predictors and avoid any overestimation of the responses. A threshold of 0.79 was considered in previous works to justify the removal of any variables when fitting a model (Lehmann et al. 2002). The absolute values of the correlation in our data set were higher than 0.9 (between *rfann* and *rfwm*), 0.8 (*rfdm* with *mxtwm* and *rfwm*) and 0.7 (between *mxtwm* and *tseas*, *mntcm* and *tseas*, *rfseas* and *rfann*). Therefore, model-fitting was constrained to include only one variable of each cited pair.

Ordination analysis and synthesis map

To first analyze the data and investigate the potential relationships with selected environmental variables, the diversity of mitochondrial DNA at the various sampling sites (first diversity matrix) was spatially clustered using a canonical component analysis (CCA: Angers et al. 1999; Jombart et al. 2009). The "Species" in the ordination analysis were represented by mitochondrial haplotypes at each sampling site, each lineage being coded as "A" for African or "M" for west European, plus a number (see Cánovas et al. 2008). The ordination analysis summarized all the variation in the study area and allocated each site as a study unit. A map was then generated by interpolation of the first principal component value at each sampling site, which synthesized the detail of the spatial patterns (Hanotte et al. 2002). CCA analyses were performed using the "ade4" (Cheesel 1992) and "BiodiversityR" (Kindt and Coe 2005) packages from R statistical software v.2.9.2 (R Development Core Team 2009).

Generalized additive models

GAMs were developed by using (1) the first principal component of the CCA as a dependent variable, and (2) the diversity of each mitochondrial lineage (second diversity matrix). The inclusion of explanatory variables in the models was considered as a smoothed function of the predictor (environmental variable), following a step by step procedure initially with one degree of freedom and with α set at 0.05.

GAM models were validated by examining the proportion of the deviance explained, and by minimizing both generalized cross validation (GCV: Wood 2006; Wood and Augustin 2002) and Akaike information criterion scores (AIC: Venables and Ripley 2004). Standard errors of the smooth terms were computed as the 95 % bootstrap confidence interval based on 1,000 replications (Hastie and Tibshirani 1990). The evaluation was focused on the predictive performance at each site using the correlation coefficients (COR), which were calculated as a Pearson correlation coefficient between the observations in the presence-absence dataset and the model prediction (Elith et al. 2011). This parameter is similar to the areas under the receiver operating characteristic (ROC) curves (AUC) (Fielding and Bell 1997), but it takes into account the variation of the prediction from the observation.

Each prediction map was classified into four habitat suitability classes (unsuitable, marginal, suitable and optimal), using the method described by Boyce et al. (2002) and further developed by Hirzel et al. (2006). The predicted to expected ratio (P/E) was calculated for each habitat suitability class (F_i), and an incremental curve shape and its confidence interval were used to define the boundaries of each class. Spearman rank correlation coefficient was then used to estimate the monotonicity of the curve (continuous Boyce index; Boyce et al. 2002), as well as the adjusted r^2 to evaluate the fit of the linear model to the increments of each suitability class break. An exhaustive search for the best suitability class distribution was performed using interval thresholds at every 0.01 increment of the prediction results.

GAM models were calculated using the "mgcv" package (Wood 2006). An interface between R and GRASS GIS v.6.4.0 (GRASS Development Team 2008) was used for analyses and prediction over the potential distributions of both evolutionary lineages.

Results

Ordination analysis and synthesis map

The first component of the CCA accounted for 64.82 % of the variance in the diversity of the mitochondrial haplotype. A single haplotype from each lineage provided the highest contributions to the gradient described by that component. Temperature descriptors contributed with higher scores to the ordination, although the entire environmental data set adjusted significantly to the component that described the genetic gradient (Table 1). A positive contribution to the gradient was described by the first component for the *mntcm*, *isotherm* and *rfseas* variables (higher than 0.9), followed by *mtcq* (0.84). A negative contribution was detected for two variables, *rfcq* and *rfdm*, while the contribution of *tmseas* was close to 0 (Table 1).

The background of Fig. 1a illustrates the geographical variation for the first correspondence component (CCA). The synthesis map which was obtained by interpolation, followed a gradient that peaks in the south/southwest and west end (darkly-shaded areas belonging to the African lineage), and northwest of the Iberian Peninsula (lightly-shaded areas belonging to the European lineage). That map can be viewed as a cline from the south-west, where localities showed higher number of individuals belonging to the African lineage, to the north-east where the west European lineage was predominant. The transitional area between the two evolutionary lineages becomes narrower in NW Iberia, close to the Atlantic coast.

Modeling

The relationship between smooth terms in the GAM and the principle component scores was plotted in Fig. 2. All the tested variables significantly fitted the GAM terms and the deviance described by the four-variables model increased throughout the add-on procedure (Table 2). Temperature descriptors were directly related to the first component of the CCA, while the relationships of the precipitation descriptors were more complex (non-linear response: Fig. 2). The prediction map was generated using

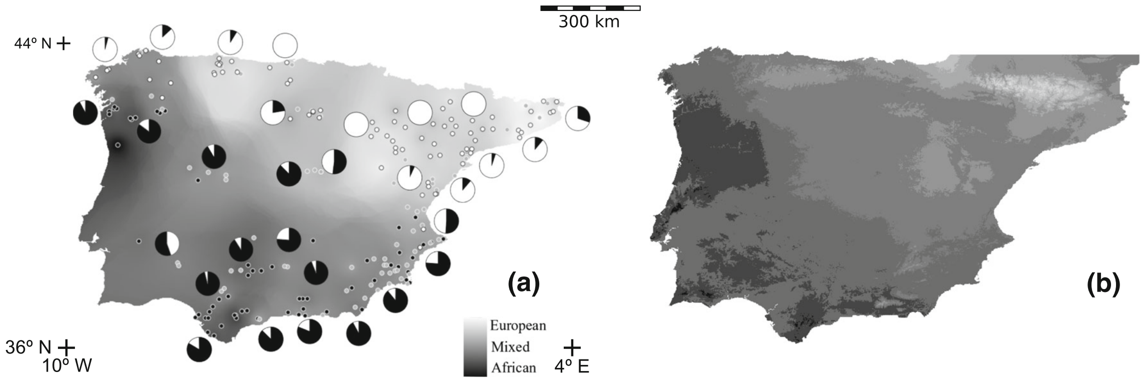

Fig. 1 **a** Localities included in this study (*dots*) and the proportion of haplotypes belonging to each evolutionary lineage (*fill color for dots* following the *grey scale bar*; Franck et al. 1998, 2001; Cánovas et al. 2008). The background follows the *grey scale bar* and illustrates the geographical variation for the first correspondence component (CCA), which accounted for 64.82 % of the total mtDNA haplotype variation in *A. mellifera* across the Iberian Peninsula. Pie charts show the proportion of haplotypes belonging to each evolutionary lineage by Province in Appendix I (*black* for African and *white* for European). **b** Prediction map of the GAM fitted model for the first component values of the ordination analysis (Table 2). The first component follows a gradient that peaks (*darkly-shaded* areas corresponding to the *grey scale bar*) in the south, and that is probably correlated with known waves of mitochondrial colonization of the African lineage from North Africa (Franck et al. 1998, 2001)

the four-variables model (Table 2) and it retained the main features of the haplotype gradient between both evolutionary lineages (Fig. 1b): extremes of the gradient and extension of the transitional area.

GAM models were also fitted to independently analyze the relationships between the environmental data set and the presence/absence of both African and European evolutionary lineages (Table 2; Fig. 3). In both cases, two of the variables (*tann* and *rfseas*) were the same as those fitted by the GAM model for the ordination component, although the proportion of the deviance explained was significantly lower (Table 2). Those two variables showed symmetrical responses between lineages, implying distinct ecological optima for each lineage (Fig. 3). The final GAM model for the European lineage fitted fewer variables than the one for the African lineage, although they show similar explained deviance and significance (Table 2). The estimates of accuracy are shown in Table 2 for all the models in both evolutionary lineage predictions.

Lineage prediction maps were calculated for the entire Iberian Peninsula and classified by using the continuous Boyce index (Fig. 4). GAM models predicted large areas of optimal occurrence for each lineage: these were located in the south–west (African lineage) and north–east (west European lineage) of Iberia (Fig. 4a, b). Suitable areas were estimated by the sum of suitable and optimal classified areas for each lineage. African lineage showed slightly weaker occupancy in the final models (west European lineage: GAM = 429,916 km^2; and African lineage: GAM = 402,945 km^2).

Past predictions during LGM

Lineage predictive models showed slight differences between both climate scenarios (CCSM and MIROC; Fig. 5). African lineage predictions were classified, showing a reduced area for the optimal habitat during the LGM: two patches located at the southwest of the Iberian Peninsula (Fig. 5a, c). West European lineage predictions described almost the entire area of study in the optimal category, except for a small patch located close to the Strait of Gibraltar which were classified as suitable (Fig. 5b, d).

Discussion

The availability of accurate genetic data for Iberian honeybees (Franck et al. 1998, 2001; Cánovas et al. 2008) allowed us to test the spatial relationships using presence-absence modeling. The combination of two different techniques (ordination and GAM) is preferable, as the limitations of one approach can be compensated by the strengths of another.

Lineages' distribution and habitat preferences

The first component of the multivariate analysis based on haplotype frequencies, described the gradient between the African and west European evolutionary lineages of the Iberian honeybees in detail for the first time. Its spatial interpretation predicted a wide inter-gradation area, which crosses the Iberian Peninsula from northern Portugal to southeastern Spain. The lineage gradient distribution

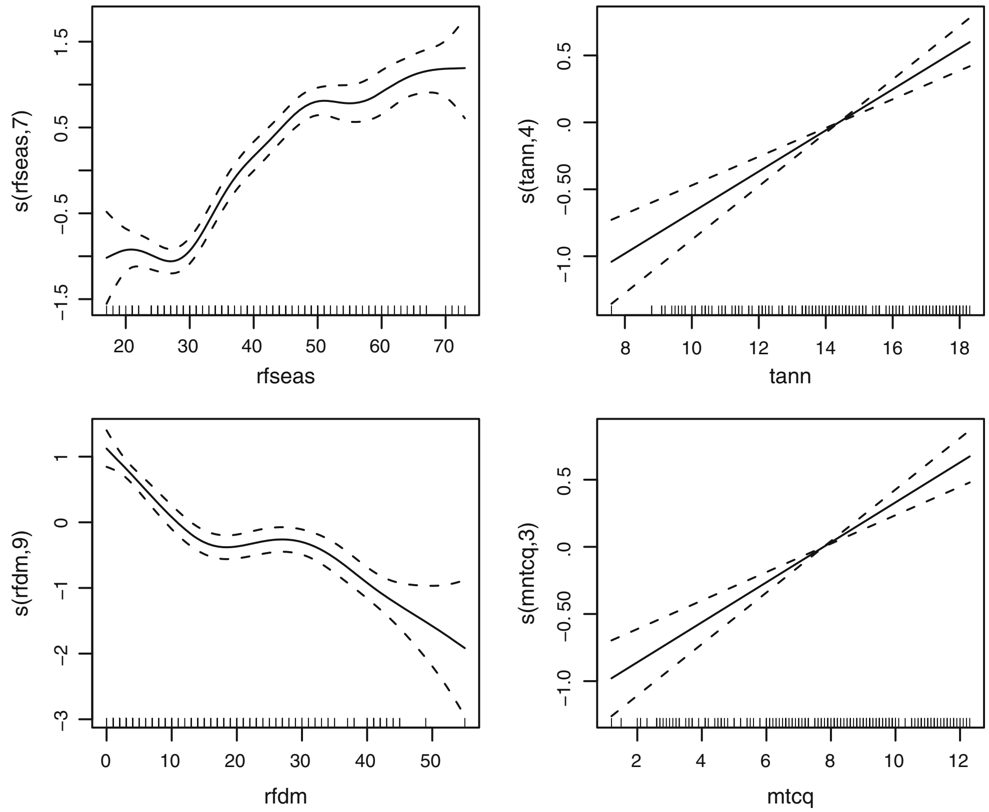

Fig. 2 Fitted smooth terms (shown as s[name of the predictor, number degrees of freedom]) for the first correspondence component derived from the abundance of mitochondrial haplotypes versus variables added to the final GAM model, by a stepwise procedure (*solid lines*), and ±95 % boot-strapped confidence intervals (*dashed lines*). The *vertical axes* represent the variable contribution of rainfall seasonality (*rfseas*) and rainfall in the driest month (*rfdm*), and annual (*tann*) and coldest quarter mean (*mtcq*) temperatures (expressed on the scale of the linear predictor). See Table 2 for details

revealed a complex pattern, also fitting previous and detailed information from the Atlantic coast (Pinto et al. 2013). That spatial distribution fitted the environmental dataset (Table 1), as the cline between the west European and African evolutionary lineages appears directly related to the minimum temperatures in the study area although it is marked by day-night rather than summer–winter oscillations (isothermality). The last estimate, isothermality, varies from 0 to 100, which describes a site where the diurnal temperature range is equivalent to the annual temperature range. The distribution of the African lineage showed values above 40, indicating that the diurnal temperature range represents over 40 % of the annual temperature range. This finding is perhaps related to the ability of this lineage to survive in a variety of sub-optimal environmental conditions. The gradient was also indirectly related to precipitation descriptors and it was marked by seasonality.

Suitability of modelling approaches

Ordination together with GAM appeared to be an accurate approach, constituting a single model that summarizes and describes the transition as it occurs naturally, as a continuum between both evolutionary lineages (see discussion below). Moreover, with this approach there are no a priori assumptions regarding the relationships between external variables and the ordination configuration. A four-variables GAM model significantly explained about 90 % of the variance, although the model fits more complex polynomial functions for the rainfall descriptor. GAMs also resulted in appropriate models for each lineage separately, as expected from species modeling with complex distribution patterns and the availability of presence-absence data (Segurado and Araujo 2004). Precipitation relationships were more complex in those models too. Distributions were limited by the seasonality of the precipitation, indicating that rainfall tends to increase in certain months or seasons where the African lineage can be found. The European lineage also shows a pattern linked to wetter areas, inversely related to the rainfall of the driest month. Temperature was linearly correlated with the lineage distribution; it was directly related to the African lineage and inversely related to the west European lineage.

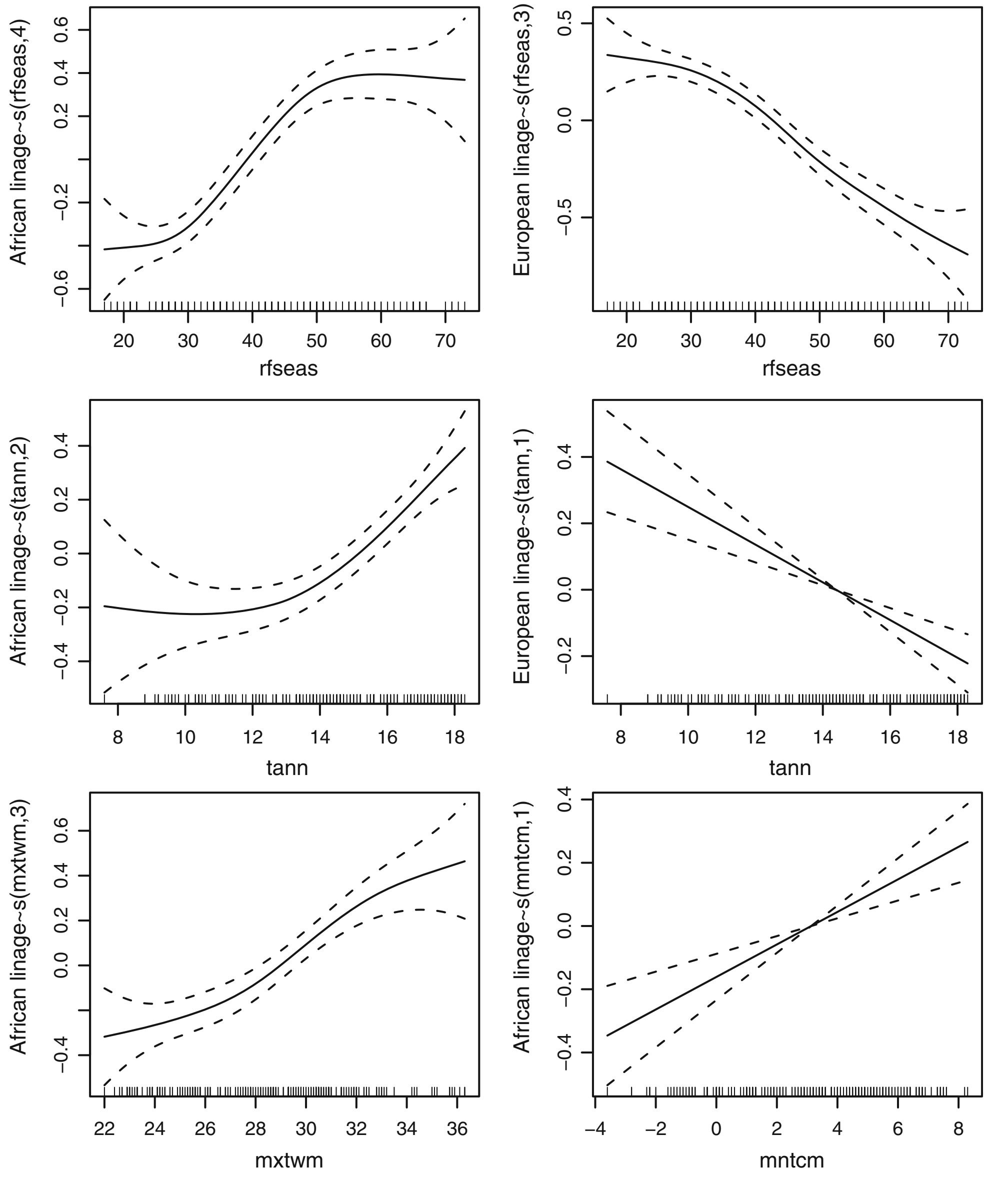

Fig. 3 Fitted smooth terms (indicated as s[name of the predictor, number degrees of freedom]), for the presence-absence of each mitochondrial evolutionary lineage versus the four variables added to the final GAM models by a stepwise procedure (*solid lines*), and ±95 % boot-strapped confidence intervals (*dashed lines*). The *vertical axes* represent the variable contribution of rainfall seasonality (*rfseas*), annual temperature (*tann*) and maximum temperature during the hottest month (*mtcq*) relative to the component (expressed on the scale of the linear predictor). See Table 2 for details

A contact area between evolutionary lineages: an ecological perspective

Those results suggest two complementary distributions at the edges of each lineage, although both lineages are widely distributed across the Iberian Peninsula. Further studies will be necessary to determine whether the current diversity distribution is maintained by other isolation processes (e.g., competition driven by selection, unsynchronized reproductive cycles).

On one hand, a suite of theoretical and empirical studies have focused on the evolution of certain characteristics of social insects, and the underlying adaptive mechanisms in honeybee phenology and behavior (see extended bibliography in Gordo and Sanz 2006; El-Niweiri and Moritz 2011). Phenology is significantly affected by temperature; honeybees appear sooner in the year in southern localities and at lower altitudes, and earlier in years with warmer springs (Gordo and Sanz 2005). Such behavior can be simply explained by honeybees exiting the hive in response

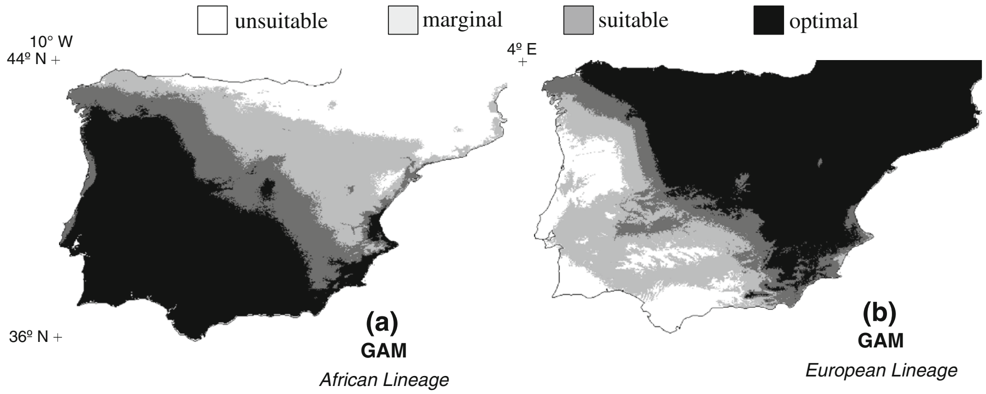

Fig. 4 Predicted potential geographic distributions for both African (**a**) and European (**b**) mitochondrial evolutionary lineages of *Apis mellifera* in the Iberian Peninsula. Predicted occurrence probabilities for each evolutionary lineage range from 0 (*lightest areas*) to 1 (*darkest areas*). The maps were reclassified in four habitat suitability classes (unsuitable, marginal, suitable, optimal: Boyce et al. 2002; Hirzel et al. 2006). See also details about model performance in Table 2

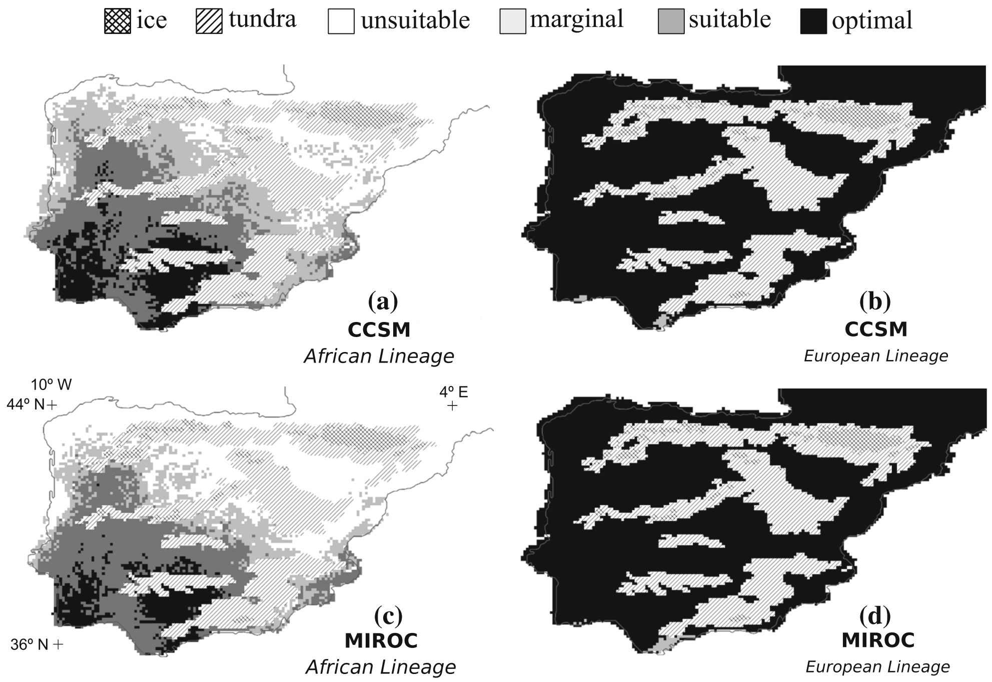

Fig. 5 Simulated potential geographic distributions for both African and European mitochondrial evolutionary lineages of *A. mellifera* in the Iberian Peninsula under Community Climate System Model (CCSM; Collins et al. 2006) and Model for Interdisciplinary Research on Climate (MIROC; Hasumi and Emori 2004) climate scenarios from the LGM. Maps were reclassified as habitat suitability classes (unsuitable, marginal, suitable, optimal: Boyce et al. 2002; Hirzel et al. 2006). Hypothesized location of glaciated mountains during the Pleistocene in the Iberian Peninsula is also shown (Schütt 2005; González-Trueba et al. 2008)

to increasing temperatures. Modeling results support these observations, as the suboptimal area with higher temperatures and rainfall concentrated in particular months (typical characteristics of Mediterranean climate) is preferred by the A lineage, which is closer to the ancestral lineages originated in Africa than the M lineage. Southern Iberian honeybee populations also show behavioral characteristics of African subspecies, *A. m. intermissa* and *A. m. adansonii*, such as the willingness to forage at low temperatures, quick defense reaction on the comb and propensity to swarm (Ruttner 1988; Engel 1999). In contrast, a flexible thermal strategy allows west European honeybees of the M lineage to regulate their head temperature at low environmental temperatures, guaranteeing proper suction, reducing energy costs and improving efficiency (Kovac et al. 2010). The predicted areas of distribution of this lineage support these observations, since the conditions in suitable and optimal habitats were cloudy with fewer areas of solar radiance, and precipitation more homogeneously distributed throughout the year and in higher amount.

On the other hand, the introduction of intensive practices (mainly transhumance and queen trade) and their consequences (homogenization of the nuclear genome) will lead to the disappearance of the gradient originated by biological mechanisms that specifically prevent mitochondrial homogeneization (see introduction and Cánovas et al. 2011).

Changing climate: genetic imprints of colonization and expansion

A reconstruction of the relationships between eco-geographical variables and genetic gradients provided valuable information about habitat conditions and the transition area between the A and M lineages, which covers almost the entire Iberian Peninsula. However, the extension or latitudinal position of transition areas might be modified in the future by environmental (climate change) or man-made (management) changes, which can critically affect the diversity of populations (Le Conte and Navajas 2008).

Mitochondrial haplotypes with the highest contribution to the canonical component analysis were those ones considered ancestral in previous studies (A2 and M4: Miguel et al. 2007; Cánovas et al. 2008). Therefore, our modeling results support the hypothesis of the natural origin of the genetic gradient studied here. It should be noted that the optimal habitat of the African lineage differs from the average conditions found in the Iberian Peninsula, particularly in the past (the predicted area occupied by the African lineage is smaller than that of the European one during LGM conditions). These results suggest a northwards expansion of the African lineage inhabiting the southern part of the Iberian Peninsula, and a simultaneous colonization of west Europe by populations of northern Iberia belonging to the M lineage since the LGM, as hypothesized by several authors (Franck et al. 1998; Garnery et al. 1998b; Miguel et al. 2007; Cánovas et al. 2008; Pinto et al. 2013).

Conclusions

The study of changes in insect distribution has attracted increased attention, mainly due to the strong influence of climate, which allows biogeographical hypotheses to be tested through environmental modeling. The variability of suitable honeybee habitats, in terms of temperature and precipitation, reveals that current distribution of the two main evolutionary lineages of the honeybee in the Iberian Peninsula shows a high degree of plasticity despite the two distinct climatic optima identified. The association of that distribution pattern with habitat differences and genotypic variations, suggests that selection could play an important role in shaping the life-history evolution of this species, particularly affecting the mitochondria. Such evolutionary process could be also resulting in strong hitchhiking effects on particular regions of the mitochondrial genome. The present study highlights the importance and utility of modeling approaches to evaluate and interpret genetic variation, and in turn, to develop and improve conservation programs, using the honeybee *A. mellifera* as a model to predict the consequences of climatic change.

Acknowledgments We thank all the bee-keepers and associations (particularly Coordinadora de Organizaciones de Agricultores y Ganaderos, COAG) that generously contributed samples and their time. We thank O. Rodríguez, J. Ornia, E. Muñiz and U. Mediel. Thanks to R. Angus, who helped us improve the manuscript with their comments and suggestions during the revision process. This work was supported by projects RZ00-013 (Spanish National Institute of Agricultural Investigation, INIA) and EVK2 2000-00628 (BABE, V European Union Framework). F. Cánovas is supported by a postdoctoral fellowship from the FCT (Fundacão para Ciência e Tecnologia, Portugal). This study was conducted entirely using open source software.

References

Angers B, Magnan P, Plantes M, Bernatchez L (1999) Canonical correspondence analysis for estimating spatial and environmental effects on microsatellite gene diversity in brook charr (*Salvelinus fontinalis*). Mol Ecol 8:1053–1403

Arias M, Rinderer T, Sheppard W (2006) Further characterization of honey bees from the Iberian Peninsula by allozyme, morphometric and mtDNA haplotype analyses. J Apic Res 45(4):188–196

Ballard W, Kreitman M (1995) Is mitochondrial DNA a strictly neutral marker? Trends Ecol Evol 10:485–488

Barrett R, Schluter D (2007) Adaptation from standing genetic variation. Trends Ecol Evol 23:38–44

Boyce M, Vernier P, Nielsen S, Schmiegelow F (2002) Evaluating resource selection functions. Ecol Model 157:281–300

Cánovas F, De la Rúa P, Serrano J, Galián J (2002) Variabilidad del ADN mitocondrial en poblaciones de *Apis mellifera iberica* de Galicia (NW España). Arch Zootec 51:441–448

Cánovas F, De la Rúa P, Galián J, Serrano J (2008) Geographical patterns of mitochondrial DNA variation in *Apis mellifera iberiensis* (Hymenoptera: Apidae). J Zool Syst Evol Res 46:24–30

Cánovas F, De la Rúa P, Serrano J, Galián J (2011) Microsatellite variability reveals beekeeping influences on Iberian honeybee populations. Apidologie 42:235–251

Chávez-Galarza J, Henriques D, Johnston JS, Azevedo JC, Patton JC, Muñoz I, De la Rúa P, Pinto MA (2013) Signatures of selection in the Iberian honey bee (*Apis mellifera iberiensis*) revealed by a genome scan analysis of single nucleotide polymorphisms. Mol Ecol 22(23):5890–5907

Cheesel D (1992) The ade4 package-I: one-table methods. R News 4:5–10

Collins WD, Bitz CM, Blackmon ML, Bonan GB, Bretherton CS, Carton JA, Chang P, Doney SC, Hack JJ, Henderson TB, Kiehl

JT, Large WG, McKenna DS, Santer BD, Smith RD (2006) The community climate system model version 3 (CCSM3). J Clim 19:2122–2143

Danforth B (2007) Bees. Curr Biol 17:156–161

De la Rúa P, Galián J, Serrano J (1999) Variabilidad del ADN mitocondrial en poblaciones de abejas de la miel (*Apis mellifera* L.) de la Región de Murcia. Invest Agric Prod San Anim 14:41–49

De la Rúa P, Galián J, Serrano J, Moritz RFA (2001) Molecular characterization and population structure of the honeybees from the Balearic Islands (Spain). Apidologie 32:417–427

De la Rúa P, Jiménez Y, Galián J, Serrano J (2004) Evaluation of the biodiversity of honey bee (*Apis mellifera*) populations from eastern Spain. J Apic Res 43:162–166

De la Rúa P, Hernández-García R, Jiménez Y, Galián J, Serrano J (2005a) Biodiversity of *Apis mellifera* iberica (Hymenoptera: Apidae) from northeastern Spain assessed by mitochondrial analysis. Insect Syst Evol 36:21–28

De la Rúa P, Fuchs S, Serrano J (2005b) Biogeography of European honey bees. In: Lodesani M, Costa C (eds) Beekeeping and conserving biodiversity of honeybees. Sustainable bee breeding, theoretical and practical guide. Northern Bee Books, Hebden Bridge, pp 15–52

Dowling D, Friberg U, Lindell J (2008) Evolutionary implications of non-neutral mitochondrial genetic variation. Trends Ecol Evol 23:546–554

Elekonich M, Roberts S (2005) Honey bees as a model for understanding mechanisms of life history transitions. Comp Biochem Physiol A 141:362–371

Elith J, Phillips S, Hastie T, Dudík M, Chee Y, Yates CJ (2011) A statistical explanation of MaxEnt for ecologists. Divers Distrib 17(1):43–57

El-Niweiri M, Moritz R (2011) Mating in the rain? Climatic variance for polyandry in the honeybee (*Apis mellifera jemenitica*). Pop Ecol 53(3):421–427

Engel MS (1999) The taxonomy of recent and fossil honey bees (Hymenoptera: Apidae; *Apis*). J Hymenopt Res 8(2):165–196

Fielding A, Bell J (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. Environ Conserv 24:38–49

Franck P, Garnery L, Solignac M, Cornuet J (1998) The origin of West European subspecies of honeybees (*Apis mellifera*): new insights from microsatellite mitochondrial data. Evolution 52(4):1119–1134

Franck P, Garnery L, Loiseau A, Hepburn HR, Solignac M, Cornuet JM (2001) Genetic diversity of the honeybee in Africa: microsatellite mitochondrial data. Heredity 86:420–430

Garnery L, Solignac M, Celebrano G, Cornuet JM (1993) A simple test using restricted PCR-amplified mitochondrial DNA to study the genetic structures of *Apis mellifera* L. Experientia 49:1016–1021

Garnery L, Frank P, Baudry E, Vautrin D, Cornuet J, Solignac M (1998a) Genetic diversity of the west European honeybee (*Apis mellifera mellifera* and *A. m. iberica*). I. Mitochondrial DNA. Genet Sel Evol 30(1):31–47

Garnery L, Frank P, Baudry E, Vautrin D, Cornuet J, Solignac M (1998b) Genetic diversity of the west European honeybee (*Apis mellifera mellifera A. m. iberica*). II. Microsatellite loci. Genet Sel Evol 30(1):49–74

González-Trueba J, Martín-Moreno R, de Pisón EM, Serrano E (2008) 'Little Ice Age' glaciation and current glaciers in the Iberian Peninsula. Holocene 18:551–568

Gordo O, Sanz J (2005) Phenology and climate change: a long- term study in a Mediterranean locality. Oecologia 146:484–495

Gordo O, Sanz J (2006) Temporal trends in phenology of the honey bee *Apis mellifera* (L.) and the small white *Pieris rapae* (L.) in the Iberian Peninsula (1952–2004). Ecol Entomol 31:261–268

GRASS Development Team (2008) Geographic resources analysis support system (GRASS GIS) software. Open source geospatial foundation. http://grass.osgeo.org

Hanotte O, Bradley D, Ochieng J, Verjee Y, Hill E, Rege EO (2002) African pastoralism: genetic imprints of origins and migrations. Science 296:336–339

Hastie T, Tibshirani R (1990) Generalized additive models. Chapman and Hall, London

Hasumi H, Emori S (2004) K-1 coupled GCM (MIROC) description. Center for Climate System Research, University of Tokyo, Tokyo p 34

Hickling R, Roy D, Hill J, Fox R, Thomas C (2006) The distributions of a wide range of taxonomic groups are expanding polewards. Glob Chang Biol 12:450–455

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

Hirzel A, Le Lay G, Helfer V, Randin C, Guisan A (2006) Evaluating the ability of habitat suitability. Ecol Model 199:142–152

Jombart T, Pontier D, Dufour A (2009) Genetic markers in the playground of multivariate analysis. Heredity 102(4):330–341

Kindt R, Coe R (2005) Tree diversity analysis: a manual and software for common statistical methods for ecological and biodiversity studies. World Agroforestry Centre, Nairobi

Kovac H, Stabentheiner A, Schmaranzer S (2010) Thermoregulation of water foraging honeybees-balancing of endothermic activity with radiative heat gain and functional requirements. J Insect Physiol 56(12):1834–1845

Le Conte Y, Navajas M (2008) Climate change: impact on honey bee populations and diseases economically valuable species. Rev Sci Technol 27(2):499–510

Lehmann A, Overton JM, Leathwik JR (2002) GRASP: generalized regression analysis and spatial predictions. Ecol Model 157:187–205

Lodesani M, Costa C (2005) Beekeeping and conserving biodiversity of honeybees. Sustainable bee breeding, theoretical and practical guide. Northern Bee Books, Hebden Bridge

Markham C (1970) Seasonality of precipitation in the United States. Ann Assoc Amer Geograph 60:593–597

Miguel I, Iriondo M, Garnery L, Sheppard WS, Estonba A (2007) Gene flow within the M evolutionary lineage of *Apis mellifera*: role of the Pyrenees, isolation by distance and post-glacial recolonization routes in the Western Europe. Apidologie 38:141–155

Miguel I, Baylac M, Iriondo M, Manzano C, Garnery L, Estonba A (2011) Both geometric morphometric and microsatellite data consistently support the differentiation of the *Apis mellifera* M evolutionary branch. Apidologie 42(2):150–161

Pinto MA, Henriques D, Neto M, Guedes H, Muñoz I, Azevedo J, de la Rúa P (2013) Maternal diversity patterns of Ibero-Atlantic populations reveal further complexity of Iberian honeybees. Apidologie 44:430–439

R Development Core Team (2009) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/. ISBN 3-900051-07-0

Ruttner F (1988) Biogeography and taxonomy of honeybees. Springer, Berlin

Ruttner F, Tassencourt L, Louveaux J (1978) Biometrical-statistical analysis of the geographic variability of *Apis mellifera* L. Apidologie 9(4):363–381

Schütt B (2005) Late quaternary environmental change on the Iberian Peninsula. Erde 136:3–14

Segurado P, Araujo MB (2004) An evaluation of methods for modelling species distributions. J Biogeogr 31:1555–1568

Smith D, Glenn T (1995) Allozyme polymorphisms in Spanish honeybees (*Apis mellifera iberica*). J Hered 86:12–16

Smith D, Palopoli M, Taylor B, Garnery L, Cornuet J, Solignac M, Brown WM (1991) Geographical overlap of two mitochondrial genomes in Spanish honeybees (*Apis mellifera iberica*). J Hered 82:96–100

Thomas C, Bodsworth E, Wilson R, Simmons A, Davies Z, Conradt L (2001) Ecological and evolutionary processes at expanding range margins. Nature 411:577–581

Venables W, Ripley B (2004) Modern applied statistics with S. Springer, New York

Whitfield C, Behura S, Berlocher S, Clark A (2006) Thrice out of Africa : ancient and recent expansions of the honey bee, *Apis mellifera*. Science 314:642–645

Wood S (2006) Generalized additive models: an introduction with R. Chapman and Hall, London

Wood S, Augustin N (2002) GAMs with integrated model selection using penalized regression splines and applications to environmental modelling. Ecol Model 157:157–177

Zayed A, Whitfield C (2008) A genome-wide signature of positive selection in ancient and recent invasive expansions of the honey bee *Apis mellifera*. Proc Natl Acad Sci USA 105:3421–3426