

ORIGINAL ARTICLE

Candidate genes for colour and vision exhibit signals of selection across the pied flycatcher (*Ficedula hypoleuca*) breeding range

PK Lehtonen¹, T Laaksonen¹, AV Artemyev², E Belskii³, PR Berg^{4,17}, C Both⁵, L Buggiotti^{1,18}, S Bureš⁶, MD Burgess⁷, AV Bushuev⁸, I Krams^{9,1}, J Moreno¹⁰, M Mägi¹¹, A Nord¹², J Potti¹³, P-A Ravussin¹⁴, PM Sirkiä¹, G-P Sætre¹⁵, W Winkel¹⁶ and CR Primmer¹

The role of natural selection in shaping adaptive trait differentiation in natural populations has long been recognized. Determining its molecular basis, however, remains a challenge. Here, we search for signals of selection in candidate genes for colour and its perception in a passerine bird. Pied flycatcher plumage varies geographically in both its structural and pigment-based properties. Both characteristics appear to be shaped by selection. A single-locus outlier test revealed 2 of 14 loci to show significantly elevated signals of divergence. The first of these, the follistatin gene, is expressed in the developing feather bud and is found in pathways with genes that determine the structure of feathers and may thus be important in generating variation in structural colouration. The second is a gene potentially underlying the ability to detect this variation: *SWS1 opsin*. These two loci were most differentiated in two Spanish pied flycatcher populations, which are also among the populations that have the highest UV reflectance. The follistatin and *SWS1 opsin* genes thus provide strong candidates for future investigations on the molecular basis of adaptively significant traits and their co-evolution.

Heredity (2012) **108**, 431–440; doi:10.1038/hdy.2011.93; published online 26 October 2011

Keywords: SNP; melanin; ultraviolet reflectance; outlier test; passerine bird

INTRODUCTION

Animal colouration has been suggested to respond rapidly to selection pressure from environmental variation (for example, Endler *et al.*, 2005). Identifying the molecular mechanisms that underlie colour variation thus serves as a valuable means for broadening our understanding on how selection shapes genetic variation leading to evolutionary change. Recent studies of pigmentation genes in the wild have revealed that pigment gene function is much conserved across vertebrate taxa (for example, Boswell and Takeuchi, 2005) and often influences adaptive colouration in a predictable manner (reviewed by Hubbard *et al.*, 2010). The prominent colour variation seen in the animal kingdom is primarily produced by variation in the amount and density of pigment granules in the integument (pigment-based colouration) or its structural properties (structural colouration; Andersson and Prager, 2006).

The pied flycatcher (*Ficedula hypoleuca*) is a good model for studying the evolutionary significance of both these types of colour variation as the species shows spatial variation in both pigmentary and structurally based plumage characteristics. Unlike for many other

avian species (see Mundy, 2005), a number of hypotheses for the adaptive significance for both pigmentary and structural variation in pied flycatchers have been put forward. The melanin-based component of male breeding plumage has been suggested to primarily be shaped by selection pressure generated by the presence/absence of the congeneric dominant black-and-white collared flycatcher (*Ficedula albicollis*; (Sætre and Saether, 2010)). In areas where the two species live in sympatry, the vast majority of pied flycatcher males are brown and the characteristic is hypothesized to be used in species recognition. As the distance from the sympatric regions increases, so does the frequency of more darkly coloured pied flycatcher males (see, for example, Lehtonen *et al.*, 2009a). In allopatric areas, plumage colour of pied flycatcher males has been suggested to be sexually selected, but the evidence for this is mixed (Lehtonen *et al.*, 2009b; Sirkiä and Laaksonen, 2009), and it is possible that other selective forces also affect the distribution of colour phenotypes.

In addition to variation in melanin-based colouration, pied flycatchers also vary in structurally based plumage characteristics such

¹Department of Biology, University of Turku, Turku, Finland; ²Institute of Biology, Karelian Research Centre, Russian Academy of Science, Petrozavodsk, Russia; ³Institute of Plant and Animal Ecology, Russian Academy of Science, Ekaterinburg, Russia; ⁴Norwegian University of Life Sciences, CIGENE, Centre for Integrative Genetics, Ås, Norway; ⁵Centre for Ecological and Evolutionary Studies, University of Groningen, Haren, The Netherlands; ⁶Laboratory of Ornithology, Palacký University, Olomouc, Czech Republic; ⁷Centre for Research in Animal Behaviour, School of Psychology, University of Exeter, Exeter, UK; ⁸Department of Vertebrate Zoology, Faculty of Biology, Moscow State University, Moscow, Russia; ⁹Institute of Systematic Biology, University of Daugavpils, Daugavpils, Latvia; ¹⁰Departamento de Ecología Evolutiva, Museo Nacional de Ciencias Naturales-CSIC, Madrid, Spain; ¹¹Department of Zoology, Institute of Ecology and Earth Sciences, University of Tartu, Tartu, Estonia; ¹²Department of Animal Ecology, Lund University, Lund, Sweden; ¹³Department of Evolutionary Ecology, Estación Biológica de Doñana-CSIC, Sevilla, Spain; ¹⁴Rue du Theu, Baulmes, Switzerland; ¹⁵Department of Biology, Centre for Ecological and Evolutionary Synthesis (CEES), University of Oslo, Oslo, Norway and ¹⁶Institute of Avian Research, 'Vogelwarte Helgoland', Wilhelmshaven, Germany.

¹⁷Current address: Department of Biology, University of Oslo, PO Box 1066B, Oslo N-0316, Norway.

¹⁸Current address: Department of Animal Production, Tuscia University, 01100 Viterbo, Italy.

Correspondence: Professor CR Primmer, Department of Biology, University of Turku, Vesilinnantie 5, Turku 20014, Finland.

E-mail: craig.primmer@utu.fi

Received 31 May 2011; revised 23 August 2011; accepted 2 September 2011; published online 26 October 2011

as the degree to which feathers reflect light at UV (300–400 nm). This plumage characteristic has also been shown to vary geographically (Sirkiä, 2011) and has been suggested to be a quality indicator, which is influenced by sexual selection in this species (Lehtonen *et al.*, 2009a; Sirkiä and Laaksonen, 2009).

The candidate gene approach has proven a prolific means for detecting genes and gene regions that underlie phenotypic variation in adaptively significant traits (Nachman *et al.*, 2003). The method makes use of the finding that evolutionary divergence does not necessarily erode similarities in gene function between distant lineages (Fitzpatrick *et al.*, 2005). The genetic basis of the expression of melanin traits is quite well-known and more than a hundred loci that classify as 'colour genes', that is, genes that influence pigmentation, have been identified in vertebrates (Bennett and Lamoreux, 2003). Similarly, the molecular mechanisms that underlie feather morphogenesis (Widelitz *et al.*, 2003) and components of avian vision (Okano *et al.*, 1994) have received considerable attention. Thus, there exists a selection of candidate genes for examining molecular variation in gene regions, which may underlie pigment-based colouration, structural colouration and variation associated with the perception of colour.

We have previously shown that the extent of phenotypic differentiation for melanin-based dorsal plumage colouration (P_{ST}) greatly exceeds that observed for assumedly neutral genetic variation (F_{ST}) in populations of pied flycatchers across their breeding range (Lehtonen *et al.*, 2009a). Assuming that P_{ST} approximates Q_{ST} this is indicative of the non-neutral evolution of dorsal plumage colouration in pied flycatcher males. As mentioned, structural UV colouration has likewise been found to show signals of adaptive significance in pied flycatchers. UV colouration is a characteristic that appears to be used by females when choosing a mate (Siitari *et al.*, 2002; Lehtonen *et al.*, 2009b) particularly early on in the mating season (Sirkiä and Laaksonen, 2009). It is thus feasible that the sensitivity to signals conveyed by melanin-based and/or structural colours is also of adaptive significance and thus shaped by selection in this species.

Here, we examine among-population patterns of genetic diversity and differentiation in single-nucleotide polymorphism (SNP) variation within or near three different categories of candidate genes:

genes associated with melanin-based pigmentation ('pigmentation genes'), genes involved in feather morphogenesis ('structural colour genes') and genes that influence the perception of colour ('vision genes'). We also include genes not known to be associated with any of the above three phenotypes ('other genes'). Further, for the purpose of having a neutral baseline (that is, an estimate of the degree of genetic differentiation in the absence of selection) generated from a larger number of markers in some of the analyses, we include microsatellite data genotyped on the same individuals in an earlier study. Our aim was to examine whether any of the examined genes show signals of non-neutral evolution across the species range, potentially revealing genetic regions that underlie adaptively significant variation in this species.

MATERIALS AND METHODS

A total of 528 pied flycatcher males belonging to 17 distinct nest box sites across the breeding range were caught and sampled during the breeding seasons of 1994–2009 (see Table 1 for population-specific information). The population samples are as in reference Lehtonen *et al.* (2009a) except that a British population sample was added to the data set (see Figures 1 and 2 for sampling sites). A blood or feather sample was collected from each male and stored at -20°C or room temperature, respectively, until further use for genetic analyses. The dorsal colour of each male (from all populations except Lingen, Germany) was also recorded and classified according to the Drost scale (Drost, 1936; Glutz von Blotzheim and Bauer, 1993). The score has seven classes, ranging from 1 (fully black head and back) to 7 (fully brown head and back), and has been routinely used for numerically describing pied flycatcher plumage colour for decades (for example, Lundberg and Alatalo, 1992). The sample collectors were experienced users of the colour scoring technique. A subset of the birds was nonetheless re-scored from photographs (by PMS) in order to confirm that the colour scoring had been executed in a uniform manner across populations. The microsatellite data were as in reference Lehtonen *et al.* (2009a), except for the addition of genotypes for the British population sample that were generated using the same methodology.

Laboratory analyses

Candidate gene identification. The first phase of the laboratory work entailed screening of candidate gene sequences to identify polymorphic sites. Candidate genes were chosen based on literature searches for genes known to be associated

Table 1 Population details of the samples included in the study

Population details Area	Country	Abbreviation	Coordinates	Year	n_{samples}	H_e
Dartmoor	United Kingdom	UK	50°36' N 3°43' W	2009	37	0.31
Drenthe*	The Netherlands	Neth	52°52' N 6°17' E	2008	35	0.30
Jeseniky Mountains	Czech Republic	Czech	49°57' N, 17° 09' E	1994–2002	22	0.36
Karelia*	Russia	Rus (Kar)	60°46' N 32°48' E	2008	35	0.29
Kilingi-Nõmme	Estonia	Est	58°8' N 24°59' E	2008	29	0.33
Kraslava	Latvia	Lat	55°53' N 27°11' E	2008	25	0.32
Moscow region*	Russia	Rus (Mosc)	55°44' N 36°51' E	2008	33	0.30
La Hiruela	Spain	Sp (Hir)	41°4' N 3°27' W	2008	53	0.30
Lingen	Germany	Ger	52°27' N 7°15' E	1998–1999	32	0.32
Lund	Sweden	Swe (Lu)	55°39' N 13°55' E	2008	33	0.35
Sørkedalen	Norway	Nor (Sør)	60°1' N 10°37' E	1995–2008	26	0.32
Skibotn*	Norway	Nor (Ski)	69°20' N 20°44' E	2007	25	0.28
Revda*	Russia	Rus (Rev)	56° 51' N 59 53' E	2008	28	0.31
Ruissalo	Finland	Fin	60°26' N 22°10' E	2005–2006	37	0.33
Valsain	Spain	Sp (Val)	40°52' N 4°1' W	2008	36	0.29
Vaud*	Switzerland	Switz	46°50' N 6°42' E	2008	26	0.26
Öland	Sweden	Swe (Öl)	57°10' N 16°58' E	2001–2004	16	0.36

H_e : Gene diversity. The population samples that the whole genome amplification procedure was applied to prior to SNP genotyping are indicated by asterisk (*).

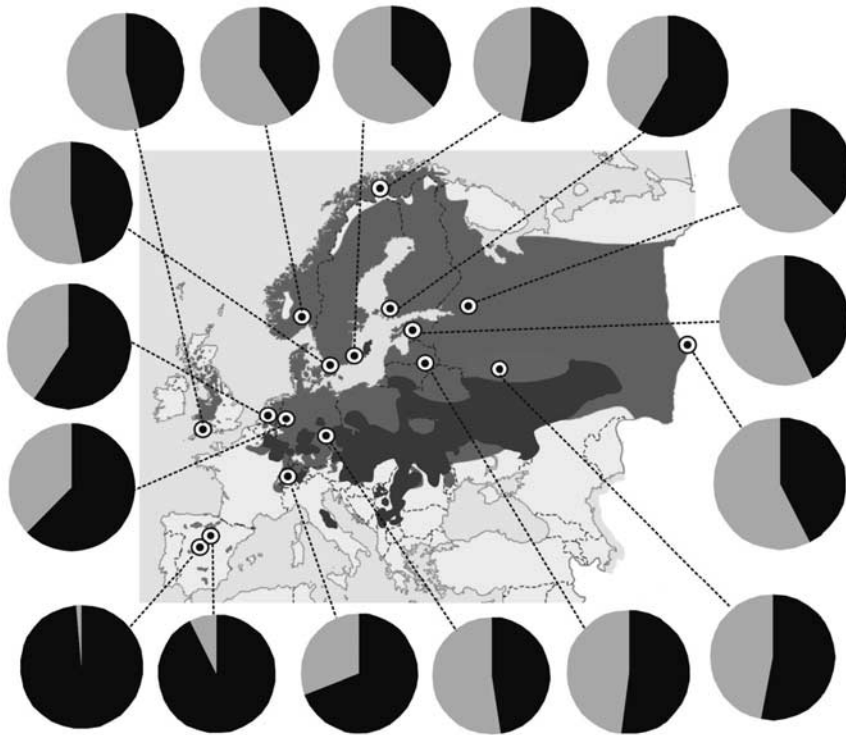


Figure 1 The allele frequencies of the SNP locus in the follistatin gene in all sampled populations. The darker grey shade on the map indicates zones where the pied flycatcher breeds in sympatry with the collared flycatcher and the lighter grey shade indicates the allopatric breeding distribution.

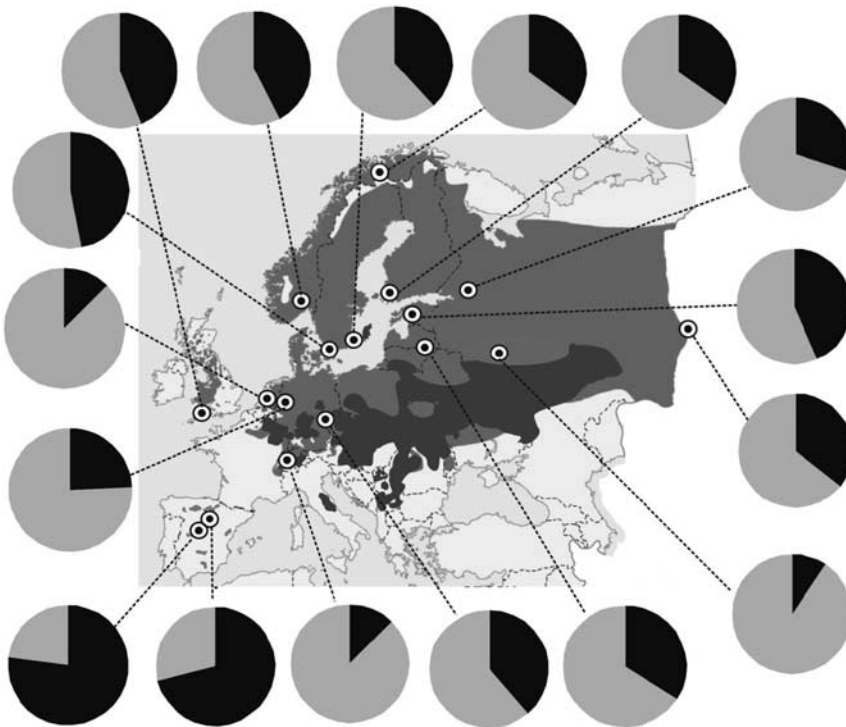


Figure 2 The allele frequencies of the SNP locus near the SWS1 opsin gene in all sampled populations. The darker grey shade on the map indicates zones where the pied flycatcher breeds in sympatry with the collared flycatcher and the lighter grey shade indicates the allopatric breeding distribution.

with each of the three above-described functional groups (pigmentary or structural colour or vision) in birds or mammals. We additionally screened genes not known to be associated with either the production or perception of

colour to provide a comparison (hereafter referred to as ‘other’ genes). Eight of the genes chosen for SNP screening had previously been characterized in the pied flycatcher (2 pigmentation genes (Buggiotti, 2007); 1 vision gene, 5 other

genes (Primmer *et al.*, 2002)). For genes not previously characterized, primers were designed in gene regions conserved between chicken (*Gallus gallus*) and zebra finch (*Taeniopygia guttata*) using the targeted gene approach (see for example, Primmer *et al.*, 2002 for more details). When no suitably long region that was conserved between the two species could be found, primer design was based on the zebra finch sequence owing to its closer evolutionary affinity with pied flycatchers. We designed 73, 2 and 4 primer pairs aiming to amplify regions from 20 pigmentation genes, 1 structural colour gene and 2 vision genes, respectively. One further primer pair was designed to amplify a gene region belonging to the 'other' category. Primers were designed using the program Primer3 (Rozen and Skaletsky, 1999).

DNA extraction was performed as described previously by Lehtonen *et al.*, (2009a). PCRs were performed in a total volume of 20 µl using 20–50 ng of DNA template, using AmpliTaq Gold, BioTaq (Bioline, London, UK), or the Qiagen (Hilden, Germany) multiplex buffer. Unincorporated primers and dNTPs were removed from PCR products by exonuclease-1-shrimp alkaline phosphatase treatment. Sequencing was performed using the BigDye terminator (version 3) chemistry (Applied Biosystems, Foster City, CA, USA) according to the manufacturer's instructions, using one of the primers that had been used to amplify the fragment. The primer sequences are listed in Supplementary Appendix 1.

SNP screening in non-model avian species has typically been performed using 6–18 individuals (for example, Primmer *et al.*, 2002). To minimize the effects of ascertainment bias, the polymorphism discovery panel consisted of individuals from geographically, genetically and phenotypically diverse populations (Rosenblum and Novembre, 2007). A total of 1–2 individuals from at least five of the populations (5–10 individuals in total) were sequenced in the initial SNP discovery phase. The polymorphism discovery panel always included at least two individuals from each of the three main population subgroups (northern, central and southern; Lehtonen *et al.*, 2009a). Additionally, each polymorphism discovery panel always contained at least two birds of each of the extreme colour phenotypes (Drost class-2 and 7). Owing to limitations on the quantity of DNA available, the same 5–10 individuals were not used for screening polymorphisms from each gene.

The criterion for SNP validation was that polymorphisms were either detectable in both the forward and reverse sequence, or, in cases when sequencing had only been successful in one direction, that the polymorphism was observed in more than one individual (Primmer *et al.*, 2002). Sequence alignment was performed using BioEdit (Hall, 1999) and SNP identification was performed by eye. For population genetic and outlier locus analyses, only loci with a rare allele frequency of 0.05 or higher were included in the data set.

Additionally, we selected SNP loci from 1 pigmentation gene (the melanocortin-1 receptor; MC1R), 1 feather morphogenesis gene (follistatin, *FST*) and 7 'other' genes (aconitase-1, *ACO1*; prolactin receptor, *PRLR*; histidine triad nucleotide-binding protein, *HINT1*; RAR-related receptor-β, *RORB*; neurotrophic kinase receptor-2, *NTRK2*; aldolase-B, *ALDOB*; very low-density lipoprotein receptor, *VLDR2*; Buggiotti, 2007) to be included in the panel of SNP markers for large-scale screening. Data from these sequences were not, however, included in the nucleotide diversity calculations described below as the SNPs in these genes had been identified based on sequence data from just two pied flycatcher individuals.

Nucleotide diversity was calculated by using the sequence data generated with the polymorphism discovery panels. The formula used for the calculations was as follows: $\theta = K / (L * [1^{-1} + 2^{-1} + 3^{-1} + \dots + (n-1)^{-1}])$, where K is the number of polymorphic sites, L is the length of the sequence in basepairs, and n is the number of chromosomes screened.

SNP genotyping. The second phase of the laboratory work was genotyping of all the individuals at the polymorphic sites that had been identified. SNP genotyping was performed by using the iPLEX Gold assay on the MassARRAY platform (Sequenom, San Diego, CA, USA) according to the manufacturer's instructions. For populations in which initial DNA quantities were insufficient for the procedure (see Table 1 for specific populations), a whole-genome amplification was performed prior to SNP analysis using the GenomiPhi V2 DNA amplification kit following the manufacturer's instructions (GE Healthcare, Waukesha, WI, USA). SNP genotyping of whole-genome-amplified samples on the MassARRAY platform has been shown to result in accurate genotyping for the vast majority of samples, but some samples tend to be prone to allelic dropout

(Schoenborn *et al.*, 2007). To examine the influence of the procedure on our data, the genotypes of whole-genome-amplified individuals were compared with the sequence information of these individuals prior to whole-genome amplification where available. No discrepancies between the two were detected. Genotypes were determined by using the Typer 4.0 (Sequenom, San Diego, CA, USA) software and exported to a spreadsheet program for further analyses.

Data analyses. The program PHASE (Stephens and Donnelly, 2003) was used to generate haplotypes for gene segments where more than one SNP had been genotyped (*ALASY*, *ALDOB*, *MITF*, *Rhodopsin*, *Tropomyosin*). The program was set to use the default values for the number of iterations (100), thinning interval (1) and burn-in (100). The minimum acceptable probability threshold was set at 0.60. Loci for which the accuracy of the constructed haplotype was estimated to be below this value were regarded as missing data in downstream analyses.

Allele frequencies and observed and expected heterozygosities were calculated for each sample and every locus using the MICROSATELLITE TOOLKIT Excel add-in and FSTAT 2.9.3 (Goudet, 1995).

The program GENEPOP 3.4 (Raymond and Rousset, 1995) was used to test for deviations from Hardy–Weinberg and genotypic linkage equilibria. A sequential Bonferroni-type method was used to correct for multiple testing. The GENEPOP 3.4 program was further used for calculating a global F_{ST} value and pairwise F_{ST} estimates for each population pair. Fisher's exact test as implemented in GENEPOP 3.4 was used to test for differences in allelic frequency distributions between all sample pairs for every locus (total number of different population pair comparisons=136). Unbiased estimates were obtained with 5000 iterations. Probability values over all loci were obtained by the Fisher method as implemented in the program.

To examine whether any of the loci had been affected by divergent selection, we used a single-locus outlier test based on among-population comparisons of genetic diversity and differentiation. One of the main challenges associated with these among-population outlier tests is obtaining the F_{ST} distribution that is expected in the absence of selection (Excoffier *et al.*, 2009). Population structure is known to influence the null distribution and this information should be incorporated into the analyses when available. To address this issue, we chose to use an outlier test, which does not assume uniform distribution of F_{ST} among populations across the sampled range. The method used is an extension to the FDIST approach of (Beaumont and Nichols, 1996) described by Excoffier *et al.*, (2009), and embedded into the ARLEQUIN v. 3.5 analysis package (Excoffier and Lischer, 2010). The population samples were grouped based on population genetic structuring observed using assumedly neutral microsatellite loci in these pied flycatcher population samples (as in reference Lehtonen *et al.*, 2009a; see Supplementary Appendix 2 for details concerning the British sample). The subgroups were as follows: 1: two Spanish populations; 2: Swiss population; 3: British population; and 4: all other populations. The expected F_{ST} distributions were obtained by performing 20 000 simulations, with four groups containing 50 demes. As contemporary outlier tests have been suggested to be relatively prone to type-I errors in the context of balancing selection (Excoffier *et al.*, 2009), we only focused on loci indicated to be affected by divergent selection. To examine the effect of an increased number of sampling sites across the genome on our results, the above-described analysis was also executed with the microsatellite loci incorporated into the data.

As application of multiple outlier tests has been suggested to reduce the number of false positives (for example, Vasemagi and Primmer, 2005), we also analysed the data set using the BAYESFST (Beaumont and Balding, 2004) analysis package. Here F_{ST} is modelled within the Bayesian framework and outliers are nominated based on locus-effect parameters, assumed to be 0 in neutrally evolving loci.

To examine the association between dorsal plumage colour and genotype on an individual level, all pied flycatcher males were assigned to one of three genotypic classes for each locus (homozygous for the common allele, homozygous for the rarer allele or heterozygous) or classified according to haplotype (1–4 for *ALASY*, *ALDOB*, *MITF* and *Tropomyosin*; 1–5 for *Rhodopsin*). The statistical association between genotype and dorsal plumage colour was examined using a Kruskal–Wallis non-parametric test. Because of the potential confounding effect of population structure on genotype–phenotype associations, only the genetically undifferentiated northern and eastern populations (see Lehtonen *et al.*, 2009a) were included in this analysis.

Table 2 Details of the genes in which the SNPs used in this study were identified

	Gene	Abbr.	A/Z	n _{SNP}	I/E	MAF	H _e	F _{ST}	Function
Pigmentation	<i>Microphthalmia-associated transcription factor</i>	MITF	A	2	I	0.10	0.20	0.006	A critical regulator of pigment cell development and survival (Widlund and Fisher, 2003)
	<i>Melanocortin-1 receptor</i>	MC1R	A	1	E	0.07	0.13	0.007	Main role as a switch affecting the type of melanin pigment produced by the melanocytes (for example, Mundy <i>et al.</i> , 2003)
	<i>Myosin-Va</i>	MyoVA	A	1	I	0.17	0.29	0.010	Involved in transport of pigment-containing granules (melanosomes) to peripheral sites (Fukuda <i>et al.</i> , 2002)
	<i>Pallidin</i>	Pldn	A	1	I	0.05	0.09	0.021	Involved in the biogenesis of melanosomes (Falcón-Pérez and Dell'Angelica, 2002)
	<i>Tyrosinase-related protein-1</i>	TYRP1	Z	1	I	0.07	0.13	0.009	Converts dopaquinone, a eumelanin and pheomelanin precursor, into eumelanin (April <i>et al.</i> , 1998; Nadeau <i>et al.</i> , 2007)
Vision	<i>Rhodopsin</i>	Rhod	A	3	I	0.48	0.59	0.014	A photoreceptor protein, which acts as a signal transducer in the rod outer segments of animal eyes. (Lamb, 1986)
	<i>Ultraviolet-sensitive opsin</i>	UV opsin	A	1	I	0.42	0.43	0.116	A visual pigment with its absorption maximum in the UV range. Forms the basis for UV vision (Yokoyama <i>et al.</i> , 1998; Wilkie <i>et al.</i> , 2000)
FM	<i>Follistatin</i>	FST	Z	1	E	0.42	0.44	0.146	A BMP antagonist, expressed in the growing feather bud (Patel <i>et al.</i> , 1999)
Other	<i>5-Aminolaevulinic acid synthase (Ala-synthase)</i>	ALASY	A	2	I	0.43	0.55	0.022	The first and rate-controlling enzyme of haeme biosynthesis (Maguire <i>et al.</i> , 1986)
	<i>Aldolase-B</i>	ALDOB	Z	2	I	0.32	0.22	0.023	An important enzyme for fructose and glucose metabolism (Ito <i>et al.</i> , 1998)
	<i>CEPUS</i>	CEPU	A	1	I	0.16	0.25	0.021	A secreted type of neural glycoprotein belonging to the immunoglobulin-like opioid-binding cell adhesion molecule (OBCAM) subfamily (Kim <i>et al.</i> , 1999)
	<i>Transforming growth factor-β2</i>	TGFβ2	A	1	I	0.10	0.19	0.024	Has a role in growth, development, repair, inflammation and immunity (Clark and Coker, 1998)
	<i>Tropomyosin</i>	Tropom	A	2	I	0.12	0.46	0.035	Binds to actin filaments and regulates the interaction of the filaments with myosin in response to Ca ²⁺ (Wegner, 1979)
	<i>Very low-density lipoprotein receptor</i>	VLDR2	Z	1	E	0.32	0.42	0.026	Associated with the metabolism of fat and cholesterol. Similar to the low-density lipoprotein (LDL) receptor (Takahashi <i>et al.</i> , 1992)

Abbreviations: A, autosomal gene; Abbr., abbreviation used in text; E, the SNP locus is in an exon of the gene; FM, feather morphogenesis-associated gene; H_e, expected heterozygosity; I, the SNP locus is in an intron or near the gene; MAF, minor allele frequency; n_{SNP}, number of SNPs in the gene region used in the study; Z, sex-linked gene. The haplotype frequencies of loci with more than one SNP are listed in Supplementary Appendix 3.

A Mantel test of matrix correspondence (Mantel, 1967) was used to examine the association between patterns of genetic differentiation at neutrally evolving loci (microsatellites; data as in reference Lehtonen *et al.*, (2009a) with the British sample added) and patterns of genetic differentiation at the SNP loci. The relationship between genetic differentiation at SNP loci and geographic distance (km) was also examined using the same methodology. The analyses were executed using GenAIE v 6 (Peakall and Smouse, 2006).

In pied flycatchers, females (the heterogametic sex) are the more extensively dispersing of the sexes (Lundberg and Alatalo, 1992) and it is thus likely that the migration rate of genes on the Z-chromosome is lower than that of genes on the autosomes. In theory, this could inflate inter-population F_{ST} values at Z-linked loci as compared with those observed at autosomal loci even in the absence of selection. To examine the effect of variable migration rates between the two types of chromosomes on the confidence intervals defined for neutrally evolving loci, we used the program FDIST (Beaumont and Nichols, 1996) to perform simulations assuming varying levels of migration for the sex-linked loci. First, we calculated the number of effective migrants per generation (N_em) for the autosomes by using the formula F_{ST}=1/(4N_em+1) (Wright, 1969). This value was then used to calculate variable F_{ST} values for the sex-linked loci assuming differing proportions of sex-linked migration (that is, autosomal N_em*0.1, autosomal N_em*0.2, ..., autosomal N_em*0.9). These F_{ST} values were then used to simulate the confidence intervals that would be expected under neutrality for sex-linked loci under various sex chromosome-to-autosome migratory ratios.

RESULTS

A total of 1–2 individuals from at least five of the populations (5–10 individuals in total) were sequenced in the initial SNP search phase,

amounting to a total of 7665 bp (previously unpublished sequences have been submitted at GenBank under accession numbers HQ659742–HQ659752 and JF305975–JF306023). We identified 34 polymorphic sites. This represents a mean frequency of one SNP per 225 bp of sequence. The mean θ-value was 1.8×10⁻³ (locus range 0 to 8.8×10⁻³), which is somewhat less than that previously reported for pied flycatchers (2.3×10⁻³; Primmer *et al.*, 2002).

Loci with a minor allele frequency of <5% were excluded from the final data set. The final data set had 20 bi-allelic SNP markers from 14 different gene regions (Table 2). Six of the SNPs were identified in five different genes that are candidates for melanin-based pigmentation (*MC1R*, *MITF*, *MyoVA*, *Pallidin* and *TYRP1*; Table 2). One SNP was in a feather morphogenesis-associated gene (*follistatin*; Table 2) and four SNPs in the vision-related genes (*Rhodopsin*, *SWS1 opsin*; Table 2). Five of the 14 gene regions (1 pigmentation gene, 1 vision gene and 2 other genes) were found to have more than one SNP (two SNPs/gene region except for *Rhodopsin*, in which three SNP sites were identified). The major allele frequency of each SNP is listed in Table 2. The allele frequencies of each of the haplotypes are listed in Supplementary Appendix 3.

None of the populations or loci showed consistent deviations from Hardy–Weinberg expectations. No deviations from linkage disequilibrium were found in any pair of loci in any population. The marker-specific heterozygosities across all populations ranged from 0.09 (*Pallidin*) to 0.59 (*Rhodopsin* haplotype; Table 2).

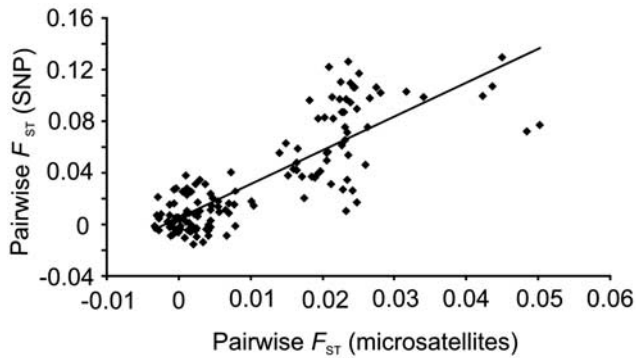


Figure 3 The relationship between pairwise F_{ST} values for SNP markers and microsatellites with the same population pairs (see Lehtonen *et al.*, 2009a).

Tests for pairwise population differentiation across all loci indicated that 84/136 (61.8%) of the population pairs were significantly differentiated at their SNP allele frequencies. The majority of these comparisons (66/84; 78.5%) were pairwise comparisons, including one of the two Spanish populations, the Swiss population or the British population sample. The global F_{ST} value across all loci and populations was 0.04. Pairwise population F_{ST} values across loci ranged from 0 (many population pairs; Supplementary Appendix 2) to 0.14 (Valsain, Spain–Dartmoor, UK; Supplementary Appendix 2). Overall, the pattern of genetic differentiation based on allelic variation at SNP loci resembled that previously reported for these pied flycatcher populations using microsatellite markers (Lehtonen *et al.*, 2009a). In the pairwise comparisons, the populations that showed the most pronounced differentiation (the Swiss, British and two Spanish populations) at their microsatellite allele frequencies were also most differentiated (as quantified by F_{ST}) at the SNP loci examined (Supplementary Appendix 2). Furthermore, two of the three central European populations (the Dutch and German populations) generally showed 10-fold greater differentiation from each of the northern populations than any north–north pairs did from each other (Supplementary Appendix 2). This was also in line with previous findings using microsatellites. Accordingly, the Mantel test of matrix correspondence revealed a significant association between the pairwise F_{ST} values previously reported for these pied flycatcher populations (Lehtonen *et al.*, 2009a) and the pairwise F_{ST} values for the SNP markers ($SP_{XY}=0.054$, $R_{XY}=0.82$, $P=0.001$; where R_{XY} gives the correlation between the two matrices; Figure 3). Genetic differentiation at SNP loci was not found to be associated with geographic distance when measured over all loci ($SP_{XY}=-975.7$, $R_{XY}=-0.19$, $P=0.10$). Kruskal–Wallis test used to test for an association between individual SNP genotypes or haplotypes, and phenotype did not reveal any of the loci to be associated with dorsal plumage colour ($P \geq 0.2$).

Locus-specific F_{ST} values ranged between 0.006 (MITF haplotype) and 0.146 (*folllistatin*; Table 2). Both outlier tests identified *folllistatin* to show significantly higher F_{ST} values than would be expected for neutrally evolving loci (Figure 4 and Supplementary Figure 1). The SWS1 opsin gene was identified as an outlier in the analysis using the hierarchical island model, but not when using the BAYESFST method (Figure 4 and Supplementary Figure 1). Addition of the microsatellite loci to the analyses reduced the width of the confidence intervals defining neutrally evolving loci, and resulted in both SWS1 opsin and *folllistatin* being outliers in both of the tests (Supplementary Figures 2 and 3). None of the microsatellites were identified as outliers by both

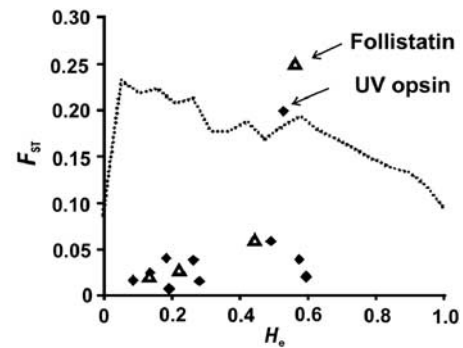


Figure 4 A plot of the relationship between heterozygosity (H_e) and F_{ST} for the SNP loci. The triangles indicate the sex-linked SNP loci and the diamonds indicate the autosomal SNP loci. The dashed line depicts the simulated 95% confidence limit for values expected under neutral evolution.

tests however. The allele frequency differences for both *SWS1 opsin* and *folllistatin* were most pronounced between the Spanish populations and the other populations (Figures 1 and 2). Excluding the population samples for which the DNA had been whole-genome-amplified did not change the results (data not shown).

Two loci were found to significantly deviate from the expectation of no selection (*SWS1-opsin* and *folllistatin*; see below for more extensive detail on the results of the outlier tests). The effect of the *SWS1-opsin* on spectral sensitivity in birds is well established (Ödeen *et al.*, 2009) and it is thus reasonable to hypothesize that this gene may influence the perception of light in pied flycatchers. The effect of the latter on phenotype is not, however, as straightforward. *Follistatin* has been shown to be expressed in the feather bud (Patel *et al.*, 1999), but the way in which this might impact feather structure has not been explored. We thus used the Ingenuity Pathway Analysis program (vs 9.0) to examine whether *folllistatin* has been shown to interact with genes that are known to link with feather structure and have been found to be expressed in the feather bud during its development in pied flycatchers (*Annexin-A2*, *Collagen- α -1(III) chain*, *Collagen- α -2(I) chain*, *Cytokeratin-14*, *Desmoplakin*, *Epidermal fatty acid-binding protein*; (Lehtonen, 2010)). More specifically, we examined whether any of these genes are downstream from *folllistatin* in genetic pathways. In short, the Ingenuity Pathway Analysis program uses information on gene–gene interactions, known from other experiments, to construct a gene network(s) around the gene of interest. We specifically used the program to examine whether any of the above-listed genes that have been detected in developing pied flycatcher feathers are downstream from *folllistatin* in genetic pathways.

The Ingenuity Pathway Analysis pathway analysis exploring the link between *folllistatin* and proteins identified in developing feathers of pied flycatchers revealed 15 different pathways through which *folllistatin* may be connected with the expression of ‘feather structure’ genes. In all cases, the connection was formed through one intermediary molecule. Seven of the pathways linked *folllistatin* with Collagen- α -1(III); 6 with Collagen- α -2(I); 1 with Desmoplakin; and 1 with the Fatty acid-binding protein. A closer examination of these pathways revealed 2 of the 14 connections could bear relevance in the context of feather generation. The first of these is a pathway that connects *folllistatin* with the localization of the tumor necrosis factor- α protein (Jones *et al.*, 2007), which in turn decreases the activation of Collagen- α -2(I) in human dermal fibroblasts (Yamane *et al.*, 2003). The second putatively relevant pathway was one that showed *folllistatin*

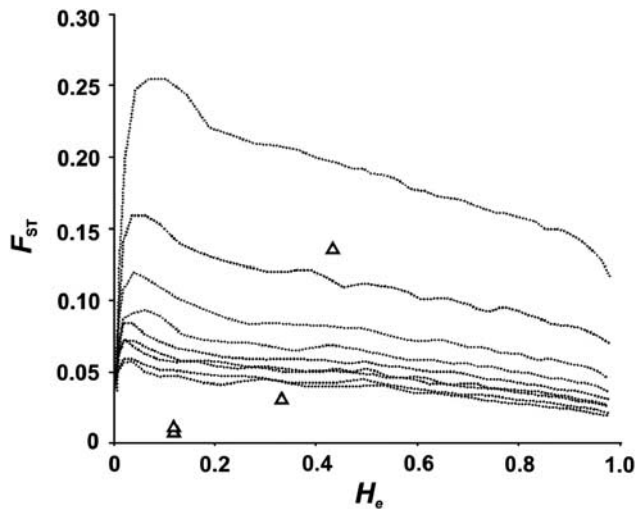


Figure 5 The upper 95% confidence limits from simulations of different levels of migration for the sex chromosomes in comparison with autosomes. Migration levels (from uppermost dashed line to lowest dashed line) are 0.1, 0.2, 0.3...0.9 of the autosomal level. Only the sex-linked markers are depicted here. *Follistatin* remains an outlier in all of the calculations, bar in the event that the sex chromosome migration rate is 0.1 of that of the autosomes.

to decrease the expression of inhibin- β -B (Bilezikjian *et al.*, 2001), which in turn increases the expression of Collagen- α -1(III) in primary mouse keratinocytes (Deng *et al.*, 2006).

The simulations under differing migratory scenarios revealed that the migration rate of sex chromosomes would have to be less than 20% of that observed for the autosomal loci for the Z-linked outlier (*follistatin*) to fall within the limits defined for neutrally evolving loci (Figure 5). This indicates that the observation that *follistatin* is an outlier, is not a false positive produced by the chromosomal location of the SNP.

DISCUSSION

We examined whether there are any signals for selection across the pied flycatcher breeding range in genomic regions near or at candidate genes for melanin-based or structural colouration, or in genes coding for vision. While none of the candidate genes for melanin pigmentation showed signals of selection, we found signals consistent with divergent selection in the *follistatin* and the SWS1 opsin genes, the former a candidate gene for variation in sexually selected structural UV colouration and the latter in the perception of this trait. This is interesting because it matches the prediction that the conspicuousness of a signal used in intra-specific communication will depend on the extent to which it matches the detection ability of the intended recipient (Boughman, 2002).

The *follistatin* gene has been demonstrated to be expressed in the feather bud (Patel *et al.*, 1999), where it antagonizes bone morphogenetic factors (bone morphogenetic proteins; Ohyama *et al.*, 2001). Different subtypes of bone morphogenetic proteins in turn are known to have a particularly prominent role in the process of feather development and re-generation (Noramly and Morgan, 1998). In this study, we further explored how *follistatin* is situated in genetic pathways containing 'feather structure' genes, which have previously been detected in developing pied flycatcher feathers (Lehtonen, 2010). The link between *follistatin* and structural

collagens was the one that featured most prominently in the results (13/15 pathways). Two of these pathways had been detected in cells, which can be hypothesized to be significant for feather structure (keratinocytes and dermal cells). Although the findings are based on research conducted on humans and rodents, they nonetheless provide starting points for further enquiries into the mechanisms through which *follistatin* might influence feather structure during feather development.

Feathers are constantly exposed to variable stressors such as bacteria and mechanical wear, and it is plausible that different environments exert different selection pressures on the properties of the feather. Preliminary results on the structure–performance relationship of flight feathers of chiffchaffs (*Phylloscopus collybita*) and willow warblers (*Phylloscopus trochilus*) are suggestive of feather shaft diameter being associated with resistance to damage (Weber *et al.*, 2005). Feather wear is likely to affect the reflective properties of the plumage and could thus function to signal individual quality. It is of interest, that allelic differentiation for the *follistatin* gene was most pronounced in the Spanish pied flycatcher populations, which have also been found to be among the most highly UV-reflectant populations of those investigated here (Sirkiä, 2011).

The SWS1 opsin gene underlies the ability to detect light at UV wavelengths in many bird species (Odeen *et al.*, 2009). Studies examining the significance of UV reflectance in avian communication have commonly measured inter-individual variation in plumage reflectance in relation to, for example, mate choice (pied flycatchers: Lehtonen *et al.*, 2009a; Sirkiä and Laaksonen, 2009; other avian species: (Bennett *et al.*, 1997; Hunt *et al.*, 1999). Evidently, the effectiveness of the signal in mate choice is dependent on the ability of members of the species to perceive the signal (for example, Cuthill *et al.*, 2000), and thus it is conceivable that both the signal and the receptor are affected by selection. The structural colouration of pied flycatcher males could impact mate choice strategy if it reveals the genetic quality ('good genes') of the mate (see for example, Lehtonen *et al.*, 2009b). Inter-individual variation in the ability to detect light at UV wavelengths would translate to inter-individual variation in the ability to choose a high-quality mate. As for the *follistatin* gene discussed above, the allele frequency differences were most pronounced in the Spanish pied flycatchers, which have relatively high UV reflectance (Sirkiä, 2011).

The sensory drive hypothesis assumes that the environment shapes the sensory system (for example, to maximize the ability to find food), creating a bias that may lead to male traits evolving in the direction that maximally stimulates it. This may either stem from both trait and sensor being shaped by the same environment, or because they are in fact co-evolving (Boughman, 2002). Specific light environments of different habitats likely shape the signals used in communication and signal design has indeed been found to be associated with detectability (Leal and Fleishman, 2004). It is thus feasible that different environments exert selection pressure not only on genes that underlie the phenotypic characteristic itself (the signal), but also on genes that underlie the ability to observe the trait (the ability to perceive the signal). Although we are unable to ascertain whether UV reflectance and the ability to perceive this trait are co-evolving, it is intriguing that allele frequency variation was most pronounced in the Spanish population for both of these genes. The Spanish populations live at a higher altitude and further south than the other populations included in this study. UV radiation increases with altitude and towards the equator (World Health Organisation; www.who.int/mediacentre/factsheets/who271/en). The higher levels

of UV reflectance in these areas could augment the utility of this signal trait in the Spanish populations. Alternatively (or additionally) high UV radiation may lead to an increased need for the birds to protect both the plumage itself and their body from the harmful effects of UV rays, resulting in selection for particular feather characteristics. In the aquatic realm, where variation in photic environments is more pronounced than in terrestrial ones, light conditions have been found to be associated with genetic variation at vision-related genes and peak sensitivities in receptors to co-vary with trait characteristics (for example, Seehausen *et al.*, 2008). Whether variation in photic conditions, and UV radiation in particular, could be a selective force resulting in divergent selection in pied flycatchers, however, remains to be investigated. An examination of several other bird species has demonstrated that cone type and plumage UV maxima (that is, reflection and perception) tend to be associated (Mullen and Pohland, 2008). A useful follow-up study to address this issue would be to sequence the parts of the SWS opsin genes, which confer spectral sensitivity (for example, Ödeen *et al.*, 2009), to examine whether the different populations have different mutations at these sites. Currently, such studies have only been conducted at the inter-specific level. Additionally, the manner in which feather structure is altered by high levels of UV radiation in the different populations could be examined in order to assess whether the feathers differ in their ability to deal with radiation-induced stress.

In addition to differences in environmental conditions, the central Spanish mountain populations also represent a clearly separate 'island' in the pied flycatcher breeding distribution (see Figures 1 and 2) and are more differentiated at assumedly neutral loci from populations in the continuous part of the breeding range (for example, Lehtonen *et al.*, 2009a; the present study). This is hypothesized to be due to historical/post-glacial reasons (Lehtonen *et al.*, 2009a). Pied flycatchers that originate from the more northern and eastern breeding populations fly through this area during their spring migration, but rarely (if ever) stop to breed there (JM personal observation). Additionally, of the pied flycatcher populations from which recruitment and return rates are known, the Spanish populations are among the most philopatric (Lehtonen *et al.*, 2009a). Thus it appears that there is reduced gene flow to and from these populations. A reduction in gene flow could in turn aid in providing the circumstances required for local adaptations to evolve if, for example, different photic conditions lead to divergent selection towards alternative trait optima.

Although outlier loci provide good starting points when aiming to detect the molecular basis of adaptively significant traits, it is important to also consider potential caveats. It would be valuable to examine the chromosomal regions surrounding the *follistatin* and *SWS1 opsin* genes to ascertain the length of the region that has potentially been affected by selection. Before this is done it is not possible to definitively state whether the signals of selection detected here stem from the selection on *follistatin* and *SWS1 opsin*, or are in fact from the SNPs being linked to other genes of different function, which have been targets of selection. Also, as the *follistatin* gene resides on the sex chromosome (Z), which has a smaller effective population size than autosomes, it is more prone to drift. This in turn could result in elevated genetic divergence even in the absence of selection (Borge *et al.*, 2005). On the other hand, sexual selection is thought to be stronger in males than in females and Z-chromosomes spend more time in males (2/3) than females (1/3). Consequently, the chromosomal location of this gene could in fact result in it being exposed to stronger selection pressure than autosomal genes. In any

case, our simulations verified that the result is unlikely to stem from differential levels of gene flow between the sex chromosomes and autosomes. It is also noteworthy that elevated levels of genetic differentiation were only observed in one of the four sex-linked loci included in the study.

Another factor that may have affected the outlier test is the fact that our SNP loci were selected from genes, which are *a priori* expected to influence phenotype. It could be argued that they may all be affected by selection to differing degrees. If this were so, the baseline level of genetic divergence would not reflect neutrality, and could result in some loci, which are affected by selection not being detected, that is, false negatives. We do not have the means to definitively ascertain whether or not this is the case here. However, the inclusion of the selectively neutral microsatellites narrowed the confidence limits set for neutrality, but did not result in the detection of additional outliers. When combining data generated using these two types of markers it is important to recognize that the mutational properties of microsatellites and SNPs differ, and they are thus not ideally suited for such comparisons. In the context of our study, the potential pitfall would be realized in the form of false positives. This was not, however, found to be the case. Furthermore, an examination of the F_{ST} values for individual SNP loci revealed that the distribution of these values was not continuous, and therefore the allelic distribution at the two outlier loci clearly differed from the others. This also speaks for the fact that the threshold set for neutrality in this test was not an arbitrary cut-off point.

In conclusion, *follistatin* and *SWS1 opsin* were found to show molecular signatures of divergent selection across the pied flycatcher breeding range. These two genes are thus strong candidates for further investigation of genes underlying adaptively significant traits in this species. Although molecular data and outlier analyses serve as evidence for selection, they do not provide information concerning the mechanisms of selection or its strength, at least when there is no additional data on the phenotypes or the association between genotype and phenotype. Further validation analyses could include sequence-based neutrality tests (for example, Wood *et al.*, 2008) and/or functional analyses of the different alleles in order to determine whether there is an association with feather morphology and/or sensitivity to reflectance at UV wavelengths (for example, Sabeti *et al.*, 2006).

DATA ARCHIVING

Data have been deposited at Dryad: doi:10.5061/dryad.n33q381s.

CONFLICT OF INTEREST

The authors declare no conflict of interest. There is no overlap between this manuscript and other articles; this manuscript has not been submitted elsewhere; and all the authors have approved its submission in its present form.

ACKNOWLEDGEMENTS

We thank Arne Roseth for assistance with the SNP genotyping and Helene Lampe for providing some extra pied flycatcher samples. We also thank Anti Vasemägi for valuable discussion that greatly improved the manuscript. CIGENE was supported by the functional genomics programme (FUGE) in the Research Council of Norway. This study was funded by the Biological Interactions Graduate School (PKL), MADfish Nordic Network (research visit to Norway; PKL), Centre of Excellence in Evolutionary Genetics and Physiology (PKL and CRP), the Emil Aaltonen foundation (PKL and TL), Academy of Finland (TL), the Russian Academy of Science (AVA), the Norwegian Research

Council (G-PS), the programs for support of Scientific Schools (NSH-3260.2010.4; EB), Scientific-Educational Centres (contract 02.740.11.0279; EB), Ministerio de Ciencia e Innovación (grant CGL2007-61251; JM), Natural England (MDB), the Estonian Science Foundation (grant no. 7476; MM), Estonian Ministry of Education and Science (SF0180004s09; MM), EU through the European Regional Development Fund (Center of Excellence FIBIR; MM), the Latvian Council of Science (IK), MSM619859212 (SB), RFBR 09-04-01690 (AVB) and Lunds Djurskyddsfond (AN).

- Andersson S, Prager M (2006). *Quantifying Colors, Vol I—Mechanisms and Measurements*. Harvard University Press: Cambridge. 589pp.
- April CS, Jackson IJ, Kidson SH (1998). The cloning and sequencing of a cDNA coding for chick tyrosinase-related protein-1. *Biochim Biophys Acta* **1395**: 7–12.
- Beaumont MA, Balding DJ (2004). Identifying adaptive genetic divergence among populations from genome scans. *Mol Ecol* **13**: 969–980.
- Beaumont MA, Nichols RA (1996). Evaluating loci for use in the genetic analysis of population structure. *Proc Natl Acad Sci USA* **263**: 1619–1626.
- Bennett ATD, Cuthill IC, Partridge JC, Lunau K (1997). Ultraviolet plumage colors predict mate preferences in starlings. *Proc Natl Acad Sci USA* **94**: 8618–8621.
- Bennett DC, Lamoreux ML (2003). The color loci of mice—a genetic century. *Pigment Cell Res* **16**: 333–344.
- Bilezikjian LM, Corrigan AZ, Blount AL, Chen Y, Vale WW (2001). Regulation and actions of Smad7 in the modulation of activin, inhibin, and transforming growth factor-beta signaling in anterior pituitary cells. *Endocrinology* **142**: 1065–1072.
- Borge T, Webster MT, Andersson G, Saetre GP (2005). Contrasting patterns of polymorphism and divergence on the Z chromosome and autosomes in two *Ficedula* flycatcher species. *Genetics* **171**: 1861–1873.
- Boswell T, Takeuchi S (2005). Recent developments in our understanding of the avian melanocortin system: its involvement in the regulation of pigmentation and energy homeostasis. *Peptides* **26**: 1733–1743.
- Boughman JW (2002). How sensory drive can promote speciation. *Trends Ecol Evol* **17**: 571–577.
- Buggiotti L (2007) Avian evolutionary genomics: studies of *Ficedula* flycatchers. PhD thesis, University of Turku, Turku.
- Clark DA, Coker R (1998). Transforming growth factor-beta (TGF-beta). *Int J Biochem Cell Biol* **30**: 293–298.
- Cuthill IC, Partridge JC, Bennett ATD, Church SC, Hart NS, Hunt S (2000). *Ultraviolet Vision in Birds Advances in the Study of Behavior*. Academic Press Inc.: San Diego. Vol. 29, pp 159–214.
- Deng MX, Chen WL, Takatori A, Peng ZM, Zhang L, Mongan M *et al.* (2006). A role for the mitogen-activated protein kinase kinase 1 in epithelial wound healing. *Mol Biol Cell* **17**: 3446–3455.
- Drost R (1936). Über das Brutkleid männlicher Trauerfliegenfänger. *Muscicapula hypoleuca*. *Vogelzug* **6**: 179–186.
- Endler JA, Westcott DA, Madden JR, Robson T (2005). Animal visual systems and the evolution of color patterns: sensory processing illuminates signal evolution. *Evolution* **59**: 1795–1818.
- Excoffier L, Hofer T, Foll M (2009). Detecting loci under selection in a hierarchically structured population. *Heredity* **103**: 285–298.
- Excoffier L, Lischer HEL (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resources* **10**: 564–567.
- Falcon-Perez JM, Dell’Angelica EC (2002). The pallidin (Pldn) gene and the role of SNARE proteins in melanosome biogenesis. *Pigment Cell Res* **15**: 82–86.
- Fitzpatrick MJ, Ben-Shahar Y, Smid HM, Vet LEM, Robinson GE, Sokolowski MB (2005). Candidate genes for behavioural ecology. *Trends Ecol Evol* **20**: 96–104.
- Fukuda M, Kuroda TS, Mikoshiba K (2002). Slac2- α /melanophilin, the missing link between Rab27 and myosin Va—implications of a tripartite protein complex for melanosome transport. *J Biol Chem* **277**: 12432–12436.
- Goudet J (1995). FSTAT (Version 1.2): a computer program to calculate *F*-statistics. *J Hered* **86**: 485–486.
- Glutz von Blotzheim UN, Bauer KM (1993). *Handbuch der Vögel Mitteleuropas*. Aula-Verlag, Wiesbaden.
- Hall TA (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Series* **41**: 95–98.
- Hubbard JK, Uy JAC, Hauber ME, Hoekstra HE, Safran RJ (2010). Vertebrate pigmentation: from underlying genes to adaptive function. *Trends Genet* **26**: 231–239.
- Hunt S, Cuthill IC, Bennett ATD, Griffiths R (1999). Preferences for ultraviolet partners in the blue tit. *Animal Behav* **58**: 809–815.
- Ito JI, Kuzumaki T, Otsu K, Iuchi Y, Ishikawa K (1998). Hormonal regulation of aldolase B gene expression in rat primary cultured hepatocytes. *Arch Biochem Biophys* **350**: 291–297.
- Jones KL, Mansell A, Patella S, Scott BJ, Hedger MP, de Kretser DM *et al.* (2007). Activin A is a critical component of the inflammatory response, and its binding protein, follistatin, reduces mortality in endotoxemia. *Proc Natl Acad Sci USA* **104**: 16239–16244.
- Kim DS, Rhew TH, Moss DJ, Kim JY (1999). cDNA cloning of the CEPUS, a secreted type of neural glycoprotein belonging to the immunoglobulin-like opioid binding cell adhesion molecule (OBCAM) subfamily. *Mol Cells* **9**: 270–276.
- Lamb TD (1986). Transduction in vertebrate photoreceptors: the roles of cyclic GMP and calcium. *Trends Neurosci* **9**: 224–228.
- Leal M, Fleishman LJ (2004). Differences in visual signal design and detectability between allopatric populations of Anolis lizards. *Am Nat* **163**: 26–39.
- Lehtonen PK (2010) The molecular mechanisms and evolutionary significance of plumage colour variation in pied flycatchers (*Ficedula hypoleuca*). PhD thesis, University of Turku, Turku.
- Lehtonen PK, Laaksonen T, Artemyev AV, Belskii E, Both C, Bures S *et al.* (2009a). Geographic patterns of genetic differentiation and plumage colour variation are different in the pied flycatcher (*Ficedula hypoleuca*). *Mol Ecol* **18**: 4463–4476.
- Lehtonen PK, Primmer CR, Laaksonen T (2009b). Different traits affect gain of extrapair paternity and loss of paternity in the pied flycatcher, *Ficedula hypoleuca*. *Animal Behav* **77**: 1103–1110.
- Lundberg A, Alatalo RV (1992). *The Pied Flycatcher*. Poyser: London.
- Maguire DJ, Day AR, Borthwick IA, Srivastava G, Wigley PL, May BK *et al.* (1986). Nucleotide sequence of the chicken-5-aminolevulinate synthase gene. *Nucleic Acids Res* **14**: 1379–1391.
- Mantel N (1967). Detection of disease clustering and a generalized regression approach. *Cancer Res* **27**(2P1): 209–220.
- Mullen P, Pohland G (2008). Studies on UV reflection in feathers of some 1000 bird species: are UV peaks in feathers correlated with violet-sensitive and ultraviolet-sensitive cones? *IBIS* **150**: 59–68.
- Mundy NI (2005). A window on the genetics of evolution: MC1R and plumage colouration in birds. *Proc R Soc B Biol Sci* **272**: 1633–1640.
- Mundy NI, Kelly J, Theron E, Hawkins K (2003). Evolutionary genetics of the melanocortin-1 receptor in vertebrates. In: Cone RD (ed) *Melanocortin System*, Vol. 994, pp 307–312.
- Nachman MW, Hoekstra HE, D’Agostino SL (2003). The genetic basis of adaptive melanism in pocket mice. *Proc Natl Acad Sci USA* **100**: 5268–5273.
- Nadeau NJ, Mundy NI, Gourichon D, Minvielle F (2007). Association of a single-nucleotide substitution in TYRP1 with roux in Japanese quail (*Coturnix japonica*). *Animal Genet* **38**: 609–613.
- Noramly S, Morgan BA (1998). BMPs mediate lateral inhibition at successive stages in feather tract development. *Development* **125**: 3775–3787.
- Odeen A, Hart NS, Hastad O (2009). Assessing the use of genomic DNA as a predictor of the maximum absorbance wavelength of avian SWS1 opsin visual pigments. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* **195**: 167–173.
- Ohyama A, Saito F, Ohuchi H, Noji S (2001). Differential expression of two BMP antagonists, gremlin and Follistatin, during development of the chick feather bud. *Mech Dev* **100**: 331–333.
- Okano T, Yoshizawa T, Fukada Y (1994). Pinopsin is a chicken pineal photoreceptive molecule. *Nature* **372**: 94–97.
- Patel K, Makarenkova H, Jung HS (1999). The role of long range, local and direct signalling molecules during chick feather bud development involving the BMPs, Follistatin and the Eph receptor tyrosine kinase Eph-A4. *Mech Dev* **86**: 51–62.
- Peakall R, Smouse PE (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes* **6**: 288–295.
- Primmer CR, Borge T, Lindell J, Saetre GP (2002). Single-nucleotide polymorphism characterization in species with limited available sequence information: high nucleotide diversity revealed in the avian genome. *Mol Ecol* **11**: 603–612.
- Raymond M, Rousset F (1995). GENEPOP (VERSION 1.2)—population-genetics software for exact tests and ecumenicism. *J Hered* **86**: 248–249.
- Rosenblum EB, Novembre J (2007). Ascertainment bias in spatially structured populations: a case study in the eastern fence lizard. *J Hered* **98**: 331–336.
- Rozen S, Skaletsky H (1999). *Primer3 on the WWW for General Users and for Biologist Programmers*, Vol. 132, pp 365–386.
- Sabeti PC, Schaffner SF, Fry B, Lohmueller J, Varilly P, Shamovsky O *et al.* (2006). Positive natural selection in the human lineage. *Science* **312**: 1614–1620.
- Saetre GP, Saether SA (2010). Ecology and genetics of speciation in *Ficedula* flycatchers. *Mol Ecol* **19**: 1091–1106.
- Schoenborn V, Gohlke H, Heid IM, Illig T, Utermann G, Kronenberg F (2007). Sample selection algorithm to improve quality of genotyping from plasma-derived DNA: to separate the wheat from the chaff. *Hum Mutat* **28**: 1141–1149.
- Seehausen O, Terai Y, Magalhaes IS, Carleton KL, Mrosso HDJ, Miyagi R *et al.* (2008). Speciation through sensory drive in cichlid fish. *Nature* **455**: 620–U623.
- Siitari H, Honkavaara J, Huhta E, Viitala J (2002). Ultraviolet reflection and female mate choice in the pied flycatcher, *Ficedula hypoleuca*. *Animal Behav* **63**: 97–102.
- Sirkkiä PM (2011) Maintenance of phenotypic variation in plumage colouration in a passerine bird. PhD thesis, University of Turku, Turku.
- Sirkkiä PM, Laaksonen T (2009). Distinguishing between male and territory quality: females choose multiple traits in the pied flycatcher. *Animal Behav* **78**: 1051–1060.
- Stephens M, Donnelly P (2003). A comparison of Bayesian methods for haplotype reconstruction from population genotype data. *Am J Hum Genet* **73**: 1162–1169.
- Takahashi S, Kawarabayasi Y, Nakai T, Sakai J, Yamamoto T (1992). Rabbit very low-density-lipoprotein receptor—a low-density-lipoprotein receptor-like protein with distinct ligand specificity. *Proc Natl Acad Sci USA* **89**: 9252–9256.
- Vasemagi A, Primmer CR (2005). Challenges for identifying functionally important genetic variation: the promise of combining complementary research strategies. *Mol Ecol* **14**: 3623–3642.

- Weber TP, Borguud J, Hedenstrom A, Persson K, Sandberg G (2005). Resistance of flight feathers to mechanical fatigue covaries with moult strategy in two warbler species. *Biol Lett* **1**: 27–30.
- Wegner A (1979). Equilibrium of the actin–tropomyosin interaction. *J Mol Biol* **131**: 839–853.
- Widelitz RB, Jiang TX, Yu MK, Shen T, Shen JY, Wu P *et al.* (2003). Molecular biology of feather morphogenesis: A testable model for evo-devo research. *J Exp Zool B Mol Dev Evol* **298B**: 109–122.
- Widlund HR, Fisher DE (2003). Microphthalmia-associated transcription factor: a critical regulator of pigment cell development and survival. *Oncogene* **22**: 3035–3041.
- Wilkie SE, Robinson PR, Cronin TW, Poopalasundaram S, Bowmaker JK, Hunt DM (2000). Spectral tuning of avian violet- and ultraviolet-sensitive visual pigments. *Biochemistry* **39**: 7895–7901.
- Wood HM, Grahame JW, Humphray S, Rogers J, Butlin RK (2008). Sequence differentiation in regions identified by a genome scan for local adaptation. *Mol Ecol* **17**: 3123–3135.
- Wright S (1969) vol. 2, the Theory of Gene Frequencies University of Chicago Press: Chicago, IL.
- Yamane K, Ihn H, Asano Y, Jinnin M, Tamaki K (2003). Antagonistic effects of TNF-alpha on TGF-beta signaling through downregulation of TGF-beta receptor type II in human dermal fibroblasts. *J Immunol* **171**: 3855–3862.
- Yokoyama S, Radlwimmer FB, Kawamura S (1998). Regeneration of ultraviolet pigments of vertebrates. *FEBS Lett* **423**: 155–158.
- Ödeen A, Hart NS, Håstad O (2009). Assessing the use of genomic DNA as a predictor of the maximum absorbance wavelength of avian SWS1 opsin visual pigments. *J Comp Physiol* **195**: 167–173.

Supplementary Information accompanies the paper on Heredity website (<http://www.nature.com/hdy>)