MOLECULAR ECOLOGY

Molecular Ecology (2011) **20**, 193–205 doi: 10.1111/j.1365-294X.2010.04936.x

Multiple approaches to detect outliers in a genome scan for selection in ocellated lizards (*Lacerta lepida*) along an environmental gradient

VERA L. NUNES,* MARK A. BEAUMONT,†‡ ROGER K. BUTLIN§ and OCTÁVIO S. PAULO*

**Computational Biology and Population Genomics Group, Centro de Biologia Ambiental, DBA/FCUL, 1749-016 Lisboa, Portugal, †School of Animal and Microbial Sciences, University of Reading, PO Box 228, Reading RG6 6AJ, UK, ‡Schools of Mathematics and Biological Sciences, University of Bristol, Bristol, UK, §Department of Animal and Plant Sciences, University of Sheffield, Sheffield S10 2TN, UK*

Abstract

Identification of loci with adaptive importance is a key step to understand the speciation process in natural populations, because those loci are responsible for phenotypic variation that affects fitness in different environments. We conducted an AFLP genome scan in populations of ocellated lizards (*Lacerta lepida*) to search for candidate loci influenced by selection along an environmental gradient in the Iberian Peninsula. This gradient is strongly influenced by climatic variables, and two subspecies can be recognized at the opposite extremes: *L. lepida iberica* in the northwest and *L. lepida nevadensis* in the southeast. Both subspecies show substantial morphological differences that may be involved in their local adaptation to the climatic extremes. To investigate how the use of a particular outlier detection method can influence the results, a frequentist method, DFDIST, and a Bayesian method, BayeScan, were used to search for outliers influenced by selection. Additionally, the spatial analysis method was used to test for associations of AFLP marker band frequencies with 54 climatic variables by logistic regression. Results obtained with each method highlight differences in their sensitivity. DFDIST and BayeScan detected a similar proportion of outliers (3–4%), but only a few loci were simultaneously detected by both methods. Several loci detected as outliers were also associated with temperature, insolation or precipitation according to spatial analysis method. These results are in accordance with reported data in the literature about morphological and life-history variation of *L. lepida* subspecies along the environmental gradient.

Keywords: adaptive divergence, AFLP, environmental gradient, landscape genetics, natural selection, outlier loci

Received 16 September 2009; revision received 7 October 2010; accepted 14 October 2010

Introduction

Uncovering the genetic basis of important adaptive traits in natural populations is a major goal for evolutionary biology in order to better understand how populations adaptively diverge in heterogeneous environments and eventually give rise to new species. It is now well accepted that differentiation can occur in the presence of gene flow if adaptively driven (Schluter 2009). This mode of speciation is likely to produce genomically heterogeneous divergence, unlike the classic view of allopatric speciation, where the whole genome should behave as a cohesive unit in the development of reproductive isolation (Wu 2001; Nosil *et al.* 2009; Via 2009). When populations face different environments, ecologically based selection can arise, leading to local adaptations. These may select for assortative mating or other pre-zygotic forms of isolation, reducing and perhaps eventually eliminating gene flow between the emerging species (Gavrilets 2004).

Correspondence: Vera L. Nunes, Fax: +351 217500028;
E-mail: vlnunes@fc.ul.pt

Identifying which regions of the genome are under selection and understanding what selective pressures are acting upon natural populations is challenging, in particular for non-model species. Although searching for selection signatures in candidate genes for adaptive traits has been fruitful in some studies (Rosenblum *et al.* 2004; Hoekstra *et al.* 2006), this kind of approach depends on the availability of well-characterized genes and is limited to well-studied metabolic pathways that underlie measurable phenotypic traits. Association of phenotypes with genotypes by quantitative trait loci analysis also relies on well-studied species, which are easy to manipulate and cross (Stinchcombe & Hoekstra 2008), although future prospects in this field are promising (see Ellegren 2008). In the last few years, population genomic approaches have become accessible and popular for searching for genes influenced by selection, even in non-model organisms (Black *et al.* 2001; Luikart *et al.* 2003). In particular, for species with little genetic information, scans using AFLP markers have become a fast and economic tool to survey several hundreds of random loci across the whole genome (Bensch & Akesson 2005), although a full coverage of the species' genome can hardly be attained with this kind of approach. Demography and the neutral evolutionary history of populations affect neutral loci across the genome in the same way, while loci under selection and closely linked loci will exhibit an outlier pattern of variation (Luikart *et al.* 2003). This strategy has been successfully applied in natural populations to detect candidate loci underlying adaptation to altitude (Bonin *et al.* 2006) or temperature (Jump *et al.* 2006), or to investigate ecotype-based differentiation (Wilding *et al.* 2001; Campbell & Bernatchez 2004; Egan *et al.* 2008; Herrera & Bazaga 2008; Nosil *et al.* 2008) and species boundaries (Murray & Hare 2006; Savolainen *et al.* 2006; Minder & Widmer 2008). However, as well as identifying outliers as candidate loci under selection, it is crucial to investigate which selective forces are acting upon them (Joost *et al.* 2007) and disentangle the possible functional roles of outlier loci (Vasemagi & Primmer 2005; Jensen *et al.* 2007). As genome scans by AFLPs rely on anonymous markers, follow-up analyses are needed to identify and characterize those loci. This task may still be challenging in non-model organisms, but some studies have already accomplished it (Minder & Widmer 2008; Wood *et al.* 2008).

The species studied here, *Lacerta lepida*, is the only species of ocellated lizard occurring in Europe. It is also designated *Timon lepidus* since Mayer & Bischoff (1996) upgraded the subgenus *Timon* (we will continue to use the previous designation hereafter, as it is still the more widely used). The species' distribution is limited to the Iberian Peninsula and some regions in the South of France and North of Italy (Fig. 1). The species occurs in a wide set of environmental conditions. Mateo (1988) emphasized the coincidence of a climatic gradient running from southeast (SE) to northwest (NW) across the Iberian Peninsula with the major morphological differences observed among populations, coinciding also with the geographical distribution of three currently

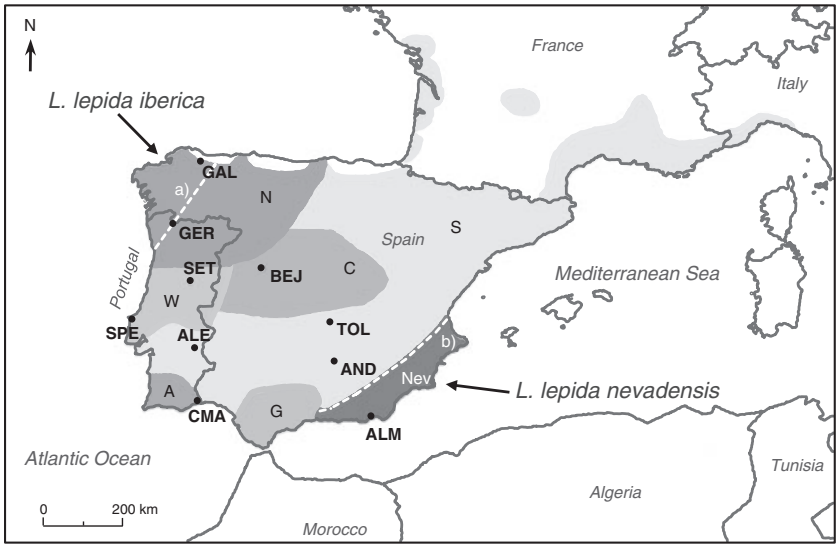

Fig. 1 Map of the western Mediterranean region showing the distribution area of *Lacerta lepida* in Europe. Distribution limits of *L. lepida* subspecies based on morphology are indicated by dashed lines: (a) *L. l. iberica*; (b) *L. l. nevadensis*. *L. l. lepida* occupies the remaining shaded area in Iberian Peninsula, France and Italy. Predicted geographical distribution of each mitochondrial clade (cyt *b*) (Paulo 2001; Paulo *et al.* 2008) is represented by different grey shades. Sampled populations are represented by black dots (codes as in Table 1).

accepted subspecies. The nominal subspecies, *L. lepida lepida*, occurs in the typical Mediterranean climate that characterizes most of the distribution area. *L. lepida nevadensis* occupies the most arid areas in the SE along the Mediterranean Sea coast and *L. lepida iberica* is limited to the occidental coast of Galicia, in the NW, characterized by a rainy Atlantic climate (Mateo & Castroviejo 1990). Several morphological traits show clinal variation, including average body size (decrease in size towards NW), colour pattern (increase of black scales in the dorsum in NW and their replacement by brown scales in SE) and dentition (morphological differentiation among teeth increases but their number decreases, both from NW to SE) (Mateo 1988; Mateo & Castroviejo 1990; Mateo & López-Jurado 1994). Some karyological (Mateo *et al.* 1999), allozyme (Mateo *et al.* 1996) and even reproductive strategy (Castilla & Bauwens 1989; Mateo & Castanet 1994) differences between subspecies have also been recorded. Neither the genetic basis nor the effect on fitness have yet been investigated for the morphological traits that show clinal variation, but some of these morphological differences may lead us to suspect selective forces shaping each subspecies' local adaptation to major climatic and ecological differences (the latter being in turn influenced by climatic conditions, such as vegetation cover, food availability and thermoregulatory conditions). This effect may be more dramatic at the opposite extremes of the cline, where *L. l. iberica* and *L. l. nevadensis* replace the nominal and widespread subspecies, *L. l. lepida*.

Studies conducted on *L. lepida* with the cytochrome *b* mitochondrial gene (Paulo 2001) reveal this species as highly structured (seven clades), with strong geographical subdivisions and limited gene flow. The separation of *L. l. nevadensis* (clade Nev), with about 13% mtDNA sequence divergence, is the oldest split in the Iberian Peninsula, estimated to have occurred about 9.43 million years ago (Ma), while subspecies *L. l. iberica* (clade N), with <3% mtDNA sequence divergence from the nominal subspecies, is estimated to be derived from a Plio-Pleistocenic splitting event, corresponding to a much more recent differentiation, around 2 Ma (Paulo *et al.* 2008).

The genetic and geographical structure of ocellated lizards in the Iberian Peninsula probably reflects the existence of multiple refugia during the last glaciations, which would have suffered several demographic contractions and expansions through the successive climatic oscillations (Paulo *et al.* 2008). Given the unique geographical characteristics of the Iberian Peninsula, the severity of the climatic changes affecting each population would depend on the latitude, the topography and the influence of Atlantic Ocean or Mediterranean Sea proximity. Therefore, ocellated lizards from NW and SE extremes of the Iberian Peninsula evolved under different environmental conditions.

Here, we present the results of a genome scan using AFLPs, with populations collected along the SE–NW environmental gradient in the Iberian Peninsula, to detect candidate loci influenced by selection. Several methods for outlier detection are now available, but the consequences for the results of choosing a particular detection method have not been fully investigated. If, as in this case, the purpose of the study is to detect a set of candidate loci influenced by selection for future confirmatory investigations, it is important that by choosing a specific detection method, the results will not be dramatically different from other methods. Here, a frequentist method, DFDIST (Beaumont & Nichols 1996), and a Bayesian method, BayeScan (Foll & Gaggiotti 2008), were used for outlier detection. Results from both methods were compared, assessing the advantages and limitations of the use of each method. To investigate which selective pressures may be acting upon ocellated lizards along the climatic gradient, associations of AFLP band frequencies with climatic variables, such as temperature or precipitation, were tested by logistic regression as implemented in the spatial analysis method (SAM) (Joost *et al.* 2008). By comparing and combining results from the three methods, we aimed to obtain a list of candidate loci for further investigation of ocellated lizard's adaptation to different environments along the Iberian Peninsula.

Materials and methods

Sampling and DNA isolation

A total of 10 populations of *Lacerta lepida* were sampled in the Iberian Peninsula along a SE–NW transect and in a north–south transect along the Atlantic coast, covering the distribution of the three subspecies and all cyt *b* mitochondrial clades, except the Gualdaquivir clade (Paulo 2001) (Fig. 1). Population locations, sample sizes and corresponding mtDNA clades are listed in Table 1. In each location, tissue samples from tail were collected from free-living adult lizards from both sexes, in approximately equal proportions. The animals were immediately released back into the wild. Whole genome DNA was extracted from tail tissue using the Jetquick Tissue DNA kit (Genomed).

Environmental data

At least one measure of GPS coordinates was recorded in the field for each sampled location. Lizards were collected no more than 3 km away from the GPS coordinates. Climatic data was obtained from public

Table 1 Populations used in the genome scan. Sample sizes (N) and mitochondrial clade (cyt *b*) (Paulo 2001; Paulo *et al.* 2008) are indicated for each location. Values of annual mean temperature (Temp), annual precipitation (Prec), annual insolation (Ins) and annual relative humidity (Hum) are also presented

Population	Code	Latitude	Longitude	Taxa	N	mtDNA clade (Cyt *b*)	Temp (°C)	Prec (mm)	Ins (h)	Hum (%)
Galicia	GAL	43° 21′ 60″ N	7° 22′ 03″ W	*L. l. iberica*	19	N (Northern)	10.6	1233	1800	75
Gerês	GER	41° 43′ 23″ N	8° 06′ 50″ W	*L. l. iberica*	23	N (Northern)	10.8	1228	2000	75
Béjar	BEJ	40° 40′ 15″ N	5° 36′ 32″ W	*L. l. lepida*	22	C (Central)	10.5	565	2500	70
Serra da Estrela	SET	40° 19′ 24″ N	7° 36′ 44″ W	*L. l. lepida*	22	W (Western)	10.3	1563	2500	70
Peniche	SPE	39° 19′ 41″ N	9° 20′ 45″ W	*L. l. lepida*	16	W (Western)	15.1	616	2400	85
Toledo	TOL	39° 15′ 32″ N	3° 44′ 01″ W	*L. l. lepida*	22	S (Southern)	13.9	477	2700	60
Alentejo	ALE	38° 35′ 55″ N	7° 33′ 40″ W	*L. l. lepida*	13	S (Southern)	16.5	622	3000	70
Andalucia	AND	38° 16′ 51″ N	3° 37′ 02″ W	*L. l. lepida*	16	S (Southern)	15.9	646	2700	60
Castro Marim	CMA	37° 14′ 08″ N	7° 26′ 43″ W	*L. l. lepida*	25	A (Algarve)	17.3	488	2900	80
Almería	ALM	36° 49′ 54″ N	2° 31′ 32″ W	*L. l. nevadensis*	18	Nev (Nevadensis)	17.8	226	2900	65

databases for GPS coordinates of the 10 sampling sites in the Iberian Peninsula. Precipitation (annual and monthly values) and temperatures (annual and monthly values of maximum, mean and minimum temperatures) were obtained from the Digital Climatic Atlas of the Iberian Peninsula (Ninyerola *et al.* 2005, available at http://www.opengis.uab.es/WMS/iberia/en_index.htm). It presents a continuous distribution of values with a resolution of 200 m, obtained by spatial interpolation from data collected by meteorological stations for the period 1950–1999. Two other climatic variables, annual insolation (hours of sunshine) and annual relative humidity, were obtained from the Atlas of Portugal (APA 1975, scale 1:1 000 000) and the Atlas of Spain (IGN 1992, scale 1:4 500 000), where data is represented by isoline maps. In the Atlas of Portugal, for insolation (data from 1931–1960) and humidity (data from 1938–1970), the isoline of lower value flanking each sampling point was chosen. The same strategy was used to collect humidity data in the Atlas of Spain (data from 1956–1985), but for insolation, as isolines delimit intervals of 200 h instead of 100 h as in the Atlas of Portugal, the median value between flanking isolines was chosen for sampled locations in Spain. Mean annual values at each population main coordinates for temperature, precipitation, insolation and relative humidity are listed in Table 1.

AFLP genotyping

AFLP markers were obtained with a modified version of the original protocol in Vos *et al.* (1995). Digestion of 50–200 ng of genomic DNA was performed with 10 units each of restriction enzymes, *Eco*RI and *Mse*I, and 2× NEBuffer 2 (NEB) in 20 μL reaction for 3 h at 37 °C, followed by enzyme inactivation at 70 °C for 15 min. Ligation of adaptors was conducted for 16 μL of digested DNA in a 40-μL reaction with 1× T4 DNA Ligase Reaction Buffer, 75 pmol of each adaptor and 40 units of T4 DNA Ligase (NEB). The ligation reaction took place overnight at 16 °C and products were diluted 10-fold. For selective amplification, reactions of 10 μL were performed with 3.3 μL of template, 1× PCR buffer (Promega), 0.5 units of *Taq* polymerase (Promega), 1.5 mM $MgCl_2$, 0.3 mM of DNTPs, 1 pmol of *Mse*I selective primer and 0.2 pmol of *Eco*RI selective primer labelled with fluorescent dye (6-FAM or VIC). AFLP markers were generated for the following eight primer combinations (*Eco*RI-*Mse*I): ACA-CAG, AAC-CAC, ACT-CTG, AAC-CTC, ACT-CTT, AAG-CAC, ACA-ACA, and AAG-CTA. Fragments were separated by electrophoresis on an ABI PRISM 310 Genetic Analyser (Applied Biosystems) with Genescan-500 LIZ as internal size standard. Only polymorphic markers were considered and presence versus absence of peaks was called automatically by GeneMapper 3.7 (Applied Biosystems), setting a fluorescent signal detection threshold of 100 units to avoid background noise. All loci for all individuals were visually inspected and corrected for peak miscalls.

The error rate was estimated as the ratio of mismatches in 22 replicates (10.8% of total samples) to total number of replicated markers (Pompanon *et al.* 2005). After removal of markers with high mismatch rates between replicates, the error rate for 392 markers (35–64 per primer combination) was estimated at 5.6%, lower than the maximum value of 10% recommended by Bonin *et al.* (2007). The size range of AFLP markers was 62–303 bp, and 92% of markers had a fragment size above 100 bp.

Outlier detection

A modification of the FDIST2 software (Beaumont & Nichols 1996) for dominant markers, DFDIST (http://

www.rubic.rdg.ac.uk/~mab/stuff), was used to detect outlier loci, potentially influenced by selection, as loci with unusually high or low F_{ST} values when compared to neutral expectations. The first step of the analysis consists of the calculation of the neutral target F_{ST} for the empirical distribution, after excluding loci with critical frequency for the most common allele equal to or above 0.98. Second, a null distribution of F_{ST} close to the empirical distribution is obtained by coalescent simulation (50 000 realizations). Simulations were performed with a mean F_{ST} similar to the trimmed mean F_{ST}. The latter is intended to be an estimate uninfluenced by outlier loci (Bonin *et al.* 2006). Therefore, it was computed by removing the 30% highest and the 30% lowest F_{ST} values observed in the empirical distribution. Different values for the $4N_e\mu$ parameter were initially tested for the simulations (0.04, 0.06 and 0.4, corresponding to an N_e of 10 000, 15 000 and 100 000, respectively), but results remained robust as previously reported (Beaumont & Nichols 1996) and a value of 0.06 was chosen to perform all the simulations. Finally, empirical and simulated distributions were compared so that loci lying outside the upper and lower confidence levels were highlighted as outliers, and thus as candidate loci influenced by selection. A global analysis was performed with all 10 populations sampled in the Iberian Peninsula (Fig. 1). Both 95% and 99% confidence intervals (CI) were considered. However, the concerns about high false detection rates (type-I error) by DFDIST pointed out by several authors (Bonin *et al.* 2007; Caballero *et al.* 2008; Pérez-Figueroa *et al.* 2010) suggest the use of more restrictive significance levels to control for false positives. Therefore, a false discovery rate (FDR) of 5% was adopted for DFDIST analysis by applying the Benjamini & Hochberg (1995) method, as implemented by Chiurugwi *et al.* (2010).

Detection of outliers was also performed with the BayeScan software (http://www-leca.ujf-grenoble.fr/logiciels.htm), which implements the method of Foll & Gaggiotti (2008). This method directly estimates the probability that each locus is subject to selection using a Bayesian method. The method uses population-specific and locus-specific components of F_{ST} coefficients and assumes that allele frequencies follow a Dirichlet distribution. BayeScan takes all loci into account for the analysis and seems to be robust when dealing with complex demographic scenarios for neutral genetic differentiation (Foll & Gaggiotti 2008). As in DFDIST analysis, the 10 populations sampled in the Iberian Peninsula were used in a global analysis, performed with default parameters. Outliers detected with posterior probabilities above 0.99, but also above 0.95 or 0.90, were taken into consideration.

Association with environmental variables

The program SAM (Joost *et al.* 2008), available at http://www.econogene.eu/software/sam, was used to test for associations between the frequency of AFLP bands' presence/absence and data from environmental variables at sample locations. SAM computes multiple univariate logistic regression models. To ensure the robustness of the method, likelihood ratio (G) and Wald statistical tests are implemented to assess the significance of coefficients calculated by the logistic regression function. A model is considered significant only if the null hypothesis is rejected by both tests, after Bonferroni correction. For both tests, the null hypothesis is that the model with the examined variable does not explain the observed distribution better than a model with a constant only (Joost *et al.* 2007).

The 392 AFLP markers were tested against 54 climatic variables consisting of monthly precipitation and temperature (maximum, mean and minimum), and annual precipitation, temperature, insolation and relative humidity. All 10 populations from the Iberian Peninsula were used in the SAM analysis. The significance threshold was set to 4.72E-7, corresponding to a 99% CI after Bonferroni correction.

Neutral differentiation and genetic structure

Loci under divergent selection are expected to show larger differentiation among populations than neutral loci. To compare levels of differentiation between neutral AFLP markers and outliers potentially under selection, pairwise F_{ST} was calculated in AFLP-SURV 1.0 (Vekemans 2002). AFLP markers were grouped in two data sets: neutral loci and outlier loci. To minimize the presence of loci under weak selection or false negatives in the neutral loci data set, only AFLP markers that were never detected as outliers in any of the DFDIST (95% CI), BayeScan (with posterior probability above 0.90) or SAM (99% CI with Bonferroni correction) analyses were considered. The outlier loci data set corresponds to a conservative list intended for future validation studies, with outliers detected at more restrictive significance levels. The outliers considered were detected with a posterior probability higher than 0.99 by BayeScan or with a 5% FDR by DFDIST.

In order to evaluate whether the AFLP markers displayed the same genetic structure as that obtained with mtDNA (Table 1) or if population clustering would follow the geographical location of the climatic extremes, both neutral and outlier loci data sets were tested with analyses of molecular variance (AMOVAs) in Arlequin 3.5 (Excoffier *et al.* 2005). Analyses were performed with populations grouped in (i) a single group; (ii) in

three groups corresponding to each subspecies, which also correspond to the major climatic regions; and (iii) in six groups corresponding to each mtDNA clade.

Results

Outlier detection

A total of 392 polymorphic AFLP markers were successfully scored for 10 populations of ocellated lizards sampled along the climatic gradient in the Iberian Peninsula. The same data set was used in a global analysis for outlier detection, both with DFDIST and BayeScan.

DFDIST analysis produced 26 outliers at 99% CI and 62 at 95% CI. Among these outliers, some presented lower F_{ST} than expected under neutrality (one at 99% CI and 14 at 95% CI), corresponding to candidate loci influenced by balancing selection, which promotes locus polymorphism and the maintenance of similar allele frequencies across populations. The remaining outliers presented higher F_{ST} than expected under neutrality, being therefore potentially influenced by directional selection. When applying a 5% FDR, only 16 detections remained significant, all corresponding to outliers potentially influenced by directional selection, representing 4.1% of the investigated loci.

BayeScan analysis detected 30 outliers with posterior probability above 0.90, 20 above 0.95 and only 12 loci remained as outliers when considering posterior probabilities higher than 0.99, corresponding to 3.1% of the investigated loci. Outliers detected by BayeScan were all candidate loci potentially under divergent selection. The correspondence between loci detected as outliers by DFDIST (5% FDR) and by BayeScan (with posterior probability above 0.99) was limited. Among the 23 loci detected as outliers, only five (21.7%) were detected by both programs (Table 2). Even when less stringent significance criteria were considered (10% FDR for DFDIST, posterior probability > 0.90 for BayeScan), the proportion of loci simultaneously detected as outliers by both detection methods remained low (26%).

Table 2 List of 23 outliers selected for future validation studies. The list includes outliers detected by DFDIST with a FDR of 5% or with posterior probability above 0.99 by BayeScan. For each outlier, values of posterior probability from BayeScan (values above 0.99 in bold) and *P*-values from DFDIST (significant values with a 5% FDR in bold) are indicated. When outliers were detected by spatial analysis method (SAM), the climatic variables most strongly associated with locus' band frequency are indicated: maximum temperature (T_{max}), mean temperature (T_{mean}), minimum temperature (T_{min}), insolation (Ins), precipitation (Prec) and relative humidity (Hum)

Outlier	DFDIST *P*-value	BayeScan posterior probability	SAM
30	**0.000**	0.935	
52	0.031	**1.000**	T_{max}
75	0.024	**0.999**	Prec
140	**0.000**	0.604	
201	**0.001**	0.402	
209	**0.000**	0.684	Prec
220	0.106	**0.998**	
223	**0.000**	0.979	
228	**0.000**	0.908	
235	0.042	**0.998**	
245	**0.001**	**1.000**	T_{max}
291	**0.002**	0.509	
297	**0.000**	0.661	Prec
301	**0.000**	0.946	Prec, Ins, T_{med}
311	**0.001**	**0.999**	T_{max}, Ins
315	**0.000**	**1.000**	Ins
323	0.225	**0.992**	T_{max}
340	**0.000**	0.907	
347	**0.000**	0.710	
351	0.059	**0.991**	
353	0.037	**0.999**	Hum
386	**0.000**	**1.000**	
390	**0.002**	**0.999**	T_{max}, Ins

Association with environmental variables

AFLP marker frequency variation along the Iberian Peninsula was tested in SAM for associations with annual and monthly values of precipitation, temperature (maximum, mean and minimum), annual insolation and annual relative humidity. Significant associations were detected for 20 (5%) of 392 loci at 99% CI with Bonferroni correction (Table 3). SAM highlighted 15 loci associated with precipitation. Fifteen loci were also associated with maximum temperatures (T_{max}), but only five loci were associated with mean temperatures (T_{mean}) and only three loci with minimum temperatures (T_{min}). Eleven loci showed an association with insolation and only two loci were associated with relative humidity. Loci 317, 315 and 245 show the highest number of associations (with 37, 26 and 22 out of 54 variables, respectively), followed by loci 311 and 75 (14 and 10 associations, respectively). At the other end of the spectrum, loci 209, 213 and 353 were associated with only one variable.

The strongest associations were observed for locus 245 with T_{max} from June and locus 315 with annual insolation. Each locus found in association with T_{max} was at least associated with T_{max} from July and August, the hottest months of the year, except for loci 317 and 388. Locus 317 was associated with T_{max} for all months, June to September excepted. Associations with T_{mean} and T_{min} were also preferentially related to summer months, with the exception of locus 317.

Table 3 Association between AFLP loci frequency and climatic variables along the Iberian Peninsula as detected by spatial analysis method (SAM). Significant associations above 99% CI with Bonferroni correction are denoted by [+] or [−] (* 99.9% CI, ** 99.99% CI, *** 99. 999% CI and **** 99.9999% CI). When the probability of the AFLP band presence is higher for higher values of the variable, it is denoted by [+]. When the probability of the AFLP band presence is higher for lower values of the variable, it is denoted by [−]

	Locus	245	315	311	353	52	231	34	390	304	314	75	317	301	297	370	388	323	349	209	213
Precipitation	Annual		[+]									[+]			[−]						
	January																				
	February		[+]*									[+]*									
	March																				
	April		[+]									[+]		[−]							
	May		[+]										[+]								
	June		[+]*										[+]*	[−]							
	July	[−]	[+]***	[−]**			[+]*	[−]	[+]	[+]*	[+]*	[+]	[+]			[−]					
	August		[+]**	[−]**			[+]			[+]	[+]	[+]	[+]								
	September	[−]*	[+]***	[−]		[−]							[+]								
	October		[+]**									[+]*		[−]							
	November													[−]	[+]					[+]	
	December											[+]		[−]	[+]						
Temperature (maximum)	Annual	[+]***	[−]*	[+]		[+]							[−]								
	January												[−]								
	February												[−]								
	March	[+]*											[−]								
	April	[+]**	[−]			[+]							[−]								
	May	[+]***	[−]**	[+]*		[+]*							[−]								
	June	[+]****	[−]**	[+]**		[+]**	[−]	[+]	[−]	[−]	[+]										
	July	[+]***	[−]**	[+]***		[+]*	[−]*	[+]*	[−]*	[−]*	[+]*	[−]				[+]		[+]	[+]		
	August	[+]**	[−]**	[+]***		[+]*	[−]*	[+]*	[−]*	[−]*	[+]*	[−]				[+]	[−]	[+]	[+]		
	September	[+]***	[−]**	[+]**		[+]**	[−]	[+]	[−]								[−]				
	October	[+]*	[−]										[−]								
	November												[−]								
	December												[−]								
Insolation	Annual	[+]	[−]****	[+]***		[+]	[−]*	[+]*	[−]*			[−]	[−]	[+]			[−]				
Humidity	Annual				[+]***																[+]

Whenever a locus exhibited associations with temperature or insolation, but also with precipitation, the probability for the band presence was higher for higher temperatures and insolation hours, but lower for higher levels of precipitation, or vice-versa (Table 3). For example, for locus 315, the probability of the band's presence increased for higher values of precipitation and decreased for higher temperatures and insolation, while for locus 245 the opposite trends were observed. This observation is consistent with the inverse correlation between precipitation and temperature/insolation along the Iberian Peninsula, resulting in a gradient from north and northwest to south that decreases for precipitation but increases for temperature and insolation.

Comparing DFDIST and BayeScan with SAM analyses

Among the 16 outliers detected by DFDIST (with 5% of FDR), seven loci were also highlighted by SAM as significantly associated with some of the climatic variables tested (Table 2). For BayeScan, eight of the 12 outliers detected with posterior probabilities higher than 0.99 were highlighted by SAM. Only four loci highlighted by SAM (99% CI with Bonferroni correction) in association with climatic variables were simultaneously detected as outliers by DFDIST (5% FDR) and BayeScan (posterior probability > 0.99). These include loci 315, 245 and 311, which are among the loci associated with the highest number of climatic variables by SAM, as well as locus 390 (Table 2). All of the four outliers were associated most strongly with either T_{max} or annual insolation, or with both equally.

For follow-up research, a conservative list of 23 outliers was selected, where outliers detected by DFDIST (5% FDR) or BayeScan (posterior probability > 0.99) were retained (Table 2).

Neutral differentiation and genetic structure

Pairwise F_{ST} values between all 10 populations were calculated independently with neutral loci (318 loci)

and outlier loci (23 loci). The global F_{ST} for neutral loci was 0.05 and reached 0.24 for outlier loci. When considering the neutral loci, the highest values of F_{ST} were registered between ALM and the other nine populations (Table 4), with values ranging from 0.10 (ALM-SET) to 0.16 (ALM-GAL). For the outlier loci data set, an increase in F_{ST} was observed for almost all population pairs. The increase in F_{ST} for outlier loci was much more dramatic in population pairs involving ALM, GAL or GER, with values of F_{ST} that were, on average, five times higher than F_{ST} values obtained with the neutral loci. The highest F_{ST} value based on outlier loci was found between populations ALM and GAL (F_{ST} = 0.712), located at opposite extremes of the climatic gradient (Table 4).

Variance components from analyses of molecular variance (AMOVAs) performed with neutral and outlier loci data sets are summarized in Table 5. The percentage of variance explained among groups was considerably increased when outlier loci were used instead of neutral loci, reaching 46.85% when populations were grouped according to the subspecies, whose geographical distribution also coincides with the major climatic regions.

Discussion

The AFLP genome scan in *Lacerta lepida* revealed several candidate loci under divergent selection as expected from the high morphological and genetic differentiation observed for *L. l. iberica* and *L. l. nevadensis* subspecies, located at the opposite extremes of a climatic gradient. Moreover, a significant proportion of those loci also revealed statistically significant associations with climatic variables along the gradient.

Outlier detection

The proportion of outliers obtained with DFDIST (4.1%) is close to the proportion of outliers detected with BayeScan with posterior probability above 0.99 (3.1%) and also similar to the 5–10% reported in the generality of AFLP genome scans employing DFDIST (Nosil *et al.* 2009). However, most of these studies are not directly comparable, because chosen confidence levels may vary (α = 0.05, Savolainen *et al.* 2006; α = 0.01, Jump *et al.* 2006; α = 0.0005, Herrera & Bazaga 2008) as well as the study design (global analysis, Herrera & Bazaga 2008; pairwise comparisons, Jump *et al.* 2006; or

Table 4 Pairwise population F_{ST} (lower diagonal for 318 neutral loci and upper diagonal for 23 outliers)

	GAL	GER	BEJ	SET	SPE	TOL	ALE	AND	CMA	ALM
GAL	—	0.114	0.334	0.384	0.440	0.508	0.306	0.349	0.409	0.712
GER	0.040	—	0.141	0.139	0.203	0.283	0.085	0.097	0.190	0.608
BEJ	0.057	0.021	—	0.056	0.044	0.173	0.076	0.074	0.089	0.578
SET	0.049	0.010	0.016	—	0.035	0.107	0.015	0.003	0.068	0.579
SPE	0.051	0.021	0.048	0.022	—	0.181	0.067	0.065	0.092	0.595
TOL	0.064	0.034	0.029	0.014	0.033	—	0.080	0.069	0.083	0.626
ALE	0.077	0.032	0.038	0.017	0.027	0.036	—	0.000	0.047	0.550
AND	0.078	0.030	0.042	0.025	0.034	0.032	0.007	—	0.046	0.539
CMA	0.051	0.041	0.032	0.026	0.037	0.025	0.051	0.048	—	0.592
ALM	0.163	0.106	0.124	0.104	0.112	0.127	0.107	0.112	0.132	—

Table 5 Analysis of molecular variance (AMOVA) of the neutral loci and outlier loci data sets according to three grouping criteria: single group, by subspecies and by mtDNA clade (see Table 1 for each population's subspecies and mtDNA clade)

		318 neutral loci			23 outlier loci		
		Variance components (%)			Variance components (%)		
Grouping criteria	Number of groups	Within populations	Among populations within groups	Among groups	Within populations	Among populations within groups	Among groups
Single group	1	91.51	8.49		59.41	40.59	
By subspecies	3	88.01	4.05	7.95	45.99	7.16	46.85
By mtDNA clade	6	91.09	4.04	4.87	57.34	5.87	36.8

both, Bonin *et al.* 2006; see Nosil *et al.* 2009 for a review).

The same data set was used for both programs, but *P*-values from DFDIST and posterior probabilities obtained by BayeScan for the same loci are not directly comparable. Nevertheless, the correspondence between the most extreme outliers obtained with both programs is limited, probably reflecting the differences in their methodology. This behaviour is similar to the one evidenced by Beaumont & Balding (2004), who noted from the results of simulations that, while on average the power of the frequentist and Bayesian methods were similar, there was not necessarily a strong overlap in detected markers within simulated data sets (in which the markers under selection were known). The main difference between DFDIST and Bayescan is that the latter allows for variable within-population F_{ST}, whereas in the former it is assumed to be the same in all populations. This difference did not have a major impact in the simulations of Beaumont & Balding (2004), who simulated many-fold differences in F_{ST} among their populations, yet the overall power of both methods were quite similar. Both methods have a tendency to show false positives when there are correlations in allele frequencies among populations, because of shared recent ancestry or isolation by distance effects (Robertson 1975; Excoffier *et al.* 2009). Although Beaumont & Nichols (1996) detected no strong effect of isolation by distance or heterogeneous levels of gene flow between populations on outlier detection, the recent study by Excoffier *et al.*(2009) demonstrated some cases where neglecting population structure could lead to high rates of false positives. They implemented a new methodology for outlier detection where population structure can be taken into account for building the null distribution of F_{ST}, but the method is not yet available for dominant marker data.

In the present results, the most serious concerns about the effect of population structure in outlier detection are raised by the inclusion of the ALM population in global analyses. This population belongs to *L. l. nevadensis*, which started its divergence from the other ocellated lizards in the Iberian Peninsula long ago and has accumulated the highest levels of neutral divergence, as reported by cytochrome *b* mitochondrial gene analysis (Paulo 2001; Paulo *et al.* 2008). The exclusion of the ALM population from the global analysis did not have a major effect on results from BayeScan, but results from DFDIST were more severely affected (Table S1, Supporting information). On the other hand, ALM is also located at the SE extreme of the climatic gradient, and due to morphological differences observed in this subspecies, we expect that along with neutral differentiation, populations from the SE have also accumulated adaptive divergence by natural selection. Therefore, it is difficult to assess the extent to which the outliers detected as a result of the inclusion of the ALM population on global analysis are false positives.

Some studies have employed multiple pairwise comparisons among populations (see Nosil *et al.* 2009). Such comparisons are less susceptible to problems caused by unknown complexity in the true population structure and can strengthen evidence for candidate loci where independent comparisons can be made across the environmental transition (as in Wilding *et al.* 2001; for example). However, they are problematic when all pairs of samples are compared because comparisons are no longer independent and there is no way to correct for the large total number of comparisons. For the present data set, only five independent comparisons could be made. The proportion of outliers and the lists of AFLP markers detected when pairwise comparisons were made were very variable among methods and between global or pairwise analyses within the same detection method (Table S2, Supporting information).

Overall, DFDIST seems to be more sensitive than BayeScan to changes in the input and in stringency criteria. Caballero *et al.* (2008) also raised several concerns about the sensitivity of DFDIST and a recent simulation study by Pérez-Figueroa *et al.* (2010), comparing the efficiency of DFDIST, DETSELD and BayeScan to detect loci under directional selection with dominant markers, showed that BayeScan appears to perform more efficiently under a wide range of scenarios than the other methods. The results presented here show that using the same data set with different outlier detection methods can produce quite different results. If the purpose of a genome scan is to target candidate loci influenced by selection for further research, the most appropriate approach is probably the use of more than one method and to combine the outlier lists to avoid the risk of losing interesting candidates.

When applying SAM analysis to the AFLP data set, several loci detected as outliers by SAM in association with climatic variables were also highlighted as outliers by DFDIST or BayeScan. It is important to note that SAM is individual-centred, being independent of any notion of population and making no presumption as to the genotypic structure of the populations to which the sampled individuals belong (Joost *et al.* 2008). However, detection of significant associations by SAM depends on the variables tested. Therefore, the existence of associations with other environmental variables not tested here, such as soil colouration, or food availability, cannot be ruled out. To overcome this limitation, it seemed appropriate to set a list of

outliers for future validation that includes loci detected by DFDIST or BayeScan but not by SAM. Therefore, 23 outlier loci (6% of the investigated loci) stand out as strong candidate loci, potentially under divergent selection.

SAM is a useful method to search for associations with many environmental variables simultaneously, providing some insights into which selective forces may be in play. Since many environmental variables are closely correlated with others, information on the ecological requirements of the species is particularly important to assess which variables have the potential to be the major selective forces. Moreover, the variables tested must be ecologically relevant for the species. Otherwise, any variable that ranges along the same axis as the genetic data will turn out to be statistically associated with it, without being necessarily a selective force. For instance, if dispersal routes coincide with the environmental gradient, associations of allele frequencies with environmental variables could result from isolation by distance or founding effect rather than selection. For ocellated lizards, the colonization routes from glacial refugia were predictably located along river basins and valleys, which are mostly oriented southwest–northeast in the Iberian Peninsula (nearly perpendicular to the climatic gradient axis), rather than across the mountain systems that shape the landscape. Other methods currently available for dominant data to find associations with environmental variables, like generalized estimating equations, can take spatial autocorrelation into account by assuming that genetic similarity is normally higher for neighbouring individuals than distant ones (Poncet *et al.* 2010).

Neutral differentiation and genetic structure

Neutral loci differentiation reflects the existence of population genetic structure as predicted by mitochondrial divergence data in Paulo (2001) and Paulo *et al.* (2008), but this structure is much less strongly supported by AFLP markers. Mitochondrial clades expected within the nominal subspecies' distribution, *L. l. lepida*, are barely recognizable with neutral AFLP markers. The same situation was reported for the nuclear gene β-fibrinogen by Paulo *et al.* (2008), who argued that differences from the mtDNA pattern could result from the higher dispersal rate of males and/or the retention of ancestral polymorphism.

Genetic differentiation with outlier loci was much higher than with neutral loci, in particular for populations located at the opposite ends of the climatic gradient, ALM (*L. l. nevadensis*) and GAL (*L. l. iberica*), following the expected pattern of loci under selection.

Association with environmental variables

Results from the SAM analysis give us some clues about the possible selective forces acting upon outlier loci along the environmental gradient in the Iberian Peninsula. However, associations detected here denote correlations between environmental variables and AFLP band's frequency at each locus, not necessarily a causal relationship between them.

The strongest associations are registered for maximum monthly temperatures (T_{max}) and insolation. It is noteworthy that most T_{max} monthly variables associated with outlier locus frequencies correspond to the months when *L. lepida*'s activity is higher, from April to October (Busack & Visnaw 1989). As lizards are diurnal and ectothermic, temperature and insolation are known to play a major role in *L. lepida*'s seasonal and daily activities. Adult *L. lepida* are known to be active at the surface when ambient temperatures range from 15.6 to 42 °C (Busack & Visnaw 1989). During colder months, between November and February, they remain inactive while in the remaining months, exposure to sunshine is required to heat their body for daily activity and food search. Moreover, insects constitute *L. lepida*'s main food resource (Busack & Visnaw 1989), and their availability also depends on temperature, as they are ectothermic as well.

In the Iberian Peninsula, while the south registers hot summers, temperatures in the NW, where the *L. l. iberica* subspecies can be found, register much milder maximal temperatures, even in summer, and with much less annual insolation than in the south. Clearly, populations of *L. l. iberica* subspecies face more adverse climatic conditions, and its smaller and darker body may be a selective response to make thermoregulation faster and more efficient. Busack & Visnaw (1989) reported that sub-adults of *L. lepida* displaying the darker juvenile coloration seem to be more tolerant to cold and less tolerant to heat than adults, displaying surface activity at ambient temperatures from 11.1 to 30 °C. Clusella-Trullas *et al.* (2007) cites several studies in lizards showing that in most cases melanization increases solar absorption under cool conditions and allows melanistic animals to reach their thermal optimum more rapidly than lighter ones. Moreover, comparative studies on *L. lepida* subspecies (Mateo & Castanet 1994) show that when developing under the same controlled conditions, smaller body size can still be observed for *L. l. iberica*, indicating that besides possible environmental constraints, it is also genetically determined.

Several associations with precipitation were also found, although this variable probably does not affect lizards' activity as directly as temperature or insolation. Nevertheless, the asymmetry in precipitation distribu-

tion at the two extremes of the Iberian Peninsula can influence the reproductive cycle, the availability of food and strongly influences the vegetation cover (Hodar *et al.* 1996). Precipitation decreases from north to south, with a meridional asymmetry, leading to higher precipitation close to the Atlantic and lower close to the Mediterranean Sea (IGN 1992). In the NW, annual precipitation can reach values as high as 2000 mm. At the opposite extreme of the climatic gradient, in the SE, *L. l. nevadensis* subspecies survives with less than 300 mm of annual precipitation. The dry climate produces an arid landscape with scarce vegetation cover, where the proportion of bare ground can reach more than 40% (Hodar *et al.* 1996) and the distinct dorsal colour of *L. l. nevadensis* provides a good camouflage in a landscape where lizards are probably frequently exposed to predatory birds and mammals. Additionally, Mateo & Castanet (1994) proposed that irregular rainfall in the SE, and consequent irregular and limiting trophic resources, could explain the observed differences in *L. l. nevadensis* reproductive strategies, with annual asynchrony of clutches, an extended laying period and even the ability to lay more than one annual clutch, as opposed to the single annual clutch registered in the rest of *L. lepida*'s distribution.

Two loci were associated with relative humidity. Although the percentage of humidity in the air is normally higher in the NW than in the SE, locations in the south of the Iberian Peninsula, along the Atlantic coastline, can reach relative humidity values as high as in the NW. This is the case for SPE and CMA populations (Table 1), but morphological or life-history trait variation for these populations has not been extensively documented to date. The actual role of humidity in *L. lepida*'s local adaptation requires further investigation.

Future directions

This study highlighted a list of candidate loci suspected to be under the influence of selection in ocellated lizards, which are currently being isolated and sequenced. The characterization of outliers and the determination of their functional role, if any, is necessary to uncover their involvement in local adaptation of *L. lepida*, with particular interest for divergence of subspecies located at cline extremes. If outlier's sequence has no homology with any known gene, it may belong to an unknown regulatory region or simply a non-coding fragment that is in linkage with the actual target of selection. While in the first case, conducting further studies with the candidate gene under selection may be straightforward, in the two other cases, it implies the extension of the sequenced fragment to find the cause of marker polymorphism and map the outlier fragment in the genome. The latter situation requires the availability of more detailed genomic information on the species to proceed with the investigation, but the use of next generation sequencing offers the possibility to overcome this limitation in the near future (see review by Slate *et al.* 2010).

This study also highlights the usefulness of complementing genome scans for selection in natural populations with ecological data in order to identify selective pressures potentially acting upon candidate loci at a local scale. Public databases constitute a good source of environmental information at sample coordinates, but more detailed information, ecologically relevant for the species studied, needs to be collected in the field, such as vegetation cover, soil colouration, diversity and abundance of food sources or predation pressure. Combining ecological, phenotypic and genomic data will be very fruitful to test hypotheses regarding adaptation and speciation.

Acknowledgements

We thank A. Bonin and five anonymous referees for their helpful comments on the manuscript. This work was supported by a doctoral scholarship from FCT (SFRH/BD/21306/2005) to Nunes VL and the FCT project POCI/BIA/59288/2004- ''Process of speciation: accessing the genic view.'' coordinated by Paulo OS. All samples were collected under appropriate licenses in each country.

References

APA (1975) Atlas do Ambiente Digital. Agência Portuguesa do Ambiente (Portugal).

Beaumont MA, Balding DJ (2004) Identifying adaptive genetic divergence among populations from genome scans. *Molecular Ecology*, **13**, 969–980.

Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, **263**, 1619–1626.

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological)*, **57**, 289–300.

Bensch S, Akesson M (2005) Ten years of AFLP in ecology and evolution: why so few animals? *Molecular Ecology*, **14**, 2899–2914.

Black WC, Baer CF, Antolin MF, DuTeau NM (2001) Population genomics: genome-wide sampling of insect populations. *Annual Review of Entomology*, **46**, 441–469.

Bonin A, Taberlet P, Miaud C, Pompanon F (2006) Explorative genome scan to detect candidate loci for adaptation along a gradient of altitude in the common frog (*Rana temporaria*). *Molecular Biology and Evolution*, **23**, 773–783.

Bonin A, Ehrich D, Manel S (2007) Statistical analysis of amplified fragment length polymorphism data: a toolbox for

molecular ecologists and evolutionists. *Molecular Ecology*, **16**, 3737–3758.

Busack S, Visnaw JA (1989) Observations on the natural history of *Lacerta lepida* in Cádiz province, Spain. *Amphibia-Reptilia*, **10**, 201–213.

Caballero A, Quesada H, Rolan-Alvarez E (2008) Impact of amplified fragment length polymorphism size homoplasy on the estimation of population genetic diversity and the detection of selective loci. *Genetics*, **179**, 539–554.

Campbell D, Bernatchez L (2004) Generic scan using AFLP markers as a means to assess the role of directional selection in the divergence of sympatric whitefish ecotypes. *Molecular Biology and Evolution*, **21**, 945–956.

Castilla AM, Bauwens D (1989) Reproductive characteristics of the lacertid lizard *Lacerta lepida*. *Amphibia-Reptilia*, **10**, 445–452.

Chiurugwi T, Beaumont MA, Wilkinson MJ, Battey NH (2010) Adaptive Divergence and speciation among sexual and pseudoviviparous populations of *Festuca*. *Heredity*, In Press.

Clusella-Trullas S, van Wyk JH, Spotila JR (2007) Thermal melanism in ectotherms. *Journal of Thermal Biology*, **32**, 235–245.

Egan SP, Nosil P, Funk DJ (2008) Selection and genomic differentiation during ecological speciation: Isolating the contributions of host association via a comparative genome scan of *Neochlamisus bebbianae* leaf beetles. *Evolution*, **62**, 1162–1181.

Ellegren H (2008) Comparative genomics and the study of evolution by natural selection. *Molecular Ecology*, **17**, 4586–4596.

Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47–50.

Excoffier L, Hofer T, Foll M (2009) Detecting loci under selection in a hierarchically structured population. *Herediy*, **103**, 285–298.

Foll M, Gaggiotti O (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics*, **180**, 977–993.

Gavrilets S (2004) *Fitness Landscapes and the Origin of Species*. Princenton University Press, Princeton, NJ

Herrera CM, Bazaga P (2008) Population-genomic approach reveals adaptive floral divergence in discrete populations of a hawk moth-pollinated violet. *Molecular Ecology*, **17**, 5378–5390.

Hodar JA, Campos F, Rosales BA (1996) Trophic ecology of the Ocellated Lizard *Lacerta lepida* in an arid zone of southern Spain: relationships with availability and daily activity of prey. *Journal of Arid Environments*, **33**, 95–107.

Hoekstra HE, Hirschmann RJ, Bundey RA, Insel PA, Crossland JP (2006) A single amino acid mutation contributes to adaptive beach mouse color pattern. *Science*, **313**, 101–104.

IGN (1992) Atlas Nacional de España, Sección II: El medio terrestre—Grupo 9—Climatología, 1ª edicion. Instituto Geografico Nacional (Spain).

Jensen JD, Wong A, Aquadro CF (2007) Approaches for identifying targets of positive selection. *Trends in Genetics*, **23**, 568–577.

Joost S, Bonin A, Bruford MW *et al.* (2007) A spatial analysis method (SAM) to detect candidate loci for selection: towards a landscape genomics approach to adaptation. *Molecular Ecology*, **16**, 3955–3969.

Joost S, Kalbermatten M, Bonin A (2008) Spatial analysis method (SAM): a software tool combining molecular and environmental data to identify candidate loci for selection. *Molecular Ecology Resources*, **8**, 957–960.

Jump AS, Hunt JM, Martinez-Izquierdo JA, Penuelas J (2006) Natural selection and climate change: temperature-linked spatial and temporal trends in gene frequency in *Fagus sylvatica*. *Molecular Ecology*, **15**, 3469–3480.

Luikart G, England PR, Tallmon D, Jordan S, Taberlet P (2003) The power and promise of population genomics: From genotyping to genome typing. *Nature Reviews Genetics*, **4**, 981–994.

Mateo JA (1988) *Estudio Sistemático y Zoogeográfico de los Lagartos Ocelados, Lacerta lepida (Daudin, 1802) y Lacerta pater (Lataste 1880) (Sauria, Lacertidae)*. University of Sevilla, Sevilla.

Mateo JA, Castanet J (1994) Reproductive strategies in 3 Spanish populations of the ocellated lizard, *Lacerta Lepida* (Sauria, Lacertidae). *Acta Oecologica*, **15**, 215–229.

Mateo JA, Castroviejo J (1990) Variation morphologique et révision taxonomique de l'espèce *Lacerta lepida* Daudin, 1802 (Sauria, Lacertidae). *Bulletin du Muséum National D'Histoire Naturelle, Paris*, **12**, 691–706.

Mateo JA, López-Jurado LF (1994) Variaciones en el color de los lagartos ocelados; aproximación a la distribución de *Lacerta lepida nevadensis* Buchholz 1963. *Revista Española de Herpetología*, **8**, 29–35.

Mateo JA, López-Jurado LF, Guillaume CP (1996) Proteic and morphological variations in ocellated lizards (Lacertidae): a complex of species across the Strait of Gibraltar. *Comptes Rendus de l' Academie des Sciences Serie III-Sciences de la vie-Life Sciences*, **319**, 737–746.

Mateo JA, López-Jurado LF, Cano J, Guillaume CP (1999) Intraspecific variability in nucleolar organiser position in *Lacerta (Timon) lepida*. *Amphibia-Reptilia*, **20**, 73–77.

Mayer W, Bischoff W (1996) Beiträge zur taxonomischen Revision der Gattung *Lacerta* (Reptilia: Lacertidae) Teil 1: Zootoca, Omanosaura, Timon und Teira als eigenständige Gattungen. *Salamandra*, **32**, 163–170.

Minder AM, Widmer A (2008) A population genomic analysis of species boundaries: neutral processes, adaptive divergence and introgression between two hybridizing plant species. *Molecular Ecology*, **17**, 1552–1563.

Murray MC, Hare MP (2006) A genomic scan for divergent selection in a secondary contact zone between Atlantic and Gulf of Mexico oysters, *Crassostrea virginica*. *Molecular Ecology*, **15**, 4229–4242.

Ninyerola M, Pons X, Roure JM (2005) *Atlas Climático Digital de la Península Ibérica. Metodología y aplicaciones en bioclimatología y geobotánica*. ISBN 932860-8-7, Universidad Autónoma de Barcelona, Bellaterra.

Nosil P, Egan SP, Funk DJ (2008) Heterogeneous genomic differentiation between walking-stick ecotypes: ''Isolation by adaptation'' and multiple roles for divergent selection. *Evolution*, **62**, 316–336.

Nosil P, Funk DJ, Ortiz-Barrientos D (2009) Divergent selection and heterogeneous genomic divergence. *Molecular Ecology*, **18**, 375–402.

Paulo OS (2001). *The phylogeography of reptiles of the Iberian Peninsula*. University of London, London.

Paulo OS, Pinheiro J, Miraldo A *et al*. (2008) The role of vicariance vs. dispersal in shaping genetic patterns in ocellated lizard species in the western Mediterranean. *Molecular Ecology*, **17**, 1535–1551.

Pérez-Figueroa A, García-Pereira M, Saura M, Rolán-Alvarez E, Caballero A (2010) Comparing three different methods to detect selective loci using dominant markers. *Journal of Evolutionary Biology*, **23**, 2267–2276.

Pompanon F, Bonin A, Bellemain E, Taberlet P (2005) Genotyping errors: causes, consequences and solutions. *Nature Reviews Genetics*, **6**, 847–859.

Poncet B, Herrmann D, Gugerli F *et al.* (2010) Tracking genes of ecological relevance using a genome scan in two independent regional population samples of *Arabis alpina*. *Molecular Ecology*, **19**, 2896–2907.

Robertson A (1975) Gene frequency distribution as a test of selective neutrality. *Genetics*, **81**, 775–785.

Rosenblum EB, Hoekstra HE, Nachman MW (2004) Adaptive reptile color variation and the evolution of the *Mc1r* gene. *Evolution*, **58**, 1794–1808.

Savolainen V, Anstett MC, Lexer C *et al.* (2006) Sympatric speciation in palms on an oceanic island. *Nature*, **441**, 210–213.

Schluter D (2009) Evidence for ecological speciation and its alternative. *Science*, **323**, 737–741.

Slate J, Santure AW, Feulner PGD *et al.* (2010) Genome mapping in intensively studied wild vertebrate populations. *Trends in Genetics*, **26**, 275–284.

Stinchcombe JR, Hoekstra HE (2008) Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. *Heredity*, **100**, 158–170.

Vasemagi A, Primmer CR (2005) Challenges for identifying functionally important genetic variation: the promise of combining complementary research strategies. *Molecular Ecology*, **14**, 3623–3642.

Vekemans X (2002) *AFLP-SURV version 1.0. Distributed by the author*. Laboratoire de Génetique et Ecologie Végétale, Université Libre de Bruxelles, Belgium.

Via S (2009) Natural selection in action during speciation. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 9939–9946.

Vos P, Hogers R, Bleeker M *et al*. (1995) AFLP—a new technique for DNA-fingerprinting. *Nucleic Acids Research*, **23**, 4407–4414.

Wilding CS, Butlin RK, Grahame J (2001) Differential gene exchange between parapatric morphs of *Littorina saxatilis* detected using AFLP markers. *Journal of Evolutionary Biology*, **14**, 611–619.

Wood HM, Grahame JW, Humphray S, Rogers J, Butlin RK (2008) Sequence differentiation in regions identified by a genome scan for local adaptation. *Molecular Ecology*, **17**, 3123–3135.

Wu CI (2001) The genic view of the process of speciation. *Journal of Evolutionary Biology*, **14**, 851–865.

VLN, RKB and OSP are interested in using population genomics approaches to gain insights into the genetic basis of adaptive speciation. MAB is interested in understanding how genetic data can be used both to study adaptation and selection and also to elucidate the demographic history of populations.

Supporting information

Additional supporting information may be found in the online version of this article:

Table S1 Comparison of results from outlier detection by DFDIST and BayeScan with and without ALM population (*L. lepida nevadensis)*

Table S2 Results from outlier detection with DFDIST and BayeScan in five independent pairwise analyses along a climatic gradient in the Iberian Peninsula

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.