

# Natural selection on mitochondrial DNA in *Parus* and its relevance for phylogeographic studies

Robert M. Zink

*Bell Museum of Natural History, University of Minnesota, St. Paul, MN 55108, USA (rzink@cbs.umn.edu)*

I examined mitochondrial ND2 sequences in species of tits (genus *Parus*) to determine whether substitutions were neutral or under selection. Haplotype trees in these species are shallow and geographically unstructured, which could be a signature of recent demographic expansion or purifying selection. McDonald–Kreitman tests revealed a pattern of excess replacement polymorphisms for closely related taxa. Replacement substitutions tended to be less common in the basal parts of haplotype trees, suggesting that they are selected against over evolutionary time. Thus, the pattern of selection is consistent with a model of mildly deleterious haplotypes. To investigate the significance of this selection, the ND2 gene was subdivided into surface and transmembrane portions, and synonymous–non-synonymous sites, and  $F_{ST}$ -values (the amount of variance distributed among populations) computed for each. Among sampling localities in *P. major*, both partitions gave similar estimates of  $F_{ST}$ , which is consistent with demographic expansion as the cause of the shallow trees. In *P. montanus*, two localities contained high percentages of individuals that had a single replacement substitution in the transmembrane portion of the gene, which inflated  $F_{ST}$ -values relative to the surface sites. Thus, the interpretation of the genetic differentiation of these two sites could be either geographical isolation or selection. Given that the sole substitution causing the high  $F_{ST}$ -values was a replacement one, selection is implicated. However, this substitution occurred in only 2 of 12 localities, revealing that overall the phylogeographic pattern was not biased by selection, once the cause of the two high pairwise  $F_{ST}$ -comparisons was recognized. Investigators should consider the source of significant genetic heterogeneity within species, to determine if the variation is likely caused by demographic isolation or selection. However, it seems unlikely that most phylogeographic analyses are invalidated by the effects of natural selection.

**Keywords:** natural selection; ND2; mitochondrial DNA; selective neutrality; McDonald–Kreitman test

## 1. INTRODUCTION

A plethora of phylogeographic studies (Avice 2000) exists based on variation in the DNA found in the mitochondrion (mtDNA). A basic premise is that the variation observed is selectively neutral and that genetic patterns reflect the historical demography of populations and not the effects of natural selection. Testing for natural selection has been sporadic or superficial, despite the importance of the neutrality assumption. Several studies have found evidence for selection on mtDNA (Messier & Stewart 1997; Wise *et al.* 1998; Creevey & McInerney 2002; Eyre-Walker 2002; Jansa *et al.* 2003). This has led some researchers (e.g. Ballard & Whitlock 2004) to challenge the evolutionary interpretation of mtDNA variation in some contexts. Thus, it is timely to investigate the nature of selection and its significance for phylogeographic inquiry, which has become a major area of study (Avice 2000).

A large number of mtDNA phylogeographic studies have been conducted on birds (summaries in Avice & Walker 1998; Zink 2004). A typical result of many studies, especially those of passerine birds (songbirds), is a shallow (sequences differ by less than 1%) and geographically unstructured haplotype tree. This result is interpreted to be a signature of population increase and range expansion, and that populations are connected either currently or in the very recent past by gene flow (Zink *et al.* 2003). It is also

possible that shallow haplotype trees are due to purifying selection. In the latter case, our understanding of recent population history could be biased because the tree structure does not represent demographic history.

Wise *et al.* (1998) studied variation in ND2 in *Homo sapiens* and *Pan troglodytes* and found significant differences in the  $F$ -value of Fu (1997) in both species, indicating either non-neutral sequence evolution or recent demographic expansion. The authors subdivided the ND2 gene into surface and transmembrane portions. Using the McDonald–Kreitman (1991) test, Wise *et al.* (1998) found evidence for selection in humans, in the form of an excess of replacement polymorphisms within the transmembrane portion of the gene only. This would not be expected if demographic expansion were the cause of the significant Fu's  $F$ -values because each region should be affected concordantly. This type of result calls into question the use of mtDNA data for phylogeographic inference because divergence within and among species might not be due solely to drift.

In this study, I explore variation in newly acquired sequences of the ND2 gene in several species of tits (genus *Parus*) from Eurasia. Two of the species, *P. major* and *P. montanus* have been subjected to multiple phylogeographic analyses, based on sequence variation in the non-coding mitochondrial control region (Kvist *et al.* 2001,

Table 1. Basic sequence statistics for species of *Parus* at ND2. (Values are not presented for *P. bokharensis* because only four individuals were sequenced.)

taxon/region	S	$\pi$	$\theta$	Tajima's <i>D</i>	Fu and Li's <i>D</i>	Fu and Li's <i>F</i>	Fu's <i>F</i>
<i>P. major</i> ( <i>n</i> = 87)							
all sites	34	0.0012	0.0065	-2.57**	-5.30*	-5.10*	-26.17**
surface	10	0.0014	0.0054	-1.94*	-2.74*	-2.92*	-10.23**
transmembrane	24	0.0010	0.0072	-2.58*	-5.35*	-5.15*	-24.07*
<i>Parus minor</i> ( <i>n</i> = 29)							
all sites	13	0.0012	0.0034	-2.22**	-3.01*	-3.24*	-13.14**
surface	2	0.0005	0.0014	-1.25	-0.73	-1.01	-1.63
transmembrane	11	0.0015	0.0046	-2.20*	-3.20*	-3.40*	-12.70**
<i>Parus montanus</i> ( <i>n</i> = 139)							
all sites	63	0.0021	0.0114	-2.55**	-5.43**	-5.00*	—
surface	21	0.0024	0.0112	-2.25**	-3.80**	-3.84*	-28.40**
transmembrane	42	0.0020	0.0115	-2.51**	-5.16*	-4.87*	-34.48**

\* $p < 0.05$ ; \*\* $p < 0.01$ .

Table 2. (a) Synonymous and non-synonymous mutations in ND2 between three species of *Parus*. (b) As in (a) but for transmembrane sites only.

(Top and bottom halves of the matrix contain comparisons for surface and all sites, respectively. Groups of four numbers off diagonal refer to numbers of synonymous (top row) and non-synonymous (bottom row) mutations fixed between species (left column) and polymorphic within species (right column). For example, comparing *P. major* and *P. minor*, the ratio of synonymous : non-synonymous mutations fixed between species was 23 : 0 and the ratio for polymorphic sites was 36 : 10 ( $p = 0.014$ ).  $p$ -values are from Fisher's exact test.)

	<i>P. major</i>	<i>P. minor</i>	<i>P. montanus</i>	<i>P. bokharensis</i>
(a)				
<i>P. major</i>		6 0	10 2	23 13
<i>P. minor</i>	23 0	36 10	$p = 0.53$	$p = 0.008$
<i>P. montanus</i>	70 35	77 19	68 34	63 15
<i>P. bokharensis</i>	6 2	29 7	30 2	12 3
	$p = 0.014$	$p = 0.038$	$p = 0.05$	$p = 0.028$
		$p = 0.89$	$p = 0.38$	$p = 0.028$
				$p = 0.94$
				$p = 0.19$
				$p = 0.018$
(b)				
<i>P. minor</i>	17 0	26 8		
<i>P. montanus</i>	50 22	47 16	51 22	42 12
<i>P. bokharensis</i>	4 1	20 6	21 1	11 2
	$p = 0.040$	$p = 0.57$	$p = 0.43$	$p = 0.40$
		$p = 0.67$	$p = 0.64$	$p = 0.40$

2003; Pavlova 2004). In both species (*sensu stricto*), a shallow haplotype tree was discovered, with no significant phylogeographic divisions. This type of tree could result from recent demographic expansion, expected for species after amelioration of Ice Age climates. Alternatively, purifying selection might be responsible for the lack of genetic heterogeneity over the ranges of each species. My goals were to test for deviations from selective neutrality and to

determine whether any observed departures influenced phylogeographic interpretations.

## 2. MATERIAL AND METHODS

Three focal taxa included were great tit (*Parus major*, *n* = 87), Japanese tit (*P. minor*, *n* = 29), and willow tit (*P. montanus*, *n* = 139); details of sampling are given in Pavlova (2004). Four individuals of the Turkestan tit (*Parus bokharensis*), the sister

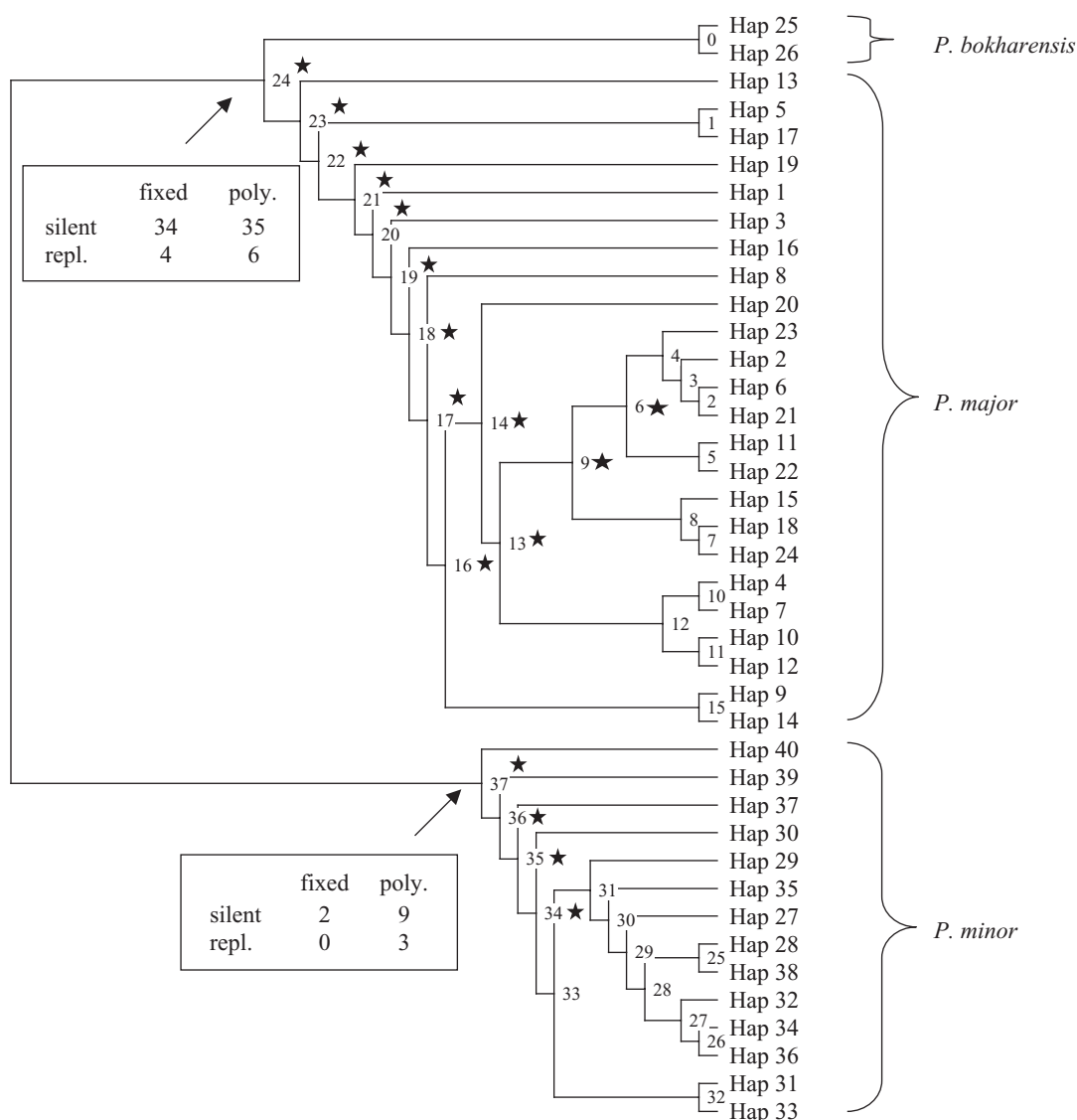


Figure 1. Neighbour-joining tree based on values of *D* for *P. major*, *P. bokharensis* and *P. minor* (root). Stars refer to nodes with significant deficiencies of replacement substitutions. The two boxes show results of M–K tests (1991) that were not significant but which show the nature of substitutions at the bases of major clades.

species to *P. major* (Kvist *et al.* 2003), were used. A single individual of the marsh tit (*P. palustris*) was used as an outgroup to *P. montanus*. GenBank accession numbers are AY732496-733007, AY734248-734251. Basic sequence statistics were calculated for each species using DnaSP (Rozas *et al.* 2003), including: nucleotide diversity ( $\pi$ ), theta ( $\theta$ ), *D* (Tajima 1989), *D* (Fu & Li 1993), *F* (Fu & Li 1993), and *F* (Fu 1997). Using a model (Persson & Argos 1994) of ND2, the sequences were divided into surface (372 bp) and transmembrane (666 bp) sections (the stop codon was ignored). McDonald–Kreitman (1991) (M–K) tests implemented in DnaSP were used to compare all base pairs, surface sites, and transmembrane sites. M–K tests compare the ratio of fixed (invariable) synonymous (SI) to non-synonymous (RI) substitutions to the ratio of polymorphic (variable) synonymous (SV) to non-synonymous variable (RV) substitutions. The SI:RI ratio is taken as the expectation for neutral evolution of sequences. Only unique haplotypes were used in M–K tests.

Because M–K tests are pairwise, one cannot infer whether departures from neutrality were in one or both taxa being

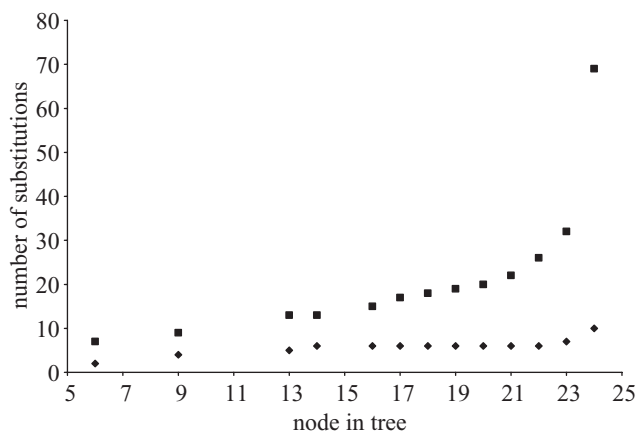


Figure 2. Numbers of synonymous (filled squares) and replacement (filled diamonds) substitutions at nodes on tree in figure 1 (values for nodes 34–37 not plotted). The number of synonymous changes increases with distance from the tip of the tree faster than replacement substitutions.

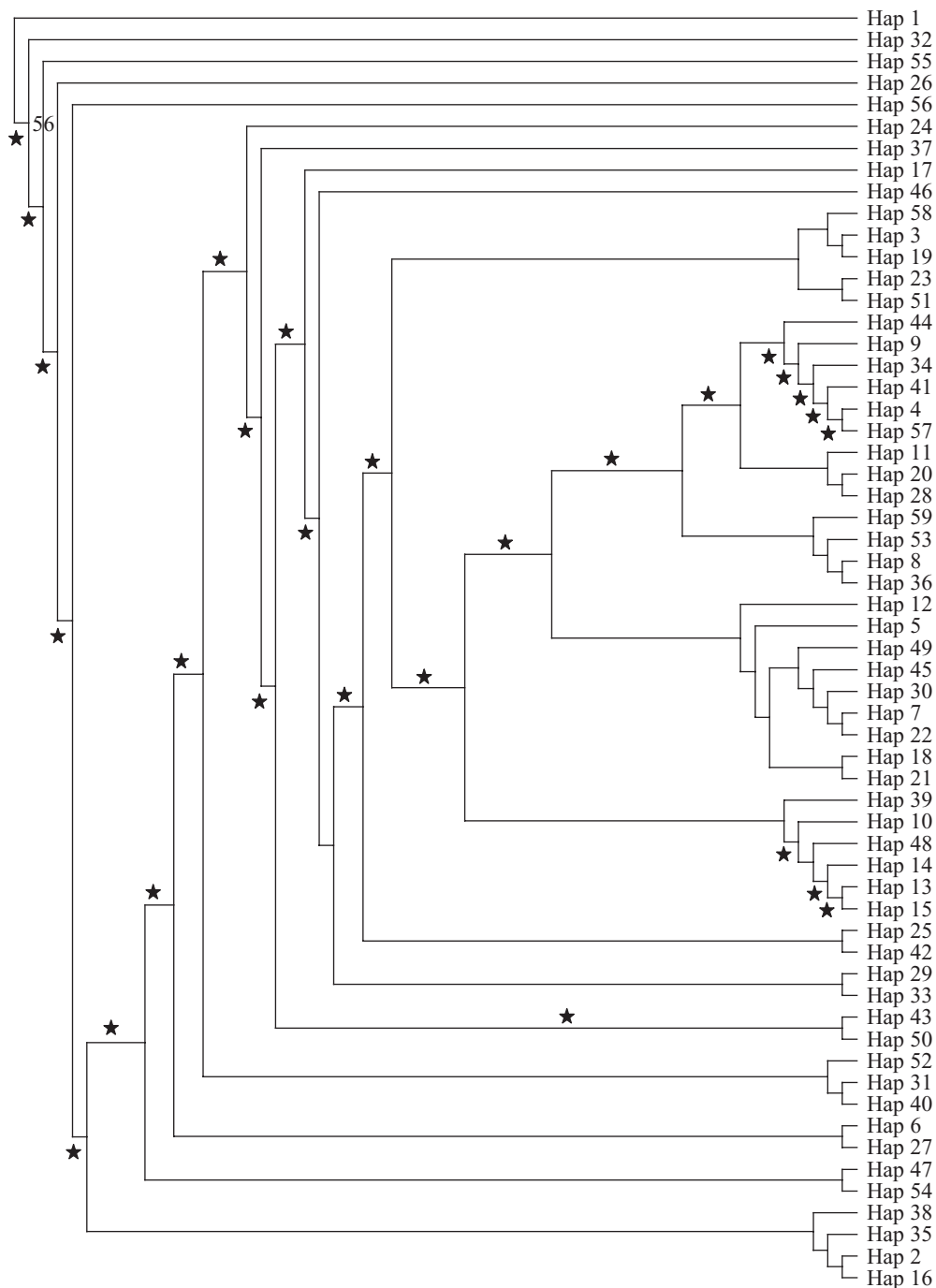


Figure 3. Neighbour-joining tree for *P. montanus* rooted with *P. palustris*. Stars refer to nodes with significant deficiencies of replacement substitutions. At node 56 there was a significant M–K test.

compared. To explore the phylogenetic deployment of putative selective effects, I used the relative ratio test of Creevey & McInerney (2002). In this method, a rooted phylogenetic tree is computed (and assumed correct). The software constructs a neighbour-joining tree using distances from synonymous sites, non-synonymous sites, or their ratio; results presented below were independent of the distance measure used for neighbour-joining trees. Hypothetical ancestral sequences are constructed for each node, and in the event of ambiguities, resolved using matrices of zero-, two- and fourfold degenerate positions. All substitutions are identified on the tree and classified as either non-synonymous or synonymous. For each internal branch of the tree, substitutions

are counted from the branch to the descendant tips, resulting in four values: RI, RV, SI and SV. A *G*-test or Fisher's exact test (Sokal & Rohlf 1981) is used to compare the ratio of SV:RV to SI:RI (the expected value under selective neutrality). If significant differences occur, one can infer whether the positive selection is directional (high RI) or non-directional (high RV). Computations are implemented in software CRANN written by C. Creevey (available at <http://bioinf.may.ie/crann/>).

Creevey & McInerney (2002) also describe a method to detect negative or purifying selection. Here, the ratio of replacement to silent sites is computed for each lineage in the tree, and the number of replacement and silent changes counted on each

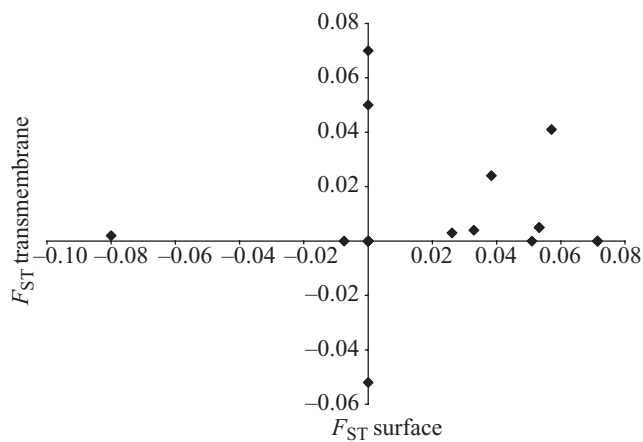


Figure 4. Plot of  $F_{ST}$ -values derived from transmembrane portion of ND2 versus surface portion for samples of *P. major*. Although there appears to be scatter, all individual  $F_{ST}$ -values are low and not significant (all values of  $\chi^2 > 0.50$ ).

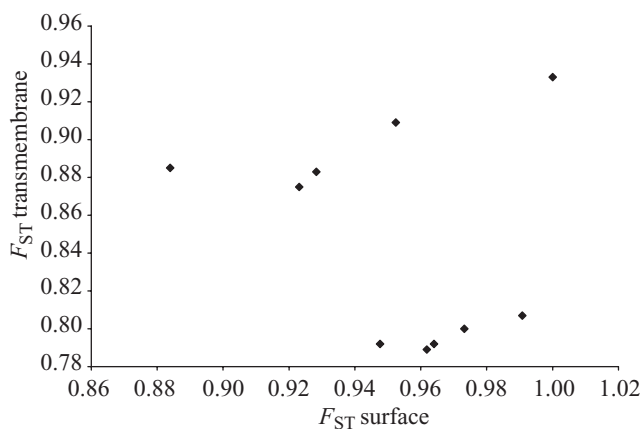


Figure 5. Plot of  $F_{ST}$ -values derived from transmembrane portion of ND2 versus surface portion for samples of *P. major*, *P. minor* and *P. bokharensis*.

branch. The expectation is that the ratio of replacement : silent changes on each branch will be predicted by the overall ratio of the two types of site on the entire tree. It is possible to observe both positive and negative selection on the same branch if overall there is strong purifying (negative) selection but a few highly favoured replacement substitutions.

Creevey & McInerney (2002) and others (Maynard Smith 1994; Hughes 1999) note that if species compared are too distant, saturation will obscure the pattern of replacements. In my analyses, I found that *P. montanus* and *P. major* differed by over 10% uncorrected sequence divergence. Hence, analyses were done independently on two major groups, *P. major*, *P. minor* and *P. bokharensis*, and *P. palustris* and *P. montanus* (which differ by ca. 7% uncorrected sequence divergence; R. M. Zink, personal observation).

I used a structural model of ND2 (Persson & Argos 1994) to divide the gene into surface (codons 1–3, 67–81, 142–183, 244–264, 325–345, 406–426, 487–561, 637–711, 796–816, 901–963, 1024–1038) and transmembrane (remaining codons) sections. To measure the pattern and extent of population subdivision, I used DnaSP to compute  $F_{ST}$  for all bases, surface and transmembrane sections; individuals were grouped by geographical locality. Because of the apparent pattern of selection in

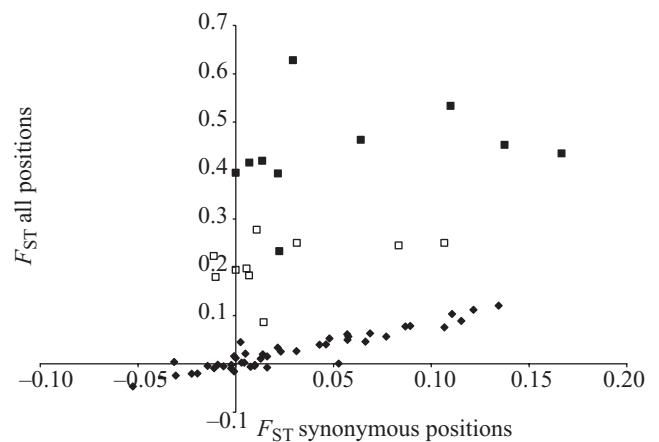


Figure 6. Plot of  $F_{ST}$ -values derived from silent sites versus all base positions in ND2 for samples of *P. montanus*. Filled squares represent comparison of the sample from Anadyr and open square comparison of samples to Kamchatka. The dash indicates comparison of Anadyr and Kamchatka. Filled diamonds represent all other comparisons.

*P. montanus*,  $F_{ST}$  was computed for synonymous sites only and a mismatch distribution (Rogers & Harpending 1992) was produced. The mismatch distribution was compared with the theoretical expectation for an expanding population implemented in DnaSP (Rozas *et al.* 2003).

### 3. RESULTS

Nucleotide diversity (table 1) was similar for *P. major* (0.0012) and *P. minor* (0.0012), whereas *P. montanus* was more variable (0.0021). For all three species, Fu's  $F$  was significant, as were Tajima's  $D$ , Fu and Li's  $D^*$ , and Fu and Li's  $F$ -values (table 1). Transmembrane sites exhibited higher nucleotide diversity than surface sites in *P. minor* but not in *P. major* or *P. montanus* (table 1). In M–K tests of *P. major* and *P. minor*, an excess of RV was observed for all sites, transmembrane and surface regions (table 2), although the test for surface sites was not significant. For example, comparing all base positions, there were no fixed replacement substitutions (out of 23 substitutions), whereas within species 10 out of 46 polymorphisms involved replacement substitutions. All M–K tests involving *P. bokharensis* were not significant (small sample size is a likely issue), although again an apparent excess of polymorphisms within species involved replacement substitutions. The M–K test for *P. palustris* versus *P. montanus* was not significant despite an apparent excess of RV substitutions (SI = 43, RI = 7, SV = 53, RV = 12; Fisher's exact test  $p$ -value 0.62). Values shown in table 2 for *P. montanus* versus *P. major*, *P. bokharensis* and *P. minor* are for comparison, but because of the likelihood of saturation (at least a 40% underestimate based on maximum-likelihood distance correction), they are not interpreted.

Relative ratio tests for branches in the haplotype tree of *P. major* and *P. bokharensis* (rooted with all *P. minor* haplotypes) were not significant, indicating no overall trend for directional (positive) selection. However, the pattern of substitutions (figure 1) suggested that non-synonymous substitutions were underrepresented in the basal parts of the haplotype tree. That is, as one proceeds from the tips to the base of the tree, silent substitutions increase but

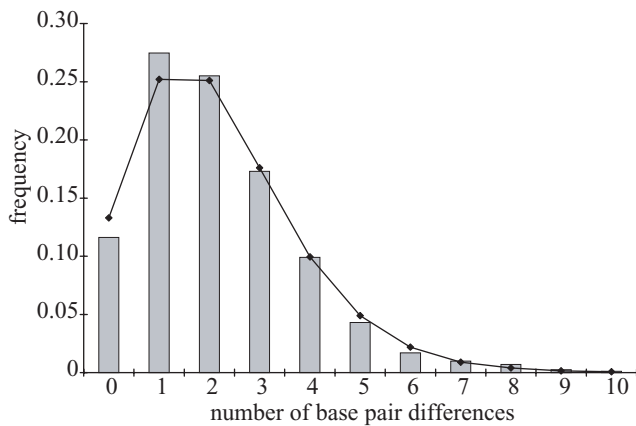


Figure 7. Mismatch distribution for 139 individual *P. montanus* using all of ND2 (observed: histogram; expected: line). Average number of pairwise differences: 2.18. The fit of observed and expected distributions (Kolmogorov–Smirnov test,  $\chi^2 = 0.72$ ,  $p > 0.9$ ) implies a recent and extensive increase in population size.

replacement substitutions increase at a much lower rate (figure 2). The same pattern was observed with surface and transmembrane sites (results not shown), irrespective of whether the neighbour-joining tree was constructed with distances calculated from synonymous substitutions non-synonymous substitutions or the ratio of the two.

The haplotype tree for *P. montanus* (rooted with *P. palustris*) showed one significant relative ratio difference due to an excess of RV substitutions, indicating positive non-directional selection (node 56, SI = 71, RI = 9, SV = 71, RV = 29). As in *P. major*, the haplotype tree (figure 3) showed a relatively large number of basal nodes in which replacement changes were underrepresented. Additionally, one terminal branch of the tree (connecting haplotypes 44, 9, 34, 41, 4 and 57) showed a preponderance of branches with underrepresented replacement substitutions. The number of synonymous substitutions increases from the tips of the tree to the base whereas replacement substitutions levels off (not shown), as in figure 2. Thus, non-synonymous substitutions are younger on average.

Multiple sampling localities are represented in the *P. major* and *P. montanus* datasets. The plot of  $F_{ST}$ -values obtained from surface and transmembrane sites among samples of *P. major* (figure 4) show that neither gene region provides qualitatively different estimates of geographical structuring. At a higher level of divergence, the same result was obtained for the comparison of *P. major*, *P. bokharensis* and *P. minor* (figure 5).

For *P. montanus*,  $F_{ST}$ -values based on the surface and transmembrane portions (not shown) differed, because in several instances, low values of the former were paired with high values of the latter. Inspection of the results showed that the high transmembrane  $F_{ST}$ -values involved comparisons of samples from Anadyr and Kamchatka with other localities. To explore this result further,  $F_{ST}$ -values based on all base pairs and synonymous sites only were compared (figure 6), with the same result—pairwise comparisons in which one of the sites was either Anadyr or Kamchatka were much higher for all base pairs. Lastly, mismatch distributions for *P. montanus* show close correspondence

between observed and expected values for all sites (figure 7) and synonymous sites (figure 8) only.

#### 4. DISCUSSION

The role of natural selection on DNA sequences in phylogenetics and phylogeography is unclear. Barton & Etheridge (2004, p. 1115) remarked that their analysis supported previous simulations (e.g. Neuhauser & Krone 1997) which showed that ‘...selection has surprisingly little effect on genealogies’. Whether natural selection also has little effect on phylogeographic studies is unknown, although Ballard & Whitlock (2004) suggest that selection on mtDNA could compromise evolutionary inferences. If true, hundreds of analyses could be biased. Divergence thought to be a result of isolation and genetic drift might instead reflect geographical patterns of selection. For *P. major*, *P. bokharensis* and *P. minor*, the M–K tests (table 2) suggest an excess of replacement substitutions segregating within species. However, these replacement substitutions tend to be at the tips of the haplotype tree rather than towards the base (figure 1). The significance of the relative basal decrease in replacement substitutions can be inflated by testing the tree at multiple nested levels. However, it is clear that replacement sites did not increase at the same rate as synonymous sites (figure 2), which themselves might be underestimated at deeper nodes (owing to saturation). Thus, the sequences appear to be evolving in a manner consistent with a model of slightly deleterious alleles. Consequently, the age of replacement substitutions is less than silent ones, the former of which are removed more quickly by selection. A similar phylogenetic distribution of replacement substitutions was noted by Nielsen & Weinreich (1999). The consequence of this relatively weak selection for phylogeographic inference appears minimal because  $F_{ST}$ -values from surface positions showed the same pattern as those from the transmembrane portion of the gene (figures 4 and 5). Thus, although replacement substitutions appear to be mildly deleterious, phylogeographic conclusions based on the sequence data appear unaffected for *P. major*. Furthermore, there were no fixed replacement substitutions between *P. major* and *P. minor*, and only one between *P. bokharensis* and *P. major*, suggesting that mtDNA divergence accompanying the early stages of lineage divergence and speciation is selectively neutral.

The situation in *P. montanus* is somewhat different. Here again there is a disproportionately large number of polymorphic replacement sites (table 2), and a strong trend for replacement substitutions to be relatively rare toward the base of the haplotype tree (figure 3). As with *P. major*, this is a signature of mildly deleterious alleles. However, when the pattern of  $F_{ST}$  is examined for surface versus transmembrane (not shown) and synonymous versus all sites (figure 6), there are clear discrepancies. Inspection of figure 6 revealed that the mismatched  $F_{ST}$  estimates occurred in comparisons of two northeastern sample localities, Anadyr and Kamchatka. The large values of  $F_{ST}$  occurred because of a replacement substitution at base position 452 (a transmembrane site) in a haplotype (#12 in figure 3) that is fixed in Anadyr (10/10) and common (4/5) in Kamchatka. There are no fixed synonymous differences for these localities. Hence, although it is possible by chance for a single fixed difference to be a replacement one, it is

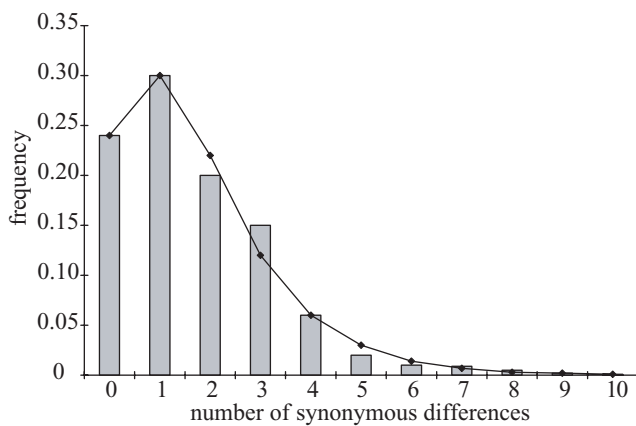


Figure 8. Mismatch distribution for 139 individual *P. montanus* using only synonymous substitutions in ND2 (observed: histogram; expected: line). Average number of pairwise differences: 1.70. The fit of observed and expected distributions (Kolmogorov–Smirnov test,  $\chi^2 = 0.18$ ,  $p > 0.9$ ) implies a recent and extensive increase in population size.

also possible that this haplotype is under positive selection, against a background of mildly deleterious alleles (Creavey & McInerney 2002). Thus, one might interpret the significant  $F_{ST}$ -values for Anadyr and Kamchatka as evidence of incipient divergence caused by geographical isolation, or selection for a new favourable haplotype. These two localities are geographically proximate, which could mean similar selective pressures are at work, or that gene flow occurs between them (or both). A definitive decision seems unlikely given that only a single substitution is involved. Nonetheless, considering the broader picture, only 2 out of 12 localities appear to be affected. Apart from being cautious about interpreting the evolutionary significance of the high  $F_{ST}$ -values for comparisons involving Anadyr and Kamchatka, overall phylogeographic inferences are not altered by the existence of mildly deleterious haplotypes.

Besides assessing population differentiation, mtDNA data are also used for assessing recent population history such as population growth. A typical procedure is to compute a mismatch distribution (Rogers & Harpending 1992) which, if unimodal, is consistent with recent large expansion in population (and likely range) size. A pattern of purifying selection could mimic population expansion by limiting mtDNA variation to a single common haplotype and a series of closely related (but neutral) ones, leading to a unimodal mismatch distribution. Analyses presented here suggest a deficit of replacement changes, which could clearly affect mismatch distributions. However, in *P. montanus*, there are only 12 non-synonymous changes, and the mismatch distributions based only on synonymous changes ( $n = 53$ ) do not differ substantially from that based on all base pairs (figures 7 and 8). Thus, it appears reasonable to interpret the unimodal mismatch distributions as evidence for recent population expansions, which makes sense for species living in environments that have expanded greatly after the last Ice Ages.

It seems premature to be overly critical of phylogeographic inferences derived from mtDNA out of concern for natural selection. There simply have not been enough studies of the effects of selection on phylogeographic

inferences based on mtDNA gene sequences. Although it would be useful to have estimates of population parameters from multiple loci, in general nuclear loci (including microsatellites) coalesce too slowly to provide tests of incipient patterns that are discernable in mtDNA gene trees. It will be critical to assess the extent of selection in other species to determine the generality of the results presented here.

The US National Science Foundation funded many of the studies summarized here. S. Rohwer, A. Pavlova and S. Drovetski contributed markedly to this work. I thank A. Pavlova for permission to analyse unpublished ND2 sequences. R. Blackwell-Rago and M. Westberg performed much of the laboratory work. B. Barber, C. Creevey, S. Drovetski, A. Jones, A. Kessen, A. Pavlova and two anonymous reviewers made comments on the manuscript.

## REFERENCES

- Avise, J. C. 2000 *Phylogeography*. Cambridge, MA: Harvard University Press.
- Avise, J. C. & Walker, D. 1998 Pleistocene phylogeographic effects on avian populations and the speciation process. *Proc. R. Soc. Lond. B* **265**, 457–463. (doi: 10.1098/rspb.1998.0317)
- Ballard, J. W. O. & Whitlock, M. C. 2004 The incomplete natural history of mitochondria. *Mol. Ecol.* **13**, 729–744.
- Barton, N. H. & Etheridge, A. M. 2004 The effect of selection on genealogies. *Genetics* **166**, 1115–1131.
- Creevey, C. & McInerney, J. O. 2002 An algorithm for detecting directional and non-directional positive selection, neutrality and negative selection in protein coding DNA sequences. *Gene* **300**, 43–51.
- Eyre-Walker, A. 2002 Changing effective population size and the McDonald–Kreitman test. *Genetics* **162**, 2017–2024.
- Fu, Y.-X. 1997 Statistical tests of neutrality against population growth, hitchhiking and background selection. *Genetics* **147**, 915–925.
- Fu, Y.-X. & Li, W.-H. 1993 Statistical tests of neutrality of mutations. *Genetics* **133**, 693–709.
- Jansa, S. A., Lundrigan, B. L. & Tucker, P. K. 2003 Tests for positive selection on immune and reproductive genes in closely related species of the murine genus *Mus*. *J. Mol. Evol.* **56**, 294–307.
- Hughes, A. L. 1999 *Adaptive evolution of genes and genomes*. Oxford University Press.
- Kvist, L., Martens, J., Ahola, A. & Orell, M. 2001 Phylogeography of a Palearctic sedentary passerine, the willow tit (*Parus montanus*). *J. Evol. Biol.* **14**, 930–941.
- Kvist, L., Martens, J., Higuchi, H., Nazarenko, A. A., Valchuk, O. P. & Orell, M. 2003 Evolution and genetic structure of the great tit (*Parus major*) complex. *Proc. R. Soc. Lond. B* **270**, 1447–1454. (doi: 10.1098/rspb.2002.2321)
- McDonald, J. H. & Kreitman, M. 1991 Adaptive protein evolution at the *Adh* locus in *Drosophila*. *Nature* **351**, 652–654.
- Maynard Smith, J. 1994 Estimating selection by comparing synonymous and substitutional changes. *J. Mol. Evol.* **39**, 123–128.
- Messier, W. & Stewart, C.-B. 1997 Episodic adaptive evolution of primate lysozymes. *Nature* **385**, 151–154.
- Neuhauser, C. & Krone, S. M. 1997 The genealogy of samples in models with selection. *Genetics* **145**, 519–534.
- Nielsen, R. & Weinreich, D. M. 1999 The age of nonsynonymous and synonymous mutations in animal mtDNA and implications for the mildly deleterious theory. *Genetics* **153**, 497–506.
- Pavlova, A. 2004 Comparative phylogeographies of Eurasian birds. Ph.D. dissertation, University of Minnesota.

- Persson, N. & Argos, P. 1994 Prediction of transmembrane segments in proteins utilising multiple sequence alignments. *J. Mol. Biol.* **237**, 182–192.
- Rogers, A. R. & Harpending, H. 1992 Population growth makes waves in the distribution of pairwise genetic differences. *Mol. Biol. Evol.* **9**, 552–569.
- Rozas, J., Sanchez-DelBarrio, J. C., Messeguer, X. & Rozas, R. 2003 DnaSP, DNA polymorphism analysis by the coalescent and other methods. *Bioinformatics* **19**, 2496–2497.
- Sokal, R. R. & Rohlf, F. J. 1981 *Biometry*. San Francisco, CA: Freeman.
- Tajima, F. 1989 Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585–595.
- Wise, C. A., Sraml, M. & Eastal, S. 1998 Departure from neutrality at the mitochondrial NADH dehydrogenase subunit 2 gene in humans, but not in chimpanzees. *Genetics* **148**, 409–421.
- Zink, R. M. 2004 The role of subspecies in obscuring avian biological diversity and obscuring conservation policy. *Proc. R. Soc. Lond. B* **271**, 561–564. (doi: 10.1098/rspb.2003.2617)
- Zink, R. M., Drovetski, S. V., Questiau, S., Fadeev, I. V., Nesterov, E. V. & Rohwer, S. 2003 Recent evolutionary history of the bluethroat across Eurasia. *Mol. Ecol.* **12**, 3069–3075.

As this paper exceeds the maximum length normally permitted, the authors have agreed to contribute to production costs.