MOLECULAR ECOLOGY

Molecular Ecology (2010) doi: 10.1111/j.1365-294X.2010.04941.x

Stabilizing selection maintains exuberant colour polymorphism in the spider *Theridion californicum* (Araneae, Theridiidae)

PETER J. P. CROUCHER,* GEOFFREY S. OXFORD,† ATHENA LAM* and ROSEMARY G. GILLESPIE*

Department of Environmental Science, Policy, and Management, University of California, Berkeley, 130 Mulford Hall, Berkeley, CA 94720-3114, USA, †Department of Biology (Area 14), University of York, Wentworth Way, Heslington, York YO10 5DD, UK

Abstract

Genetically controlled colour polymorphisms provide a physical manifestation of the operation of selection and how this can vary according to the spatial or temporal arrangement of phenotypes, or their frequency in a population. Here, we examine the role of selection in shaping the exuberant colour polymorphism exhibited by the spider *Theridion californicum*. This species is part of a system in which several distantly related spiders in the same lineage, but living in very different geographical areas, exhibit remarkably convergent polymorphisms. These polymorphisms are characterized by allelic inheritance and the presence of a single common cryptic morph and, in the case of *T. californicum* and its congener the Hawaiian happy-face spider *Theridion grallator*, numerous rare patterned morphs. We compare population differentiation estimated from colour phenotypic data to differentiation at neutral amplified fragment length polymorphisms (AFLP) loci and demonstrate that the colour polymorphism appears to be maintained by balancing selection. We also examine the patterns of selection in the genome-wide sample of AFLP loci and compare approaches to detecting signatures of selection in this context. Our results have important implications regarding balancing selection, suggesting that selective agents can act in a similar manner across disparate taxa in globally disjunct locales resulting in parallel evolution of exuberant polymorphism.

Keywords: AFLP, balancing selection, convergent evolution, population differentiation, visible variation

Received 15 August 2010; revision received 4 October 2010; accepted 13 October 2010

Introduction

Visible genetic colour polymorphism is a widespread phenomenon among animals, and—because it provides measurable variation—has been the focus of much research to understand how selection can vary according to phenotype in a population and how this can lead to polymorphism (for reviews see: Gray & McKinnon 2006; Bond 2007). While some fluctuating or transient colour polymorphisms are maintained by differences in predator pressure because of habitat differences over space and time (Bond 2007), others may be a result of context dependent sexual selection (Gray *et al.* 2008). In each of these situations, the polymorphism appears to be maintained by a tightly coupled interaction between the environment and the effectiveness of the predator or mate at finding its target. Stable colour polymorphism in which morph frequencies are subject to stabilizing selection and which show little variation over space or time is relatively uncommon (Bond 2007). However, this form of colour polymorphism is particularly intriguing because it implies frequency-dependent selection intrinsic to the selective agent (most often a predator) in the community. Recent work has examined traits of predators that may lead to the establishment of

Correspondence: Peter J. P. Croucher, Fax: +1 510 643 5438; E-mail: croucher@berkeley.edu

© 2010 Blackwell Publishing Ltd

stable colour polymorphism within species (Bond 2007; Franks & Oxford 2009). Here, we examine a stable colour polymorphism to assess the consistency across lineages. We focus on a species of spider that is part of a system we are exploring in which colour polymorphism is displayed by multiple members in the same lineage, that occur in very different geographical areas, and are only distantly related (Arnedo *et al.* 2007, R.G. Gillespie, unpublished).

Colour and pattern polymorphisms have been reported in many spider species and vary tremendously in the diversity of forms and the extent of differences in expression between the sexes (reviewed by Oxford & Gillespie 1998). In most cases in which visible polymorphisms have been examined, the number of morphs is typically two or three. However, one species, the Hawaiian happy-face spider *Theridion grallator* Simon, provides a dramatic exception with more than 20 described abdominal colour patterns created from a palette of yellow, red, white and black pigments (Oxford & Gillespie 1996a,b, 2001). This dramatically diverse or 'exuberant' (Oxford 2009) colour polymorphism raises evolutionary questions about the adaptive significance and maintenance of many rare morphs in geographically widespread populations. Laboratory crosses within and between populations of *T. grallator* have confirmed that on most islands, colour is controlled by a single Mendelian locus with multiple alleles (Oxford & Gillespie 1996a), while one island exhibits a more complex pattern of inheritance with possibly two loci involved and some patterns showing sex limitation (Oxford & Gillespie 1996b). The change in the mechanism of inheritance has led to the suggestion that some features of the colour polymorphism have evolved independently on different islands (Oxford & Gillespie 1996c).

The most widely cited mechanism for the maintenance of balanced polymorphism is through apostatic selection (Bond 2007). Typically, predators over-concentrate on detecting and consuming common prey morphs, leading to negative frequency-dependent selection in favour of less common morphs (Clarke 1962; Allen 1988; Bond 2007; Franks & Oxford 2009). Recent modelling by Franks & Oxford (2009) has suggested that a simple model of apostatic selection is insufficient to generate a large number of colour morphs. However, the introduction of dietary wariness (Mappes *et al.* 2005) into the models, whereby a predator initially avoids novel prey, leads to the maintenance of exuberant polymorphisms. Furthermore, when one morph is cryptic it typically increases to high frequency, with the numerous, non-cryptic morphs being maintained at low frequency, as is found in the *T. grallator* system in the wild. The predators most frequently implicated in these systems are birds (Bond 2007): *T. grallator* itself is nocturnal and, like most web building spiders, has poor visual acuity. We note that the colours described are based on human perception. However, while it is well known that birds have a very different and broader range of spectral sensitivity than humans (Barber *et al.* 2006), human vision can be a valid proxy for avian colour discrimination (Seddon *et al.* 2010).

Recently, an exuberant polymorphism similar to that of *T. grallator* was discovered in another theridiid spider, *Theridion californicum* Banks (Oxford 2009). This species occurs along the narrow, moist coastal strip of western North America from British Columbia to southern California (Levi 1957) and exhibits at least 11 distinct colour morphs that show striking convergence with those of *T. grallator* in both appearance and frequency and seem to be inherited in a similar Mendelian fashion (Oxford 2009). In both species, and in two other well-studied polymorphic theridiid species, *Enoplognatha ovata* Clerck and *Enoplognatha latimana* Hippa & Oksala (Oxford 1983), there is a common yellow morph (typically representing 60–70% of the population) that appears to be recessive to all other morphs, with the dominance hierarchy of the colour-patterned morphs broadly reflecting the extent of pigmentation.

The similar patterning in these species suggests some element of parallel evolution or convergence, as they are phylogenetically separated by multiple species that do not show the phenomenon: the genus *Enoplognatha* occupies a basal position within the family relative to the more derived subfamily Theridiinae (Arnedo *et al.* 2004; Agnarsson 2004). Within the Theridiinae, *T. grallator* is nested within a tightly monophyletic clade of endemic Hawaiian species, while most California *Theridion* fall within the '*T. frondeum*' clade (Arnedo *et al.* 2007). Recent analysis of COI shows that *T. californicum*, as might be expected, is nested within the *T. frondeum* clade (R.G. Gillespie, unpublished). Therefore, the apparent convergent colour patterning of these disparate species implies a commonality of selective pressures operating across geographically widespread locations and suggests a theridiid ground plan for colour patterning that may be generalized across the family.

Here, we focus explicitly on *T. californicum* to determine whether the parallelism in colour polymorphism with *T. grallator* extends to the presence, and nature, of selection acting on the variation. We test the hypothesis that the qualitative variation in the colour phenotype is under stabilizing selection by comparing population differentiation assessed from the colour phenotype with population differentiation at neutral loci (McKay & Latta 2002). We extend this approach by employing a population genomic methodology to scan a

© 2010 Blackwell Publishing Ltd

genome-wide sample of amplified fragment length polymorphisms (AFLP) for loci that exhibit putative signals of natural selection. These approaches allow us to place selection at *Colour* in context with the extent of selection detectable across the genome, while filtering out locus-specific effects such as selection, recombination and mutation from genome-wide demographic effects such as genetic drift, bottlenecks, founder effects and inbreeding that are expected to affect the majority of loci similarly (Luikart *et al.* 2003; Stinchcombe & Hoekstra 2008). The aims of this study were therefore threefold: To assess whether: (i) the *Colour* locus shows evidence of being under balancing selection among *T. californicum* populations (as it is in *T. grallator*, Gillespie & Oxford 1998); (ii) anonymous AFLP loci show evidence of being under selection; and (iii) any AFLP loci are associated with the colour polymorphism and if so, whether these show evidence of balancing selection and are therefore likely to be in linkage disequilibrium with *Colour*.

Material and methods

Study populations

Six populations from the coastal region of central northern California were sampled (Fig. 1, Table 1), chosen to represent sites of variable geographical separation. Four of these populations were from the San Francisco Bay Area: Mill Valley (MIV), Marin Co. (the type location of *Theridion californicum*); Albany Hill (ALH), Albany, Alameda Co.; and Charles Lee Tilden Regional Park (Tilden Park), Alameda Co. The latter, located on the Coastal Range and Hayward Fault area immediately east of the cities of Berkeley and Kensington, was divided into two populations: East Tilden Park (ETP) and West Tilden Park (WTP)—the two regions are genetically distinct (Croucher *et al.* unpublished data). The fifth population was from the Guerneville (GUV) region, Sonoma Co., approximately 100 km north of San Francisco Bay, and the sixth was from Russian Gulch State Park (RUG), Mendocino Co., approximately 100 km north of GUV. Specimens (adult or subadult) were collected from beneath leaves or within rolled-leaves of under-storey plants between May and July of 2007 and 2008. Specimens were scored for colour polymorphism and preserved in 95% ethanol for DNA analysis.

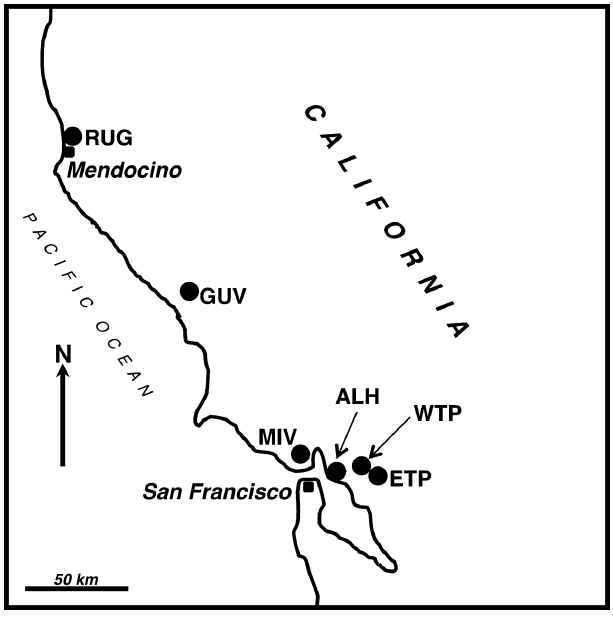

Fig. 1 Map showing the origins of the *Theridion californicum* populations. For codes to populations, see Table 1.

DNA extraction and genotyping

DNA was extracted from the legs of individual spiders using the DNeasy blood and tissue kit (Qiagen) according to the manufacturer's protocols. AFLP fragment profiles were generated using the restriction enzymes *Eco*RI and *Mse*I following Vos *et al.* (1995) and Bonin *et al.* (2005). Eight selective primer pairs were chosen from an initial screen of 64 primer combinations and selected to maximize reliability, scoring, variability and so that the AFLP profiles contained a moderate number of fragment peaks (30–100) to minimize within-peak homoplasy. The selective primer pairs and their final numbers of scored peaks are given in Table 2. The *Eco*RI selective primers were 5′-labelled with florescent dyes for analysis on an ABI 3730*ht* sequence analyser. To ensure reliability (see Bonin *et al.* 2004), DNA samples yielding low-quality profiles were excluded. Samples were randomly distributed across genotyping plates; negative controls were run at each step; and AFLP profiles were prepared and replicated for 20% of the samples. One sample was run across all plates as a positive control. The AFLP fingerprints were analysed in GeneMarker (SoftGenetics) and the matrix of raw peak intensities (relative florescence units, RFU) of all potential AFLP peaks was exported. Any peaks less than 50 bp or greater than 500 bp were excluded. The R-script (R Development Core Team 2008) AFLPscore (Whitlock *et al.* 2008) was employed as an objective way to score peak-height data. AFLPscore was used to normalize fingerprint profile intensities, prefilter the data for noise peaks and select the appropriate phenotype and relative genotype calling thresholds so as to minimize genotyping error rates on the basis of a Bayesian error assessment using the replicated samples. After scoring, the final matrices of AFLP phenotypes for each primer pair were concatenated and converted

© 2010 Blackwell Publishing Ltd

Table 1 Sample populations and colour morphs

Populations				Colour morphs*											
Locality†	Longitude	Latitude	*n*‡	Yellow	Red lines	White	Red ring B	Red ring A	Black spot	Red stripe A	Black blob	Red/ black ring	Red blob	White/ red lines	Red blob/ black spot	% Yellow
West Tilden Park (WTP)	−122.2654	37.9117	96	52	9	11	5	4	4	5	2	2	1		1	54.2
East Tilden Park (ETP)	−122.2539	37.7778	68	50	8	2	2	3		2	1					73.5
Albany Hill (ALH)	−122.3045	37.8977	30	10	10	2	1	1	4					2		33.3
Mill Valley (MIV)	−122.5638	37.9127	23	16	2	1		3	1							69.6
Russian Gulch (RUG)	−123.8033	39.3299	15	5		1	8						1			33.3
Guerneville (GUV)	−122.8692	38.4857	11	10					1							90.9
		Totals	243	143	29	17	16	11	10	7	3	2	2	2	1	
		%	–	58.8	11.9	7.0	6.6	4.5	4.1	2.9	1.2	0.8	0.8	0.8	0.4	

*Population codes are given in parentheses following the locality name.
†*n* = the total number of individuals from each sample population.
‡The numbers of each morph found in each population are given together with the overall percentage of the yellow morph in each population. The total counts of each colour morph and its overall percentage occurrence across all populations are given in the last two rows of the table. For detailed descriptions of the colour morphs, the reader is referred to Oxford (2009).

for further analyses using the R-script (R Development Core Team 2008) AFLPDAT (Ehrich 2006), yielding AFLP data for 243 individuals and 521 loci.

Data analysis

One way to infer the action and nature of the selection on a locus that is *presumed* to be under selection is to compare population differentiation at that locus with population differentiation at neutral loci (McKay & Latta 2002). The expectation is that if the trait is experiencing adaptive divergence to local conditions, then F_{ST} at the trait locus should exceed the F_{ST} estimate of population divergence by genetic drift provided by the neutral loci. Conversely, if selection is balancing—acting to maintain trait frequencies (i.e. polymorphism) across geographically spread populations or lineages—then the trait locus is expected to show less differentiation (lower F_{ST}) among populations than expected by drift (Lynch & Walsh 1998). One limitation of this approach is that the ability to detect the effects of selection is highly contingent both on the strength of selection and the degree of neutral differentiation among the study populations. If the differentiation (F_{ST}) attributed to selection is similar to the neutral differentiation, then selection will be undetectable.

Population genomics has added considerably to this approach by providing genome-wide sampling of loci. Any loci affecting fitness should behave differently and be detectable as 'outliers' when compared to genome-wide null expectations of population differentiation (Luikart *et al.* 2003; Stinchcombe & Hoekstra 2008). No a priori presumption regarding the action of selection at any one locus is made and, as a consequence, this genome-scan approach has been used to detect candidate genomic regions containing loci putatively involved in evolutionary change (Schlötterer 2003) and therefore provides a way of exploring the architecture of genomic adaptive evolution.

In all analyses presented here, the *Colour* locus was treated similarly to the AFLP data—as a dominant marker scored as either 0 (yellow—double recessive '*Yellow/Yellow*') or 1 (coloured—all colour morphs were assumed dominant to yellow and pooled). Loci were assessed as outliers using two approaches. First, the hypothesis that the *Colour* locus is subject to balancing selection was assessed by comparing θ (F_{ST}) (Weir & Cockerham 1984) at the *Colour* locus (θ_{Colour}) to a null estimate of θ calculated from the (mostly neutral) AFLP loci (θ_{AFLP}). The program TFPGA (Miller 1997) was used to estimate θ, assuming Hardy–Weinberg equilibrium and no inbreeding, with 99% confidence intervals (CI) for θ_{AFLP} obtained by bootstrapping 100 000 times across loci. This simple approach has been applied

© 2010 Blackwell Publishing Ltd

Table 2 AFLP selective primer pairs and numbers of polymorphic fragments

Primer pair	*Eco*RI primer*	*Mse*I primer*	Polymorphic fragments (#)
1	5′-FAM-GACTGCGTACCAATTC**ACA**-3′	5′-GATGAGTCCTGAGTAA**CAA**-3′	70
2	5′-VIC-GACTGCGTACCAATTC**AAG**-3′	5′-GATGAGTCCTGAGTAA**CAT**-3′	37
3	5′-NED-GACTGCGTACCAATTC**AAC**-3′	5′-GATGAGTCCTGAGTAA**CAC**-3′	90
4	5′-PET-GACTGCGTACCAATTC**ACG**-3′	5′-GATGAGTCCTGAGTAA**CAG**-3′	50
5	5′-FAM-GACTGCGTACCAATTC**ACT**-3′	5′-GATGAGTCCTGAGTAA**CAG**-3′	46
6	5′-VIC-GACTGCGTACCAATTC**AGC**-3′	5′-GATGAGTCCTGAGTAA**CCC**-3′	80
7	5′-NED-GACTGCGTACCAATTC**ACC**-3′	5′-GATGAGTCCTGAGTAA**CAT**-3′	76
8	5′-PET-GACTGCGTACCAATTC**AGG**-3′	5′-GATGAGTCCTGAGTAA**CCA**-3′	72

*Selective bases are highlighted in bold.

previously to visible polymorphisms in spiders (Gillespie & Oxford 1998) and in other taxa (e.g. Cook 1992; Abbott *et al.* 2008). Whether θ_{Colour} was an outlier (outside the θ_{AFLP} 99% CI) was assessed for all pairwise population comparisons and for the total data set.

Outliers among the AFLP loci (also including the *Colour* locus) were also identified using the population genomic approach of Beaumont & Nichols (1996) and Beaumont & Balding (2004) as implemented for dominant markers in the program DFDIST (http://www.rubic.rdg.ac.uk/~mab/stuff/). DFDIST uses the Bayesian method of Zhivotovsky (1999) to estimate allele frequencies from the recessive phenotypes in the sample and computes the empirical F_{ST} distribution from all loci. The trimmed mean of the empirical distribution is then computed by removing the 30% lowest and highest F_{ST} values observed in the data and used as an estimate of the average neutral F_{ST} uninfluenced by outliers (Bonin *et al.* 2006). This estimate is used to simulate a null distribution, from which upper and lower confidence limits (CL) are constructed, and from which outliers may be identified. DFDIST uses a hierarchical Bayesian model to estimate F_{ST} values conditional on heterozygosity in a subdivided population under Wright's (1951) symmetrical island model (Beaumont & Balding 2004). A two-step approach was employed to estimate the trimmed mean. First, the empirical trimmed mean F_{ST} was estimated and used to compute the null distribution. Second, all outlier loci from the first round of analysis were then removed and an adjusted trimmed mean calculated, as previously, and used as the basis for the final round of null F_{ST} estimation (Minder & Widmer 2008). For each analysis, all loci with an allele frequency greater than 0.98 were excluded (as recommended by the author of DFDIST) and the null F_{ST} distribution calculated with 100 000 realizations and $\theta = 0.06$ (default) ($\theta = 2\,N\mu$ for dominant AFLP markers and should realistically be much less than 1). The actual value of θ used has been shown to be very robust to misspecification (Beaumont & Nichols 1996; Minder & Widmer 2008).

Loci were considered as possible outliers when outside the 90% CL (the 5th and 95th quantiles of the null F_{ST} distribution). This yields a one-tailed test for 'negative outliers' with an F_{ST} lower than the 5th quantile. The 95% and 99% CL were also calculated. DFDIST analyses were carried out for all pairwise combinations and also for the total data set. Having a known candidate locus (*Colour*) that may be under balancing selection suggested an additional test whereby each population was divided into two populations—one containing yellow (double recessive) individuals and one containing the coloured individuals—yielding 12 populations. This creates an artificial situation in which the *Colour* locus should appear as an outlier with higher F_{ST} than expected (as if under divergent selection) along with any markers in linkage disequilibrium with the *Colour* locus.

Corrections for multiple comparisons were not applied to any of the tests presented here because these would be overly conservative and because we were merely searching for candidate loci. Rather, a weight-of-evidence approach was taken, whereby a locus was only considered a candidate outlier if it occurred in more than one pairwise comparison (loci identified as outliers in only one analysis were regarded as false). Furthermore, it became apparent from both the TFPGA and DFDIST analyses that the level of differentiation (F_{ST}) among the pairwise population comparisons varied tremendously and that the likelihood of identifying a locus under balancing selection was highly contingent on the average F_{ST}. Therefore, an additional test of significance was employed whereby the pairwise comparisons were ordered by decreasing F_{ST} and a cumulative combined *P*-value was computed from the *P*-values returned by DFDIST, incorporating each pairwise comparison in turn from highest to lowest F_{ST}. Inspection of these *P*-values

© 2010 Blackwell Publishing Ltd

Table 3 TFPGA and DFDIST assessments of selection at the *Colour* locus (in decreasing order of F_{ST_Null})

	TFPGA			*DFDIST*				
Comparison*	θ_{AFLP} (99% CI)†	θ_{COLOUR}‡	F_{ST_NULL}*	F_{ST_COLOUR}§	H_{COLOUR}¶	$P_{COMBINED}$**	No. loci F_{ST} < 5% CL††	No. loci F_{ST} > 95% CL‡‡
1. MIV-RUG	0.3173 (0.2906–0.3441)	**0.1309**	0.2363	**0.1153**	0.3897	0.3432	16	0
2. ETP-MIV	0.2510 (0.2275–0.2759)	**−0.0101**	0.2272	**−0.0134**	0.2559	0.1282	27	0
3. RUG-WTP	0.2303 (0.2106–0.2501)	**0.0429**	0.1983	**0.0264**	0.4144	0.1232	16	1
4. GUV-MIV	0.2908 (0.2651–0.3163)	**0.0370**	0.1956	**0.0346**	0.2278	0.1272	10	1
5. ETP-WTP	0.1725 (0.1573–0.1887)	**0.0434**	0.1824	**0.0420**	0.3397	0.1279	25	0
6. GUV-WTP	0.2127 (0.1938–0.2319)	**0.1017**	0.1579	**0.1429**	0.3728	0.1983	11	1
7. ALH-RUG	0.1808 (0.1617–0.2009)	**−0.0251**	0.0920	**−0.0324**	0.4958	0.1130	2	5
8. ALH-ETP	0.1129 (0.0981–0.1290)	0.1860	0.0646	0.1624	0.3435	0.1866	11	24
9. ALH-GUV	0.1589 (0.1396–0.1783)	0.2445	0.0513	0.2941	0.4291	0.2866	0	9
10. GUV-RUG	0.1781 (0.1484–0.2098)	0.2921	0.0398	0.2848	0.3769	0.4029	0	10
11. ALH-MIV	0.0505 (0.0426–0.0591)	0.1114	0.0227	0.1246	0.4257	0.5157	0	26
12. ETP-RUG	0.0920 (0.0771–0.1084)	0.2174	0.0211	0.1532	0.3063	0.6316	1	23
13. ETP-GUV	0.0798 (0.0670–0.0938)	**0.0135**	0.0174	0.0229	0.2276	0.7013	0	8
14. MIV-WTP	0.0092 (0.0058–0.0131)	0.0113	0.0059	0.0167	0.3773	0.7637	9	18
15. ALH-WTP	0.0235 (0.0192–0.0283)	0.0382	0.0057	0.0360	0.4259	0.8304	4	58
OVERALL	0.1413 (0.1311–0.1520)	**0.0701**	0.1186	**0.1020**	0.3614	0.8304	61	3

*Population comparisons are ordered by decreasing null F_{ST} (F_{ST_Null}) as determined by the DFDIST simulations.
†Estimates of θ (=F_{ST}) for 'neutral' AFLP loci, assuming HWE, as determined by TFPGA, with 99% CI determined by 100 000 bootstrap replicates.
‡Point estimates of θ (=F_{ST}) at the *Colour* locus as determined by TFPGA. Values in bold are lower then the lower 99% confidence limit for the neutral data θ_{AFLP}.
§Bayesian point estimates of F_{ST} at the *Colour* locus as determined by DFDIST. Values in bold are lower than the null estimate of F_{ST} (F_{ST_Null}) from the neutral AFLP data (does not imply statistical significance).
¶Bayesian point estimates of heterozygosity at the *Colour* locus as determined by DFDIST.
**Fisher's combined *P*-values based on the probability of obtaining an estimate of F_{ST_Colour} less than or equal to the observed value by chance, as determined by DFDIST. Combined *P*-values were computed for cumulatively increasing numbers of pairwise comparisons, from highest to lowest estimates of F_{ST_Null}.
††Number of AFLP loci with F_{ST} estimates lower than the 5% confidence limit (negative outliers) as determined by the final round of DFDIST simulation.
‡‡Number of AFLP loci with F_{ST} estimates greater than the 95% confidence limit (positive outliers) as determined by the final round of DFDIST simulation.

indicated that the lowest values were associated with $k = 8$ (the first eight most differentiated comparisons), in close agreement with the results of the initial TFPGA analysis of colour (see Results, Table 3). Combined *P*-values were computed according to Fisher (1932):

$$\chi^2 = -2\sum_{i=1}^{k} \log_e(p_i),$$

where p_i is the *P*-value for the *i*th hypothesis test and χ^2 has a chi-square distribution with two degrees of freedom. It was concluded that the seven comparisons with the smallest average F_{ST} would be unlikely to be reliable for identifying negative outliers. Therefore, the final set of highly likely outliers was determined to be those (i) observed as outliers in more than one pairwise DFDIST analysis (restricted to the eight pairwise comparisons with largest F_{ST}) and (ii) having a combined *P*-value (at $k = 8$) that was <0.05. Putative positive outliers (higher than expected F_{ST}) were subject to the same criteria as putative negative outliers. The final sets of positive and negative outliers were then subject to a confirmatory test of the hypothesis that they may be under selection by evaluating θ (F_{ST}) at each of these loci against (θ_{AFLP}) using the TFPGA approach as previously applied to the *Colour* locus.

Association among individual AFLP loci and the *Colour* locus was tested using a 2 × 2 contingency table Pearson chi-square test. This test, based on phenotypes [presence-absence of AFLP peaks or presence of colour vs. absence (yellow)], equates to a Cochran-Armitage trend test under a dominant genotype model and 1 d.f. Statistical significance was assessed by permuting 'affection status' (yellow/colour) over individuals 100 000 times (Devlin & Roeder 1999). To account for the population stratification inherent in our sampling, a

© 2010 Blackwell Publishing Ltd

simple genomic control (Devlin *et al.* 2001) was applied. The inflation factor λ was estimated as the median of the χ^2 statistics for each locus divided by 0.456 (Devlin & Roeder 1999) and was bounded as max(λ, 1) (Bacanu *et al.* 2000).

The final sets of negative and positive outliers and neutral loci were used to build neighbour-joining (NJ) trees. Nei's distance between populations was computed using AFLP-SURV version 1.0 (Vekemans 2002). One thousand bootstrapped distance matrices were created and NJ trees drawn with the NEIGHBOUR and CONSENS programs of the package PHYLIP version 3.69 (Felsenstein 2004).

Results

Signatures of natural selection were explored in a data set of 521 AFLP loci, genotyped in 243 individuals of *Theridion californicum* from six geographical populations. Each individual was scored for the abdominal colour polymorphism. The populations and the colour morphs scored in each population are given in Table 1. The overall frequencies of the various morphs were similar to those recorded by Oxford (2009), with approximately 60% of individuals exhibiting the yellow phenotype (Oxford recorded 63% yellow from Tilden Park). In the analyses performed here, the colour polymorphism was coded as a simple dominant locus [0 = yellow double recessive (*Yellow/Yellow*), 1 = any colour morph]. This was necessary as the underlying genotype of coloured individuals at the *Colour* locus was unknown. Consequently, *Colour* was treated identically to the AFLP markers and was assumed to represent a single genetic locus (see Discussion).

First, the hypothesis that the *Colour* locus exhibits a signature of balancing selection was assessed by comparing θ_{Colour} to the 99% confidence intervals of θ_{AFLP} as determined by bootstrapping across the (mostly) neutral AFLP loci using TFPGA. Table 3 (columns 2 and 3) shows that for 8 of 15 pairwise comparisons, θ_{Colour} was indeed far smaller than, and outside the 99% CI of the neutral estimate, θ_{AFLP} (Table 3, column 3, bold). This was also true for the combined estimates of θ_{Colour} and θ_{AFLP} over all six populations. This strongly supports the notion that this locus is under balancing selection. Further inspection revealed that the 'significant' comparisons were generally those where the neutral differentiation (θ_{AFLP}) was greatest. The only exceptions to this were ETP-GUV (#13), which was 'significant' with a θ_{AFLP} of only 0.0798, and GUV-RUG (#10) which was not significant but had a θ_{AFLP} of 0.1781. A loss of power to detect balancing selection with decreasing population differentiation (e.g. as θ_{AFLP} tends towards zero) is expected.

Colour was not identified as a negative outlier in any of the pairwise DFDIST genome-scan analyses (Table 3; columns 4 and 5) nor was it identified in the overall DFDIST analysis (Fig. 2a and Table 3, columns 4 and 5). However, it is noteworthy that $F_{\text{ST\_Colour}}$ was lower than the null estimate, $F_{\text{ST\_Null}}$, in eight of the nine comparisons (Table 3, column 5, bold) in which θ_{Colour} was below the 99% CI of θ_{AFLP}. This was also true for the overall DFDIST analysis. On the basis of the above results, and from inspection of the tables of combined *P*-values, it was decided that the seven least differentiated pairwise comparisons ALH-GUV (#9) through ALH-WTP (#15) were not likely to be reliable for detecting markers with unusually low values of F_{ST}. Consequently, only the first eight comparisons (MIV-RUG through ALH-ETP) were used to select potential outlying loci. The number of loci lying below the 5th quantile and above the 95th quantile in each DFDIST analysis is given in Table 3 (columns 8 and 9). The number of negative outliers identified drops dramatically for comparisons where $F_{\text{ST\_Null}} < 0.05$. In the first eight comparisons, a total of 80 potential negative outlier loci (ranging from 2 to 27) were identified. Of these, only 30 were found in more than one comparison. Five markers were found in three comparisons (114, 270, 274, 405, 412) and two markers were identified in four comparisons (324, 406). In the first eight comparisons, 30 poten-

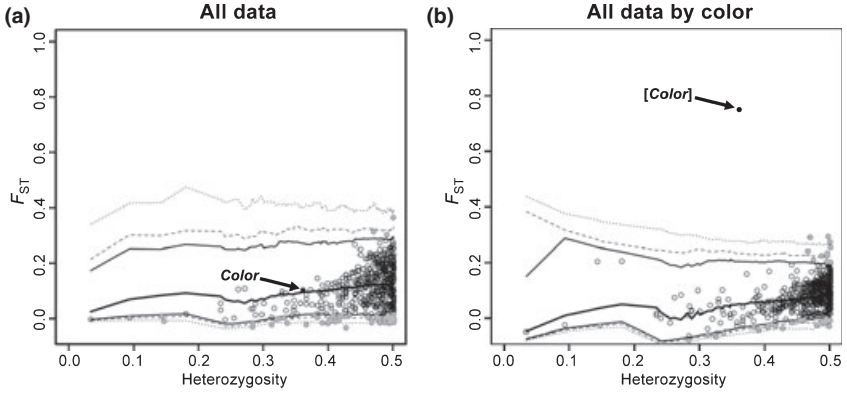

Fig. 2 Plots of F_{ST} against heterozygosity generated from DFDIST output. (a) Overall DFDIST analysis, (b) Overall DFDIST analysis with populations split into yellow and coloured subpopulations; forcing *Colour* to appear as a positive outlier. Thick solid line = median, thin solid line = 90% confidence limit, dashed line = 95% confidence limit, dotted line = 99% confidence limit. Outliers shaded gray.

© 2010 Blackwell Publishing Ltd

tial positive outlier loci (ranging from 0 to 24) were identified and of these only two markers (55, 285) were found in two comparisons.

The combined *P*-values (for negative outliers the probability of obtaining an F_{ST} the same or smaller than the observed; for positive outliers the probability of obtaining an F_{ST} the same or greater than observed) at $k = 8$ identified 65 potential negative outliers ($P_{COMBINED} < 0.05$). Of these, 26 were shared with the set of markers identified as negative outliers in more than one DFDIST comparison (Fig. 3a). These 26 markers were taken as the final set of putative markers under balancing selection. Combined *P*-values also identified nine potential positive outliers. Of these, two (55, 285) were shared with the set of markers identified as positive outliers in more than one DFDIST comparison (Fig. 3b) and were retained as the final set of putative positive outliers. The overall DFDIST analysis using all six populations identified 61 potential negative outliers. This set included 18 of 26 markers chosen as the final set of negative outliers. The overall analysis also identified three potential positive outliers. These did not include markers 55 and 285, although one of the markers (172) had been identified in the combined *P*-value analysis.

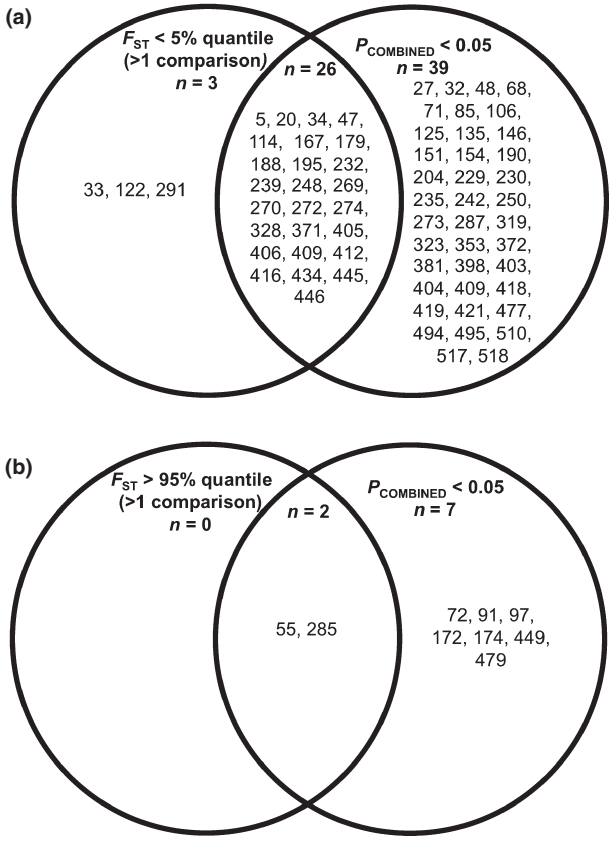

Fig. 3 Venn diagrams illustrating the intersection between outliers occurring more than once outside the 90% confidence limits in DFDIST and Fisher's combined *P*-values (at $k = 8$). (a) Negative outliers (balancing selection). (b) Positive outliers (divergent selection).

Association tests between *Colour* and the individual AFLP loci identified 12 loci (8, 62, 101, 120, 133, 138, 149, 182, 192, 243, 369, 450) that were potentially associated ($P < 0.05$) with the *Colour* phenotype. However, none of these markers was identified in any DFDIST analysis or from the combined *P*-value analysis and, consequently, are likely to represent false positives.

When populations were split into yellow and coloured subpopulations (yielding 12 populations) and subject to a DFDIST analysis, *Colour,* as expected, was detected as an extreme positive outlier (Fig. 2b). Eleven other markers were also detected as potential positive outliers and included the previously identified putative outliers 55 and 285 but included no other markers that had been detected in more than one pairwise DFDIST analysis or the combined *P*-value analysis. Twenty-nine potential negative outliers were also detected and included 12 of the 26 previously identified putative negative outliers (5, 20, 34, 179, 188, 248, 272, 274, 405, 406, 409, 416). With the exception of *Colour*, no previously identified putative negative outliers appeared as positive outliers in this analysis.

For the final set of 26 negative outliers and the set of two positive outliers, results of the combined *P*-value analysis are given in Table S1 (Supporting information). These markers were subject to a confirmatory test of the hypothesis that they are subject to selection by comparing θ_i for each marker ($i \ldots n$) against θ_{AFLP} for each pairwise population comparison and for the overall data using TFPGA (Table S2, Supporting information). In nearly all cases, θ_i of the negative outliers was far lower than θ_{AFLP} and frequently negative (zero).

NJ trees were drawn from the complete data, the set of neutral markers after removing outliers, the set of negative (balancing selection) outliers and the set of positive outliers (Fig. 4). The trees drawn from the total data set (Fig. 4a) and the neutral data set (Fig. 4b) were very similar with only slight changes in branch lengths and slightly higher bootstrap support for the GUV-RUG grouping with the neutral data. This result is not surprising given that the majority of outliers identified here were putatively under balancing selection, with only two putative positively selected markers. The removal of markers under balancing selection is likely to alter branch lengths but is less likely to alter the topology of the tree. The NJ tree drawn from the set of putative balancing selection markers was quite different, with branch lengths all close to zero, no bootstrap support and essentially a random topology, confirming the inability of these markers to distinguish the populations.

© 2010 Blackwell Publishing Ltd

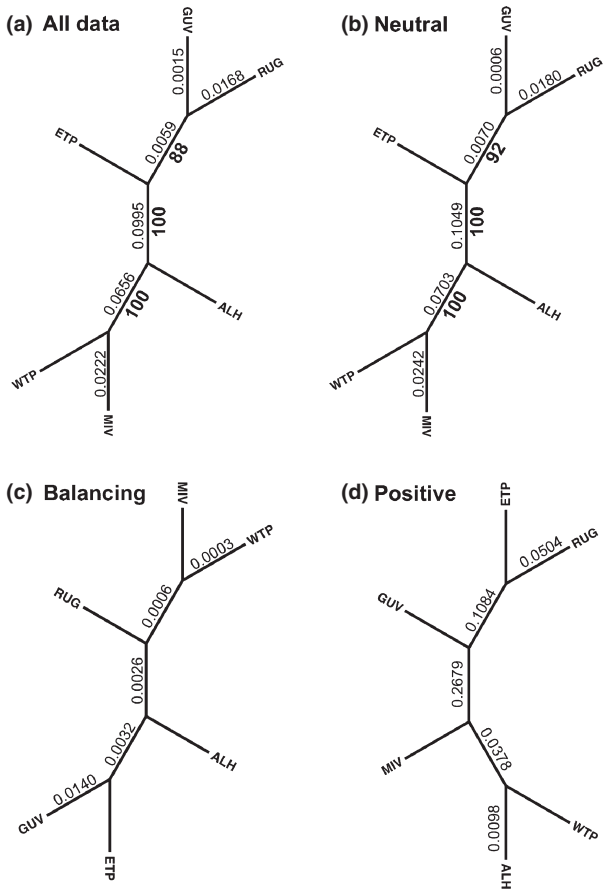

Fig. 4 Neighbour-joining trees based on Nei's distance for various marker sets with bootstrap values (>60%, 1000 replicates). (a) All data, (b) The neutral data set; (c) Balancing selection marker set (including colour); (d) Positive selection marker set. For (c) all bootstrap values were <60%. For (d) no bootstrap values were computed because there were only two loci.

The NJ tree for the two putative positive markers had a different topology again. These markers greatly exaggerated the distance between {MIV, WTP, ALH} and {GUV, ETP, RUG} and between {MIV, WTP, ALH, GUV} and {ETP, RUG} compared to the neutral and total data trees.

Discussion

The pairwise and global comparisons between F_{ST} at the *Colour* locus (θ_{Colour}) and a neutral estimate of F_{ST} from the AFLP data (θ_{AFLP}) strongly supported the conclusion that the colour polymorphism in *Theridion californicum* is indeed being maintained by some form of balancing selection. Ordering the results of the pairwise comparisons by decreasing null population differentiation also clearly indicated that there was a threshold in the value of the neutral F_{ST} estimate below which the test became inefficient at detecting this effect. This is logical. As neutral F_{ST} tends towards zero (nondifferentiation), it will become impossible to identify a locus experiencing balancing selection because the F_{ST} at the locus will be contained within the variation of the neutral estimate.

The DFDIST analyses failed to detect *Colour* as a negative outlier, which might appear to suggest that *Colour* is not under balancing selection. However, it is important to stress that the DFDIST analyses are merely a *scan* for outlier loci that may be under selection and not a test of the hypothesis that they are under selection. Furthermore, there may be a number of reasons for this apparently contradictory result. First, as Beaumont (2005) argued, there may be little power to detect outlier loci when making pairwise population comparisons because bi-allelic markers (such as AFLPs) can be subject to considerable genotyping and sampling error and skew in F_{ST} estimates. Although this difficulty may be somewhat alleviated when multiple population samples, collected across a wide geographical area, are subject to a combined analysis (Beaumont & Nichols 1996) in the current study, approximately half the pairwise comparisons may show too little differentiation to detect negative outliers reliably. Consequently, the combined analyses are likely to be confounded by the population structure (see Fig. 2a). Mäkinen *et al*. (2008) noted a similar dependence between the overall F_{ST} and the number of balancing selection signals in three-spined stickleback populations. Nonetheless, in all the most highly genetically differentiated population comparisons, F_{ST_Colour} was consistently lower than F_{ST_Null}, even though it never fell outside the confidence limits. Beaumont & Balding (2004) also argue that it may be difficult to detect balancing selection using the Bayesian F_{ST} (Beaumont & Nichols 1996) approach because the lower 95% confidence limit is typically close to zero; this is because the distribution of F_{ST} is typically positively skewed with most values being to the left of the mean. Again, combined analysis should improve the efficiency of testing for balancing selection because the overall F_{ST} estimates should tend towards normality. Despite these caveats, the approach clearly is able to detect negative outliers, albeit inefficiently. Beaumont & Balding (2004) suggest that the false discovery rate is also likely to be low (0.01%), so that when this approach does detect negative outliers, the signal is strong and they are unlikely to be false positives. The clear limitations of the two approaches employed here provide lessons for other researchers interested in exploring traits in natural populations that may be experiencing balancing selection. Particularly for species that have less clearly defined population structures, it is important to ensure that all pairwise population

© 2010 Blackwell Publishing Ltd

comparisons show moderate differentiation (i.e. as measured by neutral F_{ST}) to limit the confounding effects of comparisons with low levels of differentiation in the analysis. Of course, the limitations and development of these statistical approaches is an ongoing area of research. For excellent discussion (albeit focused on QTL loci) of some of the issues and difficulties regarding comparisons between F_{ST} and related statistics, we refer the reader to Leinonen *et al.* (2008) and Whitlock (2008).

Colour aside, the DFDIST approach did identify 26 (5%) AFLP loci that can confidently be considered true negative outliers and presumably are under some form of balancing selection. Confidence in this assertion comes from the fact that they were (i) identified as below the 95% CL in more than one pairwise population comparison, using only the eight most differentiated comparisons; (ii) also detected by combining *P*-values from the same eight comparisons; (iii) confirmed by comparing θ_{locus} to the 99% CI of mean θ_{AFLP} using TFPGA. Furthermore, the NJ tree constructed from these markers had extremely short branch lengths and an unsupported topology.

Two putative and well-supported positive outliers were also revealed by this study. There was no a priori reason to expect that we would detect evidence for divergent selection, as populations were not selected to have obvious differences in habitat or climate. Nonetheless, these markers clearly exaggerated the differences between the {MIV, ALH, WTP} and the {GUV, ETP, RUG} populations (Fig. 4d), a division that was also evident when NJ trees were drawn using the neutral markers alone and the complete data set. This division, together with little clear pattern of isolation-by-distance, reflects a complex biogeographical structure of this species in the San Francisco Bay Area, a pattern that is likely to have been shaped by a combination of postglacial expansion, complex topography and weather patterns and considerable shifts in land use regimes in the past 200 years (in particular a shift from pasture to forest) that has allowed some local populations to expand. We are currently exploring the biogeography of *T. californicum* in more detail.

The presence of a large number of loci (5%) under balancing selection in *T. californicum* agrees with theoretical expectations. Balancing selection may be a pervasive force shaping the structure of genomes, may be the predominant mode of selection in the wild (Mäkinen *et al.* 2008) and has been implicated as playing a major role in the evolution of both acquired and innate immunity genes in humans (e.g. Garrigan & Hedrick 2003; Ferrer-Admetlla *et al.* 2008). Balancing selection can be expected to keep many phenotypes, and their associated allele frequencies, close to their population mean frequency or average values (Kimura 1981; Mäkinen *et al.* 2008). Furthermore, Akey *et al.* (2002) identified 11% of 26 530 human single nucleotide polymorphisms, among three populations, to have $F_{ST} \approx 0.0$, compared to an average F_{ST} of 0.123 over all markers. In scan for footprints of selection in three-spined sticklebacks (*Gasterosteus aculeatus*), out of 105 markers Mäkinen *et al.* (2008) identified 14.7% to show evidence of balancing selection and only 2.8% to show evidence of directional selection. One possible explanation for the excess of markers showing signals of balancing selection, compared to signals of positive selection, in our study and others, would be if some loci have a higher than average mutation rate, leading to a higher than average polymorphism level (Mäkinen *et al.* 2008). Of course, given the anonymity of AFLP loci, this is hard to evaluate and clearly caution is required in interpreting these balancing selection signatures. However, given the inefficiency of the Bayesian F_{ST} approach to detect balancing selection, coupled with our conservative approach to selecting loci (so as to avoid false positives), it is likely that the number of true negative outliers loci has been underestimated. Support for this possibility comes from the observation that when the set of 26 negative outliers were treated as candidate loci and compared to the neutral θ_{AFLP} (F_{ST}) using TFPGA, they had estimates of $\theta \approx 0.0$ in nearly all comparisons. The fact that the *Colour* locus was also significant in this test, but not with DFDIST suggests that many other loci may have been overlooked.

It is highly unlikely that any of the loci examined are linked to the colour polymorphism in *T. californicum*. None of the 26 markers under balancing selection (nor the two under positive selection) were statistically associated with *Colour*. Indeed, it would have been surprising if any AFLP marker had been associated with *Colour*, given that 521 markers is an extremely low genome coverage (~1 marker/2 Mb: assuming a 1 Gb genome) and that loci under balancing selection are often considered difficult to detect by association (linkage disequilibrium) mapping because they tend to be ancient and recombination is likely to have eroded the linkage disequilibrium around the selected locus (Nielsen 2005; Worley *et al.* 2006). Furthermore, we have assumed that the *Colour* locus does in fact represent a single discrete genetic locus. The basis of this assumption is that analyses of colour-morph segregation in progeny from both laboratory crosses and wild-caught mother-offspring groups of *T. californicum* were consistent with segregation at a single Mendelian locus (Oxford 2009). Furthermore, colour segregation in progeny from all studied populations of *Enoplognatha ovata* and *E. latimina* and most populations of *T. grallator* (with the exception of Hawai'i) have also consistently

© 2010 Blackwell Publishing Ltd

indicated a single, discrete Mendelian locus (Oxford 1983, 2005; Oxford & Gillespie 1996a,b,c, 2001). It is of course possible that the genetic architecture of the colour polymorphism in *T. californicum* may not be that simple. If more than one locus is involved, then it further erodes the likelihood of detecting association between a colour locus or loci and any AFLP loci. This would not however greatly alter our ability to detect selection acting on the colour polymorphism *per se*, as selection should act directly on the colour *phenotype* and only indirectly on the underlying loci.

Our result here, in particular the detection of a strong signal suggesting the maintenance of the colour polymorphism by balancing selection, is not entirely unexpected: other evidence supporting this conclusion includes the polymorphic nature of all populations examined to date as well as the similarity to polymorphisms in other species. In particular, numerous lines of evidence now exist for selection acting to maintain the colour polymorphism of the candy-stripe spider *E. ovata*, including the sharing of the polymorphism with its sister species, *E. latimana*, the highly visible nature of the variation, the lack of monomorphic populations and the consistent rank order of morphs (Oxford 2005). The polymorphism in *T. grallator* is also present in all examined populations, both within and among the Hawaiian islands, despite the likelihood of a high degree of population isolation, strong founder effects and even changes in the mechanism of inheritance between islands (Oxford & Gillespie 1996a,b,c; Gillespie & Oxford 1998). Furthermore, a study showing a return to original frequencies after a natural perturbation is highly indicative of balancing selection (Gillespie & Oxford 1998).

Perhaps most significantly, the confirmation that the colour polymorphism in *T. californicum* is likely subject to balancing selection in the same way as occurs in other species suggests a common mode of selection, even though the spiders are unrelated, are found in very different parts of the world, occur in remarkably different habitats and interact with an almost entirely nonoverlapping set of predators and prey. Furthermore, the occurrence of common morphs across species may imply canalization of the developmental process by which the colour patterns themselves are laid down (Oxford 2009).

It has been argued (Blackledge *et al.* 2003) that the expansion of the two-dimensional orb-web into the tangled cobweb typical of Theridiid spiders was a defensive adaptation against predators and parasites that has contributed to the phenomenal success of this group of spiders. Despite their globally disjunct occurrence, *T. grallator* and *T. californicum*, as well as *E. ovata* and other colour polymorphic spiders (Oxford & Gillespie 1998) share a common feature in that they have largely abandoned their protective webs and now live exclusively under leaves. Therefore, even though the habitats, microhabitats and predator species in each situation will differ, the selective pressure imposed by this guild of predators (presumably gleaning birds) is likely very similar. This key commonality may provide the set of conditions required to drive the parallel evolution of exuberant polymorphisms in these spiders.

Acknowledgements

The authors thank J. DiDonato of the East Bay Regional Parks District for permission to collect in Tilden Regional Park, and the California State Parks. This research was supported by the Schlinger Fund (RGG) and a grant from the National Science Foundation (DEB 0919215). Python and R-scripts for computing cumulative combined *P*-values, testing association and plotting DFDIST output are available from PJPC on request.

References

Abbott JK, Bensch S, Gosden TP, Svensson EI (2008) Patterns of differentiation in a colour polymorphism and in neutral markers reveal rapid genetic changes in natural damselfly populations. *Molecular Ecology*, **17**, 1597–1604.

Agnarsson A (2004) Morphological phylogeny of cobweb spiders and their relatives (Araneae, Araneoidea, Theridiidae). *Zoological Journal of the Linnean Society*, **141**, 447–462.

Akey JM, Zhang G, Zhang K, Jin L, Shriver MD (2002) Interrogating a high-density SNP map for signatures of natural selection. *Genome Research*, **12**, 1805–1814.

Allen JA (1988) Frequency-dependent selection by predators. *Philosophical Transactions of the Royal Society of London B*, **319**, 485–503.

Arnedo MA, Coddington J, Agnarsson I, Gillespie RG (2004) From a comb to a cree: phylogenetic relationships of the comb-footed spiders (Araneae, Theridiidae) inferred from nuclear and mitochondrial genes. *Molecular Phylogenetics and Evolution*, **31**, 225–245.

Arnedo MA, Agnarsson I, Gillespie RG (2007) Molecular insights into the phylogenetic structure of the spider genus *Theridion* (Araneae, Theridiidae) and the origin of the Hawaiian *Theridion*-like fauna. *Zoologica Scripta*, **36**, 337–352.

Bacanu S-A, Devlin B, Roeder K (2000) The power of genomic control. *American Journal of Human Genetics*, **66**, 1933–1944.

Barber CL, Prescott NB, Jarvis JR, LeSueur C, Perry GC, Wathes CM (2006) Comparative study of the photopic spectral sensitivity of domestic ducks (*Anas platyrhynchos domesticus*), turkeys (*Meleagris gallopavo gallopavo*) and humans. *British Poultry Science*, **47**, 365–374.

Beaumont MA (2005) Adaptation and speciation: what can F_{ST} tell us? *Trends in Ecology and Evolution*, **20**, 435–440.

Beaumont MA, Balding DJ (2004) Identifying adaptive genetic divergence among populations from genome scans. *Molecular Ecology*, **13**, 1619–1626.

© 2010 Blackwell Publishing Ltd

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

Blackledge TA, Coddington JA, Gillespie RG (2003) Are three-dimensional spider webs defensive adaptations? *Ecology Letters*, **6**, 13–18.

Bond AB (2007) The evolution of colour polymorphism: crypticity, searching images, and apostatic selection. *Annual Review of Ecology, Evolution, and Systematics*, **38**, 489–514.

Bonin A, Bellemain E, Bronken Eidesen P, Pompanon F, Brochmann C, Taberlet P (2004) How to track and assess genotyping errors in population genetics studies. *Molecular Ecology*, **13**, 3261–3273.

Bonin A, Pompanon F, Taberlet P (2005) Use of amplified fragment length polymorphism (AFLP) markers in surveys of vertebrate diversity. In: *Molecular Evolution: Producing the Biochemical Data, Part B. Methods in Enzymology*, Vol. 395 (eds Zimmer EA, Roalson E), pp. 145–161. Academic Press, New York.

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*, **21**, 773–783.

Clarke BC (1962) Balanced polymorphism and the diversity of sympatric species. In: *Taxonomy and Geography* (ed. Nichols D), pp. 47–70. Systematics Association, Oxford, UK.

Cook LM (1992) The neutral assumption and maintenance of colour morph frequency in mangrove snails. *Heredity*, **69**, 184–189.

Devlin B, Roeder K (1999) Genomic control for association studies. *Biometrics*, **55**, 997–1004.

Devlin B, Roeder K, Wasserman L (2001) Genomic control, a new approach to genetic-based association studies. *Theoretical Population Biology*, **60**, 155–166.

Ehrich D (2006) AFLPdat: a collection of R functions for convenient handling of AFLP data. *Molecular Ecology Notes*, **6**, 603–604.

Felsenstein J (2004) PHYLIP (phylogeny inference package). Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

Ferrer-Admetlla A, Bosch E, Sikora M *et al.* (2008) Balancing selection is the main force shaping the evolution of innate immunity genes. *The Journal of Immunology*, **181**, 1315–1322.

Fisher RA (1932) *Statistical Methods for Research Workers*, 4th edn. Oliver and Boyd, Edinburgh.

Franks DW, Oxford GS (2009) The evolution of exuberant visible polymorphisms. *Evolution*, **63**, 2697–2706.

Garrigan D, Hedrick PW (2003) Perspective: detecting adaptive molecular polymorphism: lessons from the MHC. *Evolution*, **57**, 1707–1722.

Gillespie RG, Oxford GS (1998) Selection on the colour polymorphism in Hawaiian happy-face spiders: evidence from genetic structure and temporal fluctuations. *Evolution*, **52**, 775–783.

Gray SM, McKinnon JS (2006) Linking colour polymorphism and speciation. *Trends in Ecology and Evolution*, **22**, 71–79.

Gray SM, Dill LM, Tantu FY, Loew ER, Herder F, McKinnon JS (2008) Environment-contingent sexual selection in a colour polymorphic fish. *Proceedings of the Royal Society B*, **275**, 1785–1791.

Kimura M (1981) Possibility of extensive neutral evolution under stabilizing selection with special reference to nonrandom usage of synonymous codons. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 5791–5796.

Leinonen T, O'Hara RB, Cano JM, Merilä J (2008) Comparative studies of quantitative trait and neutral marker divergence: a meta-analysis. *Journal of Evolutionary Biology*, **21**, 1–17.

Levi HW (1957) The spider genera *Enoplognatha, Theridion* and *Paidisca* in America north of Mexico. *Bulletin of the American Museum of Natural History*, **112**, 1–123.

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.

Lynch M, Walsh B (1998) *Genetics and Analysis of Quantitative Traits*. Sinauer & Associates, Sunderland, MA.

Mäkinen HS, Cano JM, Merilä J (2008) Identifying footprints of directional and balancing selection in marine and freshwater three-spined stickleback (*Gasterosteus aculeatus*) populations. *Molecular Ecology*, **17**, 3565–3582.

Mappes J, Marples N, Endler JA (2005) The complex business of survival by aposematism. *Trends in Ecology and Evolution*, **20**, 598–603.

McKay JK, Latta RG (2002) Adaptive population divergence: markers, QTL and traits. *Trends in Ecology and Evolution*, **17**, 285–291.

Miller MP (1997) Tools for Population Genetic Analysis (TFPGA), 1.3: a Windows program for the analysis of allozyme and molecular population genetic data. Distributed by the author: http://www.marksgeneticsoftware.net/tfpga.htm.

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.

Nielsen R (2005) Molecular signatures of natural selection. *Annual Review of Genetics*, **39**, 197–218.

Oxford GS (1983) Genetics of colour and its regulation during development in the spider *Enoplognatha ovata* (Clerck) (Araneae: Theridiidae). *Heredity*, **51**, 621–634.

Oxford GS (2005) Genetic drift within a protected polymorphism: enigmatic variation in colour-morph frequencies in the candy-stripe spider *Enoplognatha ovata*. *Evolution*, **59**, 2170–2184.

Oxford GS (2009) An exuberant, undescribed colour polymorphism in *Theridion californicum* (Araneae, Theridiidae): implications for a theridiid pattern ground plan and the convergent evolution of visible morphs. *Biological Journal of the Linnaean Society*, **96**, 23–34.

Oxford GS, Gillespie RG (1996a) Genetics of a colour polymorphism in the Hawaiian happy-face spider, *Theridion grallator* (Araneae: Theridiidae) from Greater Maui. *Heredity*, **76**, 238–248.

Oxford GS, Gillespie RG (1996b) Quantum shifts in the genetic control of a colour polymorphism in the Hawaiian happy-face spider, *Theridion grallator* (Araneae: Theridiidae). *Heredity*, **76**, 249–256.

Oxford GS, Gillespie RG (1996c) The effects of genetic background on the island-specific control of a colour

© 2010 Blackwell Publishing Ltd

polymorphism in *Theridion grallator* (Araneae: Theridiidae), the Hawaiian happy-face spider. *Heredity*, **76**, 257–266.
Oxford GS, Gillespie RG (1998) Evolution and ecology of spider colouration. *Annual Review of Entomology*, **43**, 619–643.
Oxford GS, Gillespie RG (2001) Portraits of evolution: studies of colouration in Hawaiian spiders. *BioScience*, **51**, 521–528.
R Development Core Team (2008) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-90051-07-08. http://www.R-project.org.
Schlötterer C (2003) Hitchhiking mapping – functional genomics from the population genetics perspective. *Trends in Genetics*, **19**, 32–38.
Seddon N, Tobias JA, Eaton M, Ödeen A (2010) Human vision can provide a valid proxy for avian perception of sexual dichromatism. *The Auk*, **127**, 283–292.
Stinchcombe JR, Hoekstra HE (2008) Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. *Heredity*, **100**, 158–170.
Vekemans X (2002) AFLP-SURV version 1.0. Distributed by the author. Laboratoire de Génetique et Ecologie Végétale, Université Libre de Bruxelles, Belgium.
Vos P, Hogers R, Bleeker M *et al*. (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research*, **23**, 4407–4414.
Weir BS, Cockerham CC (1984) Estimating *f*-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.
Whitlock MC (2008) Evolutionary inference from QST. *Molecular Ecology*, **17**, 1885–1896.
Whitlock R, Hipperson H, Mannerelli H, Butlin RK, Burke T (2008) An objective, rapid and reproducible method for scoring AFLP peak-height data that minimizes genotyping error. *Molecular Ecology Resources*, **8**, 725–735.
Worley K, Carey J, Veitch A, Coltman DW (2006) Detecting the signature of selection on immune genes in highly structured populations of wild sheep (Ovis dalli). *Molecular Ecology*, **15**, 623–637.
Wright S (1951) The genetical structure of populations. *Annals of Eugenics*, **15**, 323–354.
Zhivotovsky LA (1999) Estimating population structure in diploids with multilocus dominant DNA markers. *Molecular Ecology*, **8**, 907–913.

PJPC is an evolutionary biologist and statistical population geneticist whose interests genetic epidemiology and the genomics of selection and adaptation. He is an Associate Specialist at the University of California, Berkeley. GSO is an evolutionary biologist and geneticist whose research interests include speciation, hybridization, selection and the evolution of colour polymorphisms, especially in spiders. He is an Emeritus Reader at the University of York, UK. AL is a molecular biologist and is currently a graduate student at the University of California, San Diego. RGG is Professor and Director of the Essig Museum of Entomology, Berkeley Natural History Museum. Her research interests include evolutionary and ecological processes that shape communities, biodiversity science, and systematics, especially of spiders in the Pacific.

Supporting information

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

Table S1 Fisher's combined *P*-values, determined cumulatively from high to low F_{ST_Null}, for the final sets of negative and positive outliers

Table S2 Confirmatory TFPGA analyses of θ versus θ_{AFLP} for the final sets of negative and positive outliers

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.

© 2010 Blackwell Publishing Ltd