Multilocus sequence analyses of the near threatened Semi-collared flycatcher (*Ficedula semitorquata*) and a comparison with three other *Ficedula* flycatcher species

Silje Hogner

Master of Science thesis
2008

Centre for Ecological and Evolutionary Synthesis
Department of Biology
University of Oslo, Norway

PREFACE

This study has been conducted at the Centre for Ecological and Evolutionary Synthesis (CEES) at the Biology Department, University of Oslo. I would not have been able to conduct this study without help from a lot of people. First, I would like to thank my supervisor, Glenn-Peter Sætre for giving me this exciting project. And for being helpful during the hole project. I would like to thank Stein Are Sæther for helping me in the field, with the analyses and for giving me helpful comments during the writing process. I would like to thank Thomas Borge for valuable help during my lab work, analyses as well as the writing process.

I would also like to thank all the people at the CEES lab for a good working environment and all the people at my study room, for making these two years fun as well as educational. Thanks to Harald for emotional support, and for being a good cook.

At last, but not least, a big thanks to my mom and dad, for emotionally as well as financial support.

Silje,
Blindern 2008

Male Semi-collared flycatcher. Photograph by Stein Are Sæther

CONTENTS

ABSTRACT

The Semi-collared flycatcher (*Ficedula semitorquata*) is a member of the black-and-white flycatcher species complex. It is poorly studied, rare, and is currently classified as "near threatened", in the IUCN red lists of threatened species. The Semi-collared flycatcher has a patchy distribution in Europe and part of the Middle East. In this study, I analyzed sequence variation at nuclear loci of Semi-collared flycatchers, and compared the patterns found with those of three other black-and-white flycatcher species, the Pied (*F. hypoleuca*), Collared (*F.albicollis*) and Atlas flycatchers (*F. speculigera*). Genetic variation was found to be relatively high, compared to the other three flycatcher species and there were no signs of inbreeding. All four flycatcher species had less variation at Z-linked loci compared to autosomal loci. A comparison showed that all species combinations had fewer shared polymorphisms and more fixed substitutions at Z-linked than at autosomal loci. Selective sweeps on the Z-chromosome during the evolutionary history of these species is likely to have contributed to this pattern. A mismatch distribution showed signs of a recent population expansion in all four species, and a phylogenetic reconstruction confirmed a relatively deep split and that each species is monophyletic. This study supports the classification of the Semi-collared flycatcher as a separate species. Even though no signs of a small population size are found here, it is important to keep monitoring this bird, since it has gone through several declines across Europe during the last decades. Much of the decline is probably due to habitat destruction, so it is important to retain old forest in the habitat of the species.

INTRODUCTION

When species diverge from each other, they are expected to gradually loose shared polymorphisms and accumulate fixed substitutions, due to random genetic drift within each species (WAKELEY and HEY 1997). Divergence rates may differ between different parts of the genome. For example, loci under selection, or that are linked to regions under selection, will show different patterns of variation compared to those evolving neutrally (FAY and WU 2000). Not only selection, but also demographic events can cause deviations from patterns expected under neutrality (KREITMAN 2000). For example, population bottlenecks can lead to a reduction of genetic variation, loss of alleles (especially rare ones) and random changes in allele frequencies (FRANKHAM *et al.* 2004). Genetic markers can thus be useful in conservation biology, when it comes to determining past demography, such as bottlenecks and loss of genetic diversity.

Wild species can be driven to extinction by deterministic factors (such as habitat loss, overexploitation, pollution and introduced species) and stochastic factors associated with a small population size (e.g. BAILLIE *et al.* 2004; FRANKHAM 1999; FRANKHAM *et al.* 2004). The stochastic factors may have environmental, catastrophic, demographic or genetic origins (FRANKHAM *et al.* 2004). For birds, relatively good data exist indicating that impacts of invasive species, overexploitation and habitat destruction and degradation have been the major causes of extinction (BAILLIE *et al.* 2004). It is typical for threatened species that they have a small or declining population (SPIELMAN *et al.* 2004), and some species have been reduced to a number where they require human intervention to save them from extinction (e.g. FRANKHAM 1999). In small populations, both loss of genetic variation (as a consequence of genetic drift) and inbreeding will become unavoidable (REED and FRANKHAM 2003; SPIELMAN *et al.* 2004). Inbreeding can increase the extinction risk by several different factors. These include exposure of rare, deleterious alleles and reduction of heterozygosity (e.g. FRANKHAM *et al.* 2004; PACKER 1979). The primary cause of inbreeding depression is thought to be that homozygotes for deleterious recessive alleles become more frequent than in a random mating population (PACKER 1979). Packer listed three reasons why homozygosity might lead to an inbreeding depression. The first is that more recessive alleles would be expressed in inbred individuals. The probability of inheriting two recessive alleles is very low, except when both parents carry the allele by common decent. The second reason is that the heterozygote individuals might be more fit than homozygote individuals. In inbred populations, heterozygote advantage (heterosis) in the parents will be reduced in the offspring

since they will have a higher chance of being homozygote than outbred offspring. Third, there will be less variation between the offspring. This might lead to a reduction in survival chance for the offspring, in an event of sudden environmental change (LANDE 1988).

Due to the effects of genetic drift, the relative effect of selection is weaker in small populations compared to larger populations, leading to an increase in the proportion of deleterious alleles and their possible fixation in the small populations (FRANKHAM *et al.* 2004; REED 2005). In small populations, genetic and stochastic factors operate in a negative feedback, and might result in an extinction vortex where loss of fitness resulting from the fixation of deleterious alleles suppresses population size. This might again lead to an increased amount of drift and eventually end in extinction (GILPIN and SOULÉ 1986). REED and FRANKHAM (2003) found that fitness and future adaptability are reduced in smaller populations of plants and animals, due to drift and inbreeding depression. Since genetic variation is related to population size, endangered species should have significantly less genetic variation than non-endangered species (e.g. FRANKHAM 1996).

This study focuses mostly on the Semi-collared flycatcher (*Ficedula semitorquata*), which is a small passerine bird. The Semi-collared flycatcher is closely related to the Pied (*F. hypoleuca*) and Collared flycatchers (*F. albicollis*), which are well studied species (see e.g. BORGE *et al.* 2005; LUNDBERG and ALATALO 1992; SÆTHER *et al.* 2007; SÆTRE *et al.* 2001; VEEN *et al.* 2001). The Semi-collared flycatcher experienced a large population decline during 1970-1990, and continued to decline across most of southeast Europe also during 1990-2000 (BIRDLIFE 2008). The rapid decline in population size is most likely due to habitat destruction in some areas (BIRDLIFE 2008; TUCKER and HEATH 1994). For example, the decline in Turkey is thought mainly to be caused by habitat loss due to ongoing dam projects and rapid loss of Oak (*Quercus* spp.) forest. Further information on population trends outside of its European breeding range is needed (BIRDLIFE 2008). The Semi-collared flycatcher is currently listed as "near threatened" by the World Conservation Union (IUCN), and it almost qualifies as "threatened" under certain criteria (A2bc+3bc).

Here I present results of analyses of genetic variation in the Semi-collared flycatcher to look for possible signals of a small population size as well as signals of past demography. Most studies in conservation genetics are done with mtDNA or microsatellites. However, I chose to use nuclear sequences since I had access to such sequence data from closely related flycatcher

species that can be used for comparison. Moreover, recent advances in population genetics has made it possible to extract more information from multilocus sequence analysis than was previously the case (see e.g. SCHLÖTTERER 2003; WAKELEY and HEY 1997). The three other species of the black-and-white flycatcher species complex, the Pied, Collared and Atlas (*F. speculigera)* flycatchers, generally have larger population sizes and less fragmented populations than the Semi-collared flycatcher (e.g. LUNDBERG and ALATALO 1992). In a previous study of Pied and Collared flycatchers, BORGE *et al.* (2005) found that the genetic divergence from the outgroup, the Redbreasted flycatcher (*Ficedula parva*) was higher on the Z-chromosome, compared to the autosomes. This pattern was also found in the Atlas flycatcher by BRUVIK (2007). An explanation for this might be that the mutation rate is higher in males, due to the many more rounds of cell-divisions that the male germlines goes through, compared to the female germline. Hence, since females only have one Z-chromosome but both sexes have two copies of the other chromosomes, a male-biased mutation rate may cause Z-linked genes to evolve faster than autosomal ones. BORGE *et al.* (2005) and BRUVIK (2007) also found that the levels of polymorphism on the Z-chromosome were lower than that of the autosomes. BORGE *et al.* (2005) suggested that one explanation for this could be selective sweeps on the Z-chromosome, and that such selection events could have played an important role in the speciation process. The reason for this is that genes associated with speciation seems to be over-represented at the Z-chromosome (e.g. SÆTHER *et al.* 2007).

In this study, I will compare sequences from autosomal and Z-linked genes from the Semi-collared flycatcher to the same sequences from the Pied, Collared and Atlas flycatcher. First, I analyse if variation is reduced at Z-linked loci compared to autosomes as found in the Pied, Collared and Atlas flycatchers (BORGE *et al.* 2005; BRUVIK 2007). Second, I analyse if the genetic divergence from the outgroup is higher on the Z-chromosome than on the autosomes in the Semi-collared flycatcher, as also found in the other flycatcher species. Third, I compare the variation in the Semi-collared flycatcher with the variation in the other flycatcher species, to see if it shows any signs of a relatively smaller population size. Fourth, I will conduct various analyses to reveal past demographic patterns, based on allele frequency distributions. Finally, I will do a new phylogenetic reconstruction of the species complex, with more sequences and individuals, to see if the Semi-collared flycatcher is indeed a separate species and fits with the phylogeny previously suggested by SÆTRE *et al.* (2001), where a smaller population size (two individuals) and mtDNA were used.

MATERIALS AND METHODS

Study species

The taxonomy of the black-and-white *Ficedula* flycatcher species complex has been debated (LUNDBERG and ALATALO 1992). It was previously thought that the Atlas flycatcher was a subspecies of the Pied flycatcher and that the Semi-collared flycatcher was a subspecies of the Collared flycatcher. SÆTRE *et al.* (2001) found, using mtDNA, that they were in fact four highly divergent groups. However, these results were based on a small sample size for the Semi-collared flycatcher. It has been suggested that the ancestral flycatcher species got separated into four different refugia during the Pleistocene glaciations (SÆTRE *et al.* 2001), and that their distributions were limited by the presence of suitable forest habitats around the Mediterranean Sea. Today, the Pied flycatcher has its most southern distribution in Iberia and the Collared flycatcher in Italy. The Semi-collared flycatcher have its only distribution in the Balkans/Caucasus and Middle East, while the Atlas flycatcher only occurs in the Atlas Mountains in northern Africa (Figure 1, CRAMP and PERRINS 1993).

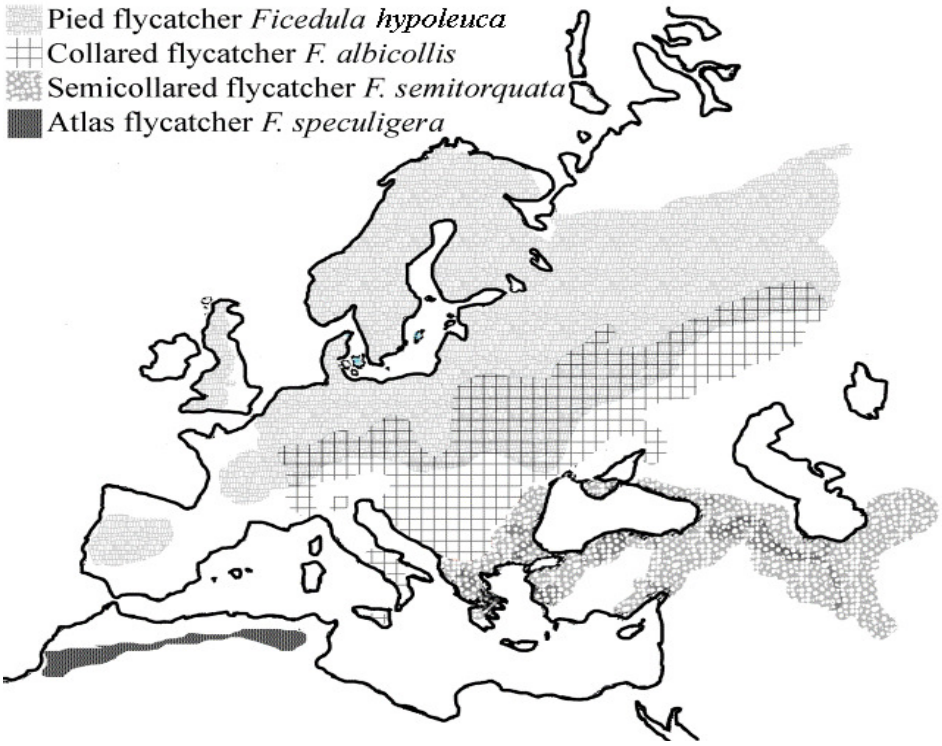


Figure 1. The geographical distribution of the four black-and-white *Ficedula* flycatchers in Europe. The distribution of the Semi-collared flycatcher is fragmented and uncertain, especially in the eastern parts of the range (From BORGE *et al.* 2005).

The Semi-collared flycatcher breeds in southeast Europe (Southern parts of the Balkans) and in the Near and Middle East (Asia Minor, Caucasus and Iran), where it favours forests, mainly on mountain slopes (up to about 2000m altitude), occupied by mature deciduous trees (notably oak *Quercus* and hornbeam *Carpinus),* as well as deciduous riverine forest on plains and old groves and gardens (IUCN 2007; TUCKER and HEATH 1994). The total number of breeding pairs is uncertain, but it is believed that the population size lies between 15 000 and 53 000 pairs in Europe, which is considered to be 50-74% of the global population (IUCN 2007). It winters in relatively small regions of East Africa, ranging from southern Sudan through western Uganda to Tanzania. There, it favours evergreen forests, gallery forest and edges of forest blocks (BURFIELD *et al.* 2004; IUCN 2007).

Semi-collared flychatcers are, like the other *Ficedula* species, solitary and territorial, and they defend a small territory around their nest (CURIO 1959). They breed in natural holes in trees, but can also be attracted to nest boxes (CRAMP and PERRINS 1993; TUCKER and HEATH 1994). Their territory serves for courtship, nesting, some feeding, and it includes a number of favoured song-posts. The Semi-collared flycatcher competes for nest-holes with other passerines, such as Redstart (*Phoenicurus phoenicurus*), Titmice (*Parus sp*), Treecreepers (*Certhia sp.*) and Nuthatch (Sitta europea, CURIO 1959).

Field work

The field work was conducted during May 2006, in the Kamcheya region in Bulgaria (43°00' N 27°50' E). The breeding ground consists mostly of old deciduous forest. I caught a total of twelve Semi-collared flycatchers; nine males and three females. In addition, three male samples from the same area were available from a study the previous year. All birds were caught using mist nets, song playback, and a male in a cage to lure the male birds down from their nest holes. The nets were placed directly underneath the nest holes. Previous studies on the closely related Pied and Collared flycatcher have shown that these birds are good model species since they are relatively tame, easy to handle in the field and since catching and handling them do not seem to affect their breeding performance or survival in any negative way (LUNDBERG and ALATALO 1992). Hence, even though the Semi-collared flycatcher is a vulnerable species I had reasons to assume that catching and handling the birds would not affect them negatively.

Phenotypic measures were taken from all birds. These included wing length, tarsus length, forehead patch height and breadth, and it was noted at which primary feather the wing patch started and whether the forehead patch consisted of one big patch or two small ones. All phenotypic measures were done using a digital caliper with an accuracy of 0.1mm, except for the wing length which was measured using a ruler with an accuracy of 1mm. However, analyses of the phenotypic measures are not presented in this thesis. A blood sample was taken from all individuals, by puncturing the brachial vein. The blood was stored in Queen's Lysis Buffer (SEUTIN *et al.* 1991). In addition, a photo was taken of all the birds. Appropriate authorization to catch birds and take blood samples was obtained from the ministry of environment and waters, Regional Inspectorate of Environment and Waters – Varna, Bulgaria.

Lab work

DNA extraction

DNA extractions were done using two different kits, QIAamp DNA Blood Mini Kit (Sample & Assay Technologies) and eZNA Blood DNA Kit (Omega Bio-Tek, Inc.). I followed the protocol for each of the two kits when extracting the DNA.

PCR

Several primer pairs that had been tested on *Ficedula* flycatchers were available from previous studies (BORGE *et al.* 2005; BRUVIK 2007). These primers were designed as described by PRIMMER *et al.* (2002), using chicken (*Gallus gallus*) sequences available from GenBank as templates. The primers were designed in exon sequences flanking introns of sizes appropriate for direct sequencing of both autosomal and Z chromosome genes. New flycatcher-specific primers were designed for long introns when the amplification successes for these were variable (BORGE *et al.* 2005). In this study, I used primers that had high amplification success in earlier studies on the Pied, Collared and Atlas flycatcher (BORGE *et al.* 2005; BRUVIK 2007). I chose the ones with readable sequences for further analyses.

The introns were amplified in PCR reaction volumes of 10 µl, consisting of 8 µl master mix and 2 µl DNA. The master mix contained 4.85 µl dH_2O, 1 µl buffer II (Applied Biosystems), 0.4 µl magnesium, 1 µl dNTP, 0.3 µl forward and reverse primer, 0.05 µl of bovine serum albumin (BSA) and 0.1 µl AmpliTaq DNA polymerase (Applied Biosystems). The

amplifications were run on a DNA Engine Tetrad 2 (MJ Research, Waterton, MA, USA). The following profile was used: 95˚C for 1 min, 94˚C for 30 sec, 55-60˚C for 30 sec, 72˚C for 1 min, then the second and third step another 34-39 cycles before the last step, 72˚C for 10 min. The program was then set to run at 10˚C forever in order to maintain the products.

Sequencing

I cleaned the PCR-products for excessive primers and nucleotides by using Exo-Sap-IT (United States Biochemical Cleveland) and a program run at 37˚C for 45 min, then 80˚C for 15 min. The PCR products were then sequenced using BigDye Terminator sequencing buffer and v 3.1 Cycle Sequencing kit (Applied Biosystems).

Analyses

Sequences from thirteen individuals of the Semi-collared flycatcher, eleven males and two females, were compared with similar data from nine male Pied flycatchers (from Spain), nine male Collared flycatchers (from Italy), nine male Atlas flycatchers (from Morocco) and one male Red-breasted flycatcher (*F. parva*). For those genes where more than one intron was sequenced (Table 1), I combined the sequences from the same genes and analyzed them as one locus. Before any analyses started, I transformed all the diploid sequences into "pseudo-haplotypes" (female Z-linked genes were kept as haploid). This was done by assigning the two alleles at each polymorphic site randomly to one of the two sequences constructed for each individual. They were then aligned using ClustalW, in the program Mega v 3.1(KUMAR *et al.* 2004) before they were adjusted manually. To get reliable alignments, nucleotides of both strand was used in all instances, except for a few strands with low quality. These strands were sequenced multiple times before alignment, to get a reliable result.

Polymorphism and divergence

I used the program DnaSP 4.0 (ROZAS and ROZAS 1999) to analyze polymorphism and divergence. In birds, females are the heterogametic sex, with a Z and a W sex chromosome. Using a standard neutral model, assuming constant population size, random mating and no migration (WATTERSON 1975) it would therefore be expected that a higher mutation rate on the Z-chromosome would increase the variation relative to the autosomes, whereas the mode of inheritance of Z-chromosome would decrease the variation by a factor of ¾, since females have only one Z chromosome (BORGE *et al.* 2005). The expected variation at each Z-linked

locus in a species was calculated based on effective population size estimated from average levels of polymorphism and divergence from the outgroup at autosomal loci. The mutation rate was estimated from levels of divergence from the outgroup at each individual Z-linked locus (BORGE *et al.* 2005). The formula used is:

$$S_Z(\exp) = \frac{3D_Z S_A}{4D_A} \qquad (1)$$

Where S_Z is the expected number of segregating sites at a specific Z-linked locus, D_Z is the divergence at Z-linked loci between the focal species and the outgroup, S_A is the total number of segregating sites at autosomal loci, and D_A is the total divergence between the focal species and the outgroup, at autosomal loci.

The formula used for calculating the expected number of segregating sites at autosomal loci from average polymorphism and divergence on Z-linked loci is

$$S_A(\exp) = \frac{4D_A S_Z}{3D_Z} \qquad (2)$$

I calculated two common measures of nucleotide polymorphism, π, the average number of nucleotide differences per site between two sequences (NEI 1987) and θ_W, calculated from the number of variable positions in the aligned sample of sequences. $\theta = 4N_e\mu$ for autosomal loci and $3N_e\mu$ for Z-linked loci, N_e is the effective population size and μ is the neutral mutation rate (NEI 1987). π and θ_W were estimated with their standard deviations, using the program DnaSP.

Demographic changes

Allele frequency distributions can be used to investigate the demographic history of populations, and identify bottlenecks and recent population expansion (KREITMAN 2000). However, deviations from expectations under neutrality can be caused both by selection and changes in population size. Even if I analysed (mostly) non-coding intron sequences in this study, they might be linked to genes under selection.

A multilocus statistical test for natural selection was performed using the Hudson-Kreitman-Aguade (HKA) test (http://lifesci.rutgers.edu/~heylab). This test was performed to test if the levels of within- and between-population DNA variation are positively correlated, as predicted by the neutral mutation hypothesis (HUDSON *et al.* 1987). Deviations from the expected relationship would indicate that selection have altered the genetic variation. If the test does not show any sign of selection, tests such as Tajima's *D* and a mismatch distribution (see below) can be used to investigate demographic changes.

I used DnaSP to compute Tajima's *D* (TAJIMA 1989) and Fu and Li's *D* and *F* (FU and LI 1993). These neutrality tests are based on the allele frequency spectrum. They can be used to infer previous evolutionary and demographic events that the populations have experienced. Negative values of Tajima's *D* reflect an excess of rare polymorphisms in the populations, while positive values indicates an excess of intermediate-frequency alleles. In a population, an excess of rare alleles might result from positive selection or an increase in population size, whereas an excess of intermediate-frequency alleles might result from balancing selection or a population bottleneck (AKEY *et al.* 2004).

To test for historical demographic events within the different flycatcher species, I compared the observed frequency distributions of pairwise nucleotide differences among individuals, with the expected distribution from a population expansion (mismatch distribution), using the Arlequin ver 3.1 program (EXCOFFIER *et al.* 2005). If a population has experienced a demographic equilibrium or been in decline, then a multimodal distribution should be displayed, whereas a unimodal distribution should be displayed if a population has experienced a sudden demographic expansion (ROGERS and HARPENDING 1992; SLATKIN and HUDSON 1991). However, recent changes are not always detectable in a mismatch distribution, because they might be masked by threshold effects, time lags or earlier demographic events (LAVERY *et al.* 1996; ROGERS and HARPENDING 1992).

A mismatch distribution is based on haplotype data. I estimated the haplotype frequencies required for mismatch distribution analysis using the Bayesian (ELB algorithm) method implemented in Arlequin ver 3.1. I chose the ELB algorithm since it is the fastest algorithm for multi-locus genotypic data when the gametic phase is unknown. This algorithm attempts to reconstruct the unknown gametic phase of multi-locus genotypes, by randomly adding loci until the best fit is found. Phase updates are made on the basis of a window of neighbouring

loci. The ELB algorithm starts with a phase where all individuals in the sample are associated with a window. The window contains each heterozygous locus and the neighbouring loci (EXCOFFIER *et al.* 2003). Before the estimation started, I removed the two female individuals, since they can never be heterozygous for Z-linked loci. I used the best haplotype estimates to calculate the observed and expected distributions of pairwise differences. However, the mismatch distribution results should be interpreted with care, since the haplotype data are not very accurate.

Since the gametic phase was unknown I could not calculate reliable F_{IS}-statistics using the sequence data. In order to calculate F_{IS}-values, I therefore applied another method. I chose the most variable site at each locus (i.e. those most likely to reveal possible deviations from Hardy-Weinberg equilibrium), for each of the different species. The F_{IS}-values were then calculated using the program Genepop on the web (http://genepop.curtin.edu.au/genepop_op1.html) and the allele identity option. To check if the F_{IS}-values were significantly different from zero, I performed a global Hardy-Weinberg test (across all loci), using the same program. For the populations with a positive F_{IS}-value, I tested for heterozygote deficiency, and for those with a negative F_{IS}-value I tested for heterozygote excess.

Population divergence

Analysis of molecular variance (AMOVA) was used to examine the genetic structure of the populations. The variance was partitioned into three different parts: between the different species, between individuals within the species, and within the different individuals. Pairwise species differentiation was calculated based on F_{ST} – values (WEIR and COCKERHAM 1984), using the default settings in the population comparisons. These F_{ST} -values can be used as short-term genetic distances between populations (REYNOLDS *et al.* 1983; SLATKIN 1995). The null distribution of pairwise F_{ST}-values is obtained by permutating haplotypes between populations, under the hypothesis of no differences between populations. The *P*-value of the test is given as the proportions of simulations giving a F_{ST}-value larger or equal to the observed one.

A speciation isolation model was conducted using the program WH (http://lifesci.rutgers.edu/~heylab) as described by WAKELEY and HEY (1997) and WANG, WAKELEY and HEY (1997). This test allows an estimate of the time since speciation and the

size of the ancestral population. Under a simple speciation isolation model (WAKELEY and HEY 1997) as two populations diverge, their shared polymorphisms are lost and, due to random drift, gradually becomes fixed differences. The isolation model assumes that an ancestral population has split into two separate populations some time in the past. These two populations have then evolved independently according to the assumptions of a standard neutral model. All variation in the populations is assumed to be neutral and the mutations occur according to the infinite site model, with a mutation rate, μ, per sequence per generation. The isolation speciation model is described by three separate estimates of the population mutation parameter, θ; θ_A for the ancestral population, and θ_1 and θ_2 for each of the descendant populations, and one parameter for the time T since speciation event. The parameter T is measured in units of $2N_1$ generations, where N_1 is the constant effective population size of species 1 (WAKELEY and HEY 1997; WANG *et al.* 1997).

I constructed Neighbour-joining trees using the Mega v 3.1 program, with 1000 bootstrap replicates, and the nucleotide Kimura 2-parameter model. These trees were made to see if nuclear genes would give the same topography as a previous study using mtDNA and a more restricted amount (3635 bp) of nuclear DNA sequences (SÆTRE *et al.* 2001).

RESULTS

Sequence data

Sequences from the Atlas, Pied and Collared flycatchers were available from earlier studies (BORGE *et al.* 2005; BRUVIK 2007), and I compared these to thirteen sequences from the Semi-collared flycatcher. Pseudo-haplotypes were constructed for all the individuals. In this study, I used a total of fifteen loci (Table 1), nine from Z-linked genes (3973bp in total length) and six from autosomal genes (2974bp in total length). All sequences for each locus were adjusted to the same length as the shortest sequences of that locus for comparisons. This resulted in loss of some sequence data (56bp on Z-linked loci and 270bp on autosomal loci). The sequences I used consist mostly of introns, though some sequences have short flanking exon fragments at the ends (Table 1).

Table 1. Sample sequences. Cytological position refers to the Z chromosome (Z) or autosomes (A).

			Length (bp)		
Gene	Sequence	Cytological position	Total	Exon	Intron
ALDOB-6	ALDOB-6	Z	465	28	437
BRM-12	BRM-12	Z	1569	0	1569
CHDZ	CHDZ-15	Z	494	33	461
	CHDZ-18	Z	208	0	208
GHR	GHR-5.1	Z	106	0	106
	GHR-5.2	Z	449	0	449
VLDLR	VLDLR-8	Z	127	59	68
	VLDLR-9	Z	409	33	376
	VLDLR-12	Z	195	71	124
ACLY-16	ACLY-16	A	358	3	355
ALAS1-8	ALAS1-8	A	297	0	297
FAS-Y	FAS-Y	A	811	0	811
RHO-1	RHO-1	A	376	0	376
RPL30-3	RPL30-3	A	996	0	996
TGFB2-5	TGFB2-5	A	402	0	402

Polymorphism and divergence

I compared the average pairwise difference between autosomal and Z-linked loci for the different species (Figure 2). The only species with a statistical difference between the two classes of loci was the Semi-collared flycatcher (t = -1.812, df = 9 and P = 0.05). Here, I found that the average pairwise difference was higher for the autosmal loci than for the Z-

linked loci. The other species showed a similar pattern, although not significant (*t*-values; -0.929, -1.128 and -1.101 for the Atlas, Collared and Pied flycatchers respectively, and corresponding *P*-values > 0.15).

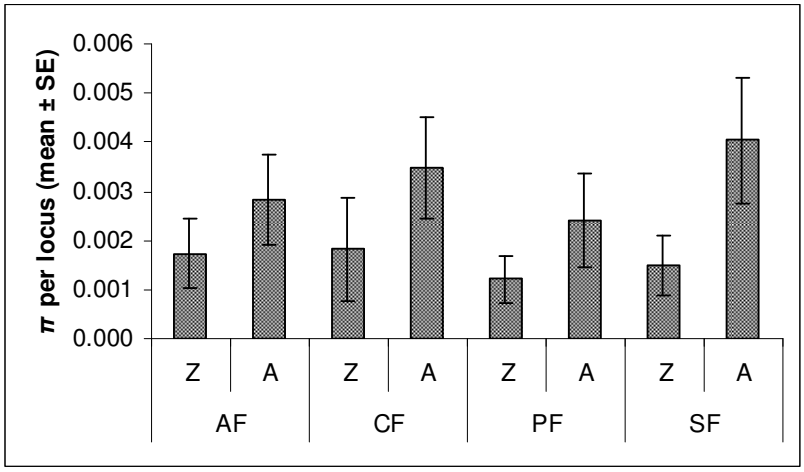


Figure 2. Average pairwise difference between autosomal (A) and Z-linked loci (Z) for the different species. AF = Atlas flycatcher, CF = Collared flycatcher, PF = Pied flycatcher, SF = Semi-collared flycatcher.

Comparing the Z-linked loci with the autosomal loci for the different species, I found that the frequency of polymorphic sites was higher for the Z-linked loci than for the autosomal for Collared, Pied and Semi-collared flycatchers, while the opposite was true for the Atlas flycatcher (Figure 3). However, the standard errors were high and overlapping between Z-linked and autosomal loci for all species, so none of the results were significant. When I compared all Z-linked loci and all autosomal loci for all species combined, a two-sample t-test was not statistically significant either ($t = 1.413$, df = 42, $P = 0.165$).

The number of segregating sites was slightly higher than expected from equation 1 and 2 (see materials and methods) for all the autosomal loci in all species (Appendix 1). Two Z-linked loci (ALDOB-6 and VLDLR) had a higher number of segregating sites than what would be expected, while the rest of the Z-linked loci generally had a slightly lower number of segregating sites than what would be expected (see Appendix 1).

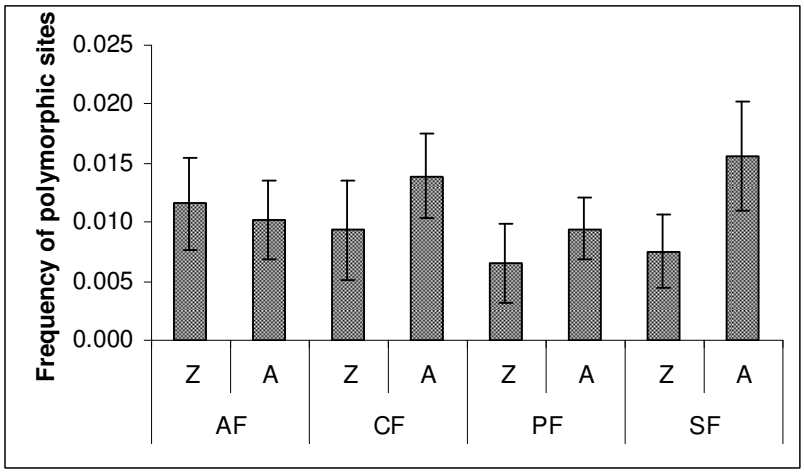
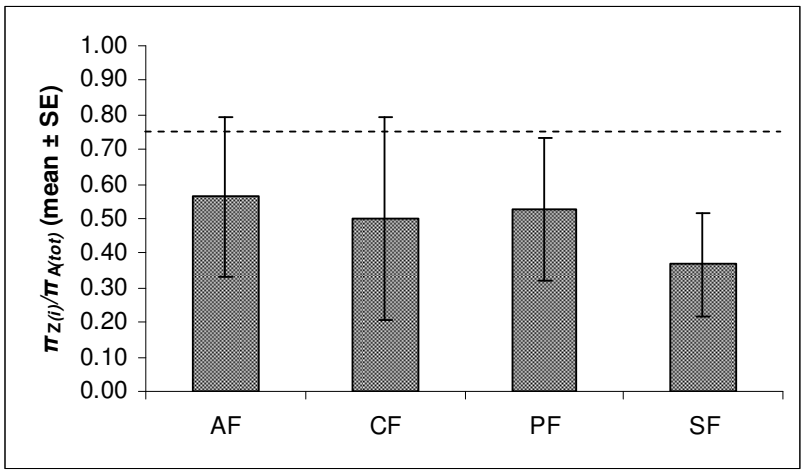


Figure 3. Mean frequency (± standard error) of polymorphic sites for Z-linked (Z) and autosomal (A) loci in four species of flycatcher. AF = Atlas flycatcher, CF = Collared flycatcher, PF = Pied flycatcher, SF = Semi-collared flycatcher.

I calculated the Z:autosomal ratio (Z_π:A_π) of average pairwise sequence difference. This was done by dividing the average pairwise sequence difference per nucleotide for each Z-linked locus by the average pairwise sequence difference per nucleotide for all autosomal loci together. I then calculated the mean of these ratios (Figure 4a). The ratio of expected heterozygosity (Z_θ:A_θ) between Z-linked and autosomal loci was calculated in the same way (Figure 4b). Both of these ratios should ideally be 0.75, if we assume random mating, constant population size and no migration. This is expected because the Z-linked loci have a N_e ¾ of the autosomal loci. For each species, these ratios were found on average to be below the expected value, but none of these results are significant except for the expected heterozygosity ratio for Pied flycatcher (ratio = 0.42, t = -2.317, df = 4, $P < 0.05$). Since I only have five Z-linked loci, these tests have low power. However, when I analyzed all species together, I found that the Z:autosomal ratio was significantly lower than the expected value 0.75 (based on π; t = -2.48, df = 19, P = 0.023 and based on θ; t = -2.172, df = 19, P = 0.043).

a)



b)

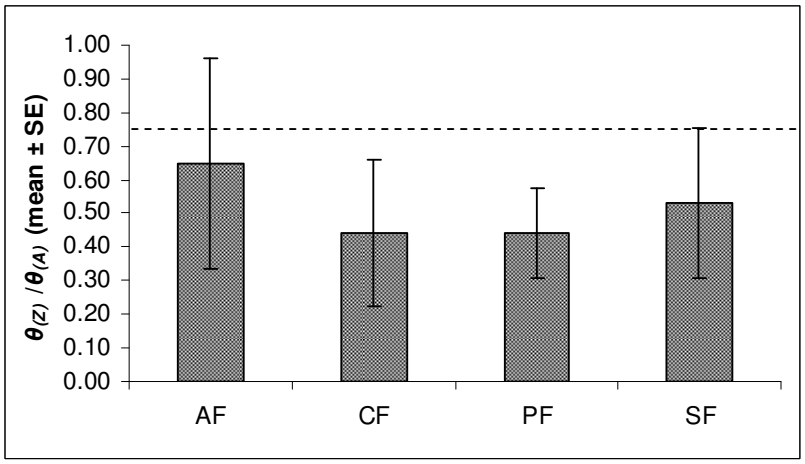


Figure 4. Comparison of Z to autosomal variation. a) The average pairwise sequence difference per nucleotide for each Z-linked locus ($\pi_{Z(i)}$) divided by the average pairwise sequence difference per nucleotide for all autosomal loci ($\pi_{A(tot)}$). b) The expected heterozygosity for each Z-linked locus ($\theta_{(Z)}$) divided by the expected heterozygosity for all autosomal loci ($\theta_{(A)}$). The dashed lines indicate the expected value 0.75. AF = Atlas flycatcher, CF = Collared flycatcher, PF = Pied flycatcher and SF = Semi-collared flycatcher.

Fixed and shared polymorphisms between species

Polymorphisms were divided into four categories in each of the four different species combination pairs: the variable sites exclusive to each of the two species; the shared polymorphisms between them; the fixed differences between the species; and the average pairwise differences between them (Appendix 2).

For all the species combinations, the average number of shared polymorphisms was found to be clearly higher on the autosomal than the Z-linked loci, with generally non-overlapping standard errors (Figure 5). In contrast, there was a higher level of fixed differences on the Z-linked than on the autosomal loci (except for the Pied and Atlas flycatchers compared), but due to high standard errors, these results are not statistically significant (Figure 6).

When I combined all species, I found that there was a significant difference between autosomal and Z-linked loci; there were more fixed ($t = 3.45$, df = 5, $P < 0.01$) and fewer shared ($t = 3.46$, df = 5, $P < 0.01$) differences on the Z-linked than on the autosomal loci.

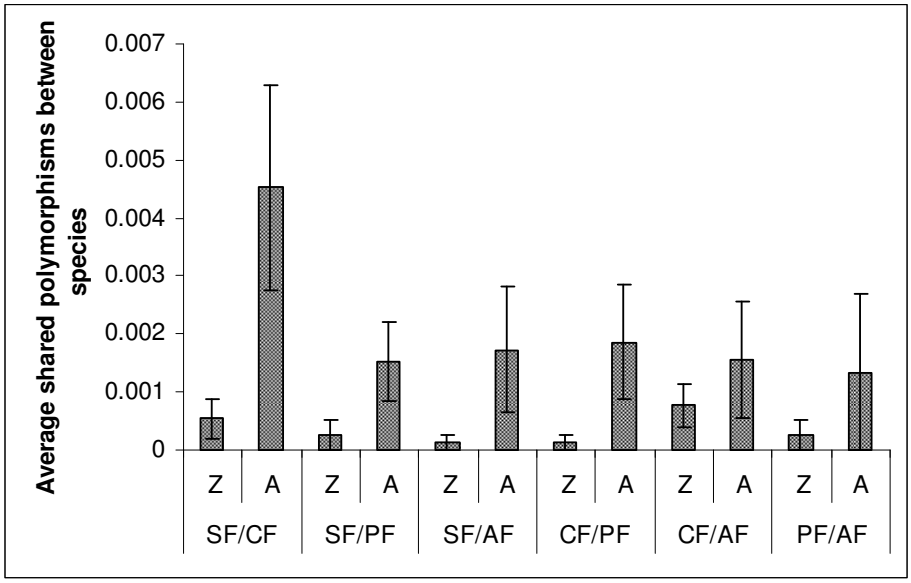


Figure 5. Average number of shared polymorphisms ± standard errors on Z-linked (Z) and autosomal (A) loci between the different flycatcher species (SF/CF; Semi-collared and Collared, SF/PF; Semi-collared and Pied, SF/AF; Semi-collared and Atlas, CF/PF; Collared and Pied, CF/AF; Collared and Pied and PF/AF; Pied and Atlas).

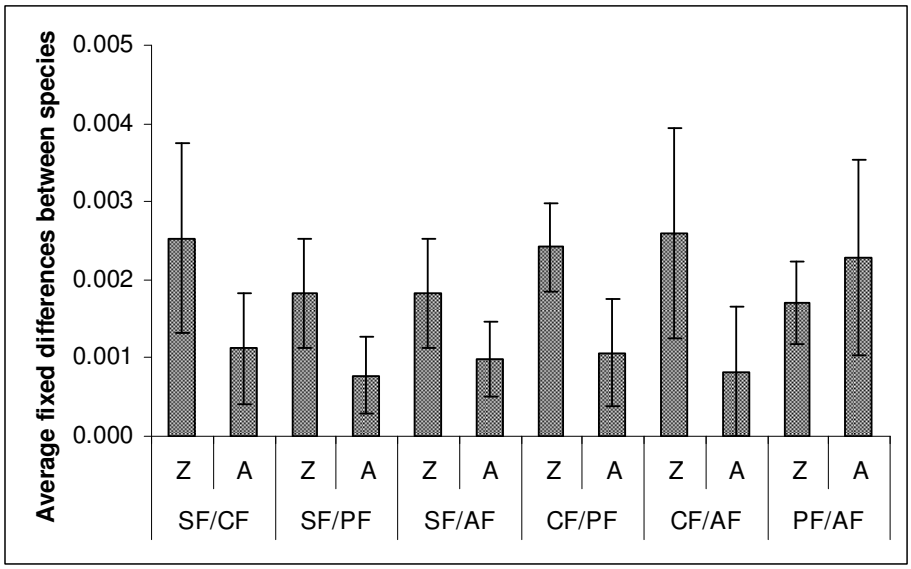


Figure 6. Average number of fixed differences ± standard errors on Z-linked (Z) and autosomal (A) loci between the different flycatcher species (SF/CF; Semi-collared and Collared, SF/PF; Semi-collared and Pied, SF/AF; Semi-collared and Atlas, CF/PF; Collared and Pied, CF/AF; Collared and Pied and PF/AF; Pied and Atlas).

Demographic changes and tests of neutrality

To see if the levels of polymorphism and divergence are correlated between loci and species as predicted under neutrality, multilocus HKA-tests were run for all species pairs for both Z-linked and autosomal loci (Appendix 3). The polymorphism and divergence data from appendix 1 and 2 were used for these tests. For all species pairs 10 000 coalescent simulations were run to assess significance. Previous studies have shown that there is a slight deviation from expected values on the Z-linked loci, but not on the autosomal loci (BORGE *et al.* 2005; BRUVIK 2007). Consistent with this I found that there was a slightly higher deviance on the Z-linked loci from what would be expected. However, none of the tests that were run here showed any statistically significant results (P-values > 0.26).

I performed three tests of neutrality based on the allele frequency distribution; Tajima's D and Fu and Li's D- and F- tests (Appendix 1). The Red-breasted flycatcher was used as an outgroup to determine the ancestral state of diallelic polymorphisms for Fu and Li's statistics. For the Collared and Atlas flycatcher, both Tajima's D and Fu and Li's statistics showed slightly negative values for all Z-linked loci combined and all autosomal loci combined. For the Semi-collared flycatcher, all tests showed negative values, except Tajima's D for all autosomal loci combined, which was positive. For the Pied flycatcher, the total Z-linked loci

showed positive values for all three tests, whereas the autosomal loci showed negative values for the tests. The slightly negative statistic values on both Z-linked and autosomal loci, indicates that the allele frequency spectrum closely matches the neutral expectations, with a small skew towards rare alleles. The positive values of the Pied flycatchers for Z-linked loci indicates a deviation from the neutral model, with a deficiency of rare alleles, These results, even though they are not significant, show the same patterns as previous studies by BORGE *et al.* (2005) and BRUVIK (2007).

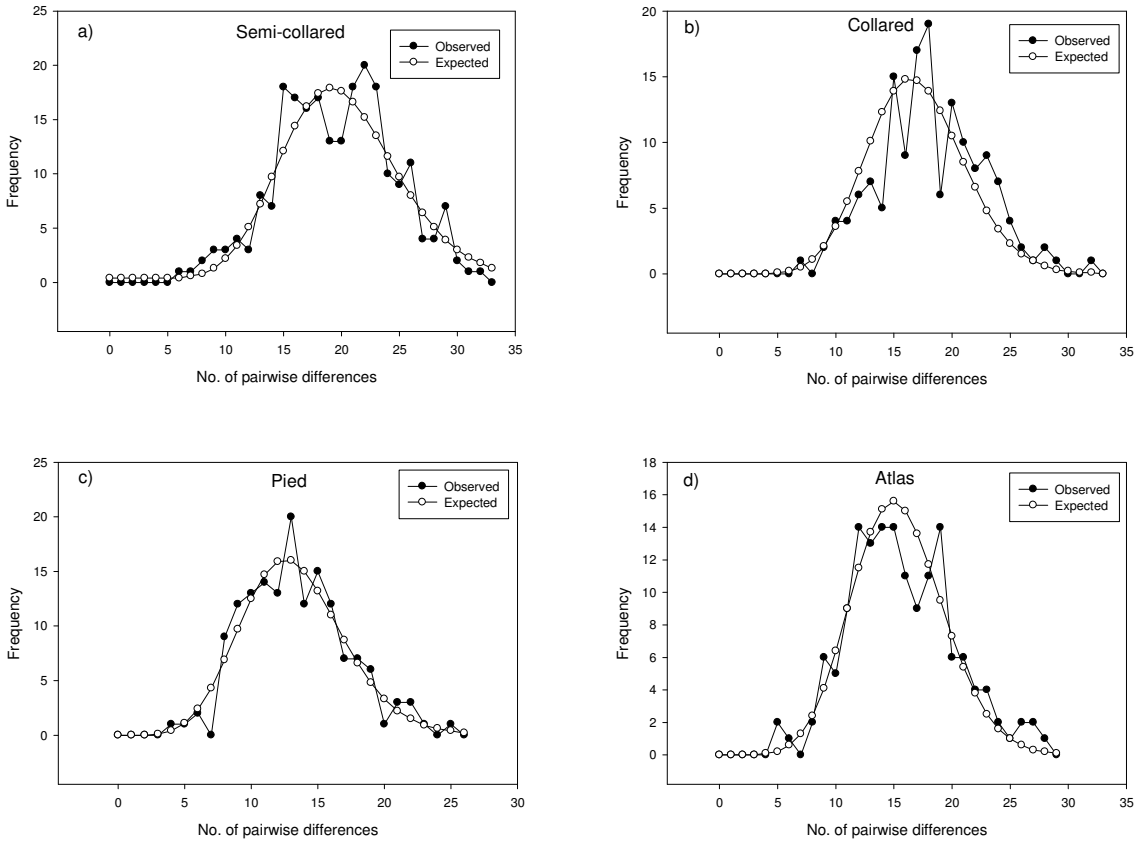


Figure 7: Comparison of the observed and the expected mismatch distribution given a null distribution consistent with a sudden population expansion for a) Semi-collared, b) Collared, c) Pied and d) Atlas flycatchers.

I calculated a mismatch distribution for each of the different species (Figure 7). None of the mismatch distributions deviated from a unimodal pattern. The raggedness index was 0.0067, 0.00125, 0.0204 and 0.0074 for the Semi-collard, Pied, Collared and Atlas flycatcher respectively, with corresponding *P*-values > 0.66, so none of the results were significantly different from an expanding population. The validity of the estimated stepwise expansion

model was also tested using a bootstrap approach, with the sum of squared deviations (SSD) between the observed and expected mismatch as a test statistic (SCHNEIDER and EXCOFFIER 1999). Here, I found the SSD to be 0.0034, 0.0035, 0.0096 and 0.0030 with the corresponding *P*-values > 0.24 for the Semi-collared, Pied, Collared and Atlas flycatchers respectively, so these results are consistent with a sudden population expansion.

I performed a speciation isolation (WH) test between all four species pairs, with all autosomal loci considered as one single locus, and all Z-linked loci considered as one single locus. Any deviations from the isolation model were tested by comparing the observed values of the WH statistics (WAKELEY and HEY 1997) with values from 10 000 coalescent simulations (Table 2). None of the test results showed any significant deviation from the isolation model so the model with no selection and no gene flow since speciation cannot be rejected.

Table 2. Isolation model fitting. θ_1 and θ_2 are the estimates θ for population 1 and 2 of the six different species pair, while θ_A is the estimate of their common ancestral population. T is the estimate of time since their speciation event (T is measured in units of $2N_1$ generations). WH is a test statistic designed for this model (WANG *et al.* 1997) and *P* is the probability calculated from this test statistics.

Species pair	θ_1	θ_2	θ_A	T	WH	*P*
SF/AF	19.921	23.478	29.772	0.560	8.000	0.488
SF/CF	9.156	10.966	59.922	0.452	12.000	0.627
SF/PF	10.184	18.851	41.542	0.685	6.000	0.751
AF/PF	8.282	12.458	45.212	0.615	12.000	0.479
AF/CF	19.084	18.803	47.512	0.531	12.000	0.507
PF/CF	12.620	15.752	43.738	0.597	11.000	0.510

Population divergence

An analysis of molecular variation (AMOVA) showed that around 67% of all variation could be explained by differences between species (F_{ST} = 0.67), around 8% could be explained by variation among individuals within populations, and around 26% could be explained by variation within individuals. All tests were statistically significant, with *P*-values < 0.001 (Table 3). I also computed F_{ST}-values for each of the species pairs (Table 4). F_{ST}-values are a measure of the genetic differentiation among populations, with values ranging from 0 to 1. All the F_{ST}-values were high, quite similar between the different species pairs and all the pairwise F_{ST}-values significantly larger than zero (*P* < 0.001). This shows that the species are genetically strongly differentiated from each other.

None of the species showed any signs of inbreeding. I found no significant deviations from Hardy-Weinberg equilibrium, although the Pied flycatcher population had a rather high F_{IS}-value approaching significance (Table 5). For the other three species only minor deviations from Hardy-Weinberg equilibrium was observed.

Table 3. Analysis of molecular variation (AMOVA). The variation is decomposed into: between species (F_{ST}), between individuals within species, and within individuals' variation.

Source of variation	d.f	Sum of squares	Variance components	Percentage of variation	*P*-value
Among populations	3	999.045	17.024	66.61	<0.001
Among individuals within populations	34	355.692	1.928	7.54	<0.001
Within individuals	38	251.000	6.605	25.85	<0.001
Total	75	1605.737	25.557		

Table 4. Pairwise F_{ST}-values between the different species.

	Semi-collared	Pied	Collared	Atlas
Semi-collared				
Pied	0.664			
Collared	0.621	0.605		
Atlas	0.698	0.717	0.696	

Table 5. **F_{IS} values (inbreeding coefficients) within each species**. *P*-values calculated from deviations from Hardy-Weinberg equilibrium.

Species	F_{IS}	*P*-value
Semi-collared	0.0179	0.4325
Pied	0.1579	0.0942
Collared	0.0704	0.0980
Atlas	-0.0449	0.2609

Phylogenetic reconstruction

I constructed consensus trees based on 1000 bootstrap replications for all autosomal loci combined, and for all Z-linked loci combined. Both the autosomal (Figure 8) and the Z-linked (Figure 9) trees show that there is a deep split between the four species, indicating that they diverged relatively simultaneously. The bootstrap values were higher for the Z-linked tree than for the autosomal tree. This might be because the loci on the Z-chromosomes evolve quicker than autosomal ones, resulting in more divergence in the Z-linked loci. There are some differences between the two trees, when it comes to which species are most closely related to each other. However, the support values on the autosomal tree were low and the branches short on both trees, suggesting that all species diverged from each other at approximately the same time.

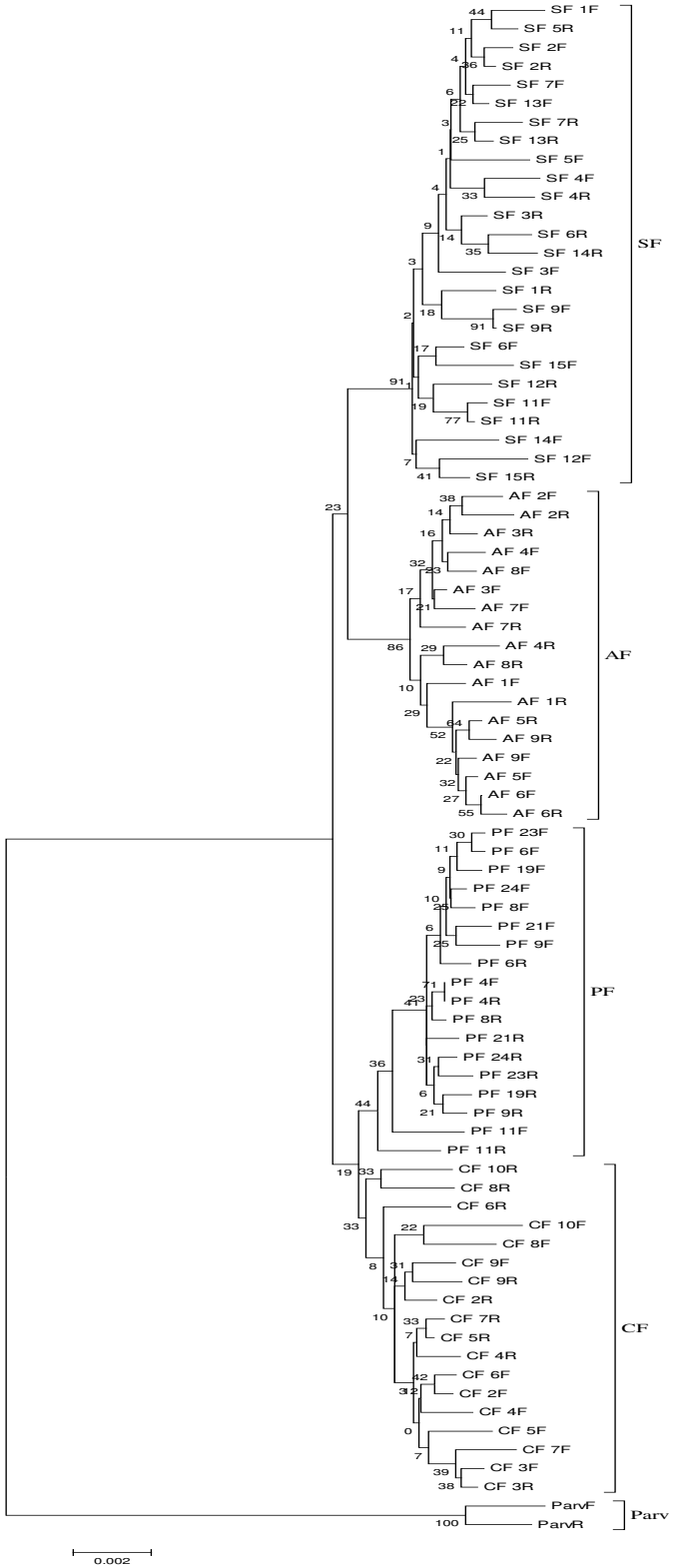


Figure 8. Neighbour joining tree for all autosomal loci combined. AF = Atlas flycatcher, CF = Collared flycatcher, PF = Pied flycatcher, SF = Semi-collared flycatcher, Parv = Red-breasted flycatcher. Length of branches indicates number of changes per site in the aligned sequences. The scale bar shows 0.002 changes per site. Values at nodes represent the bootstrap support values.

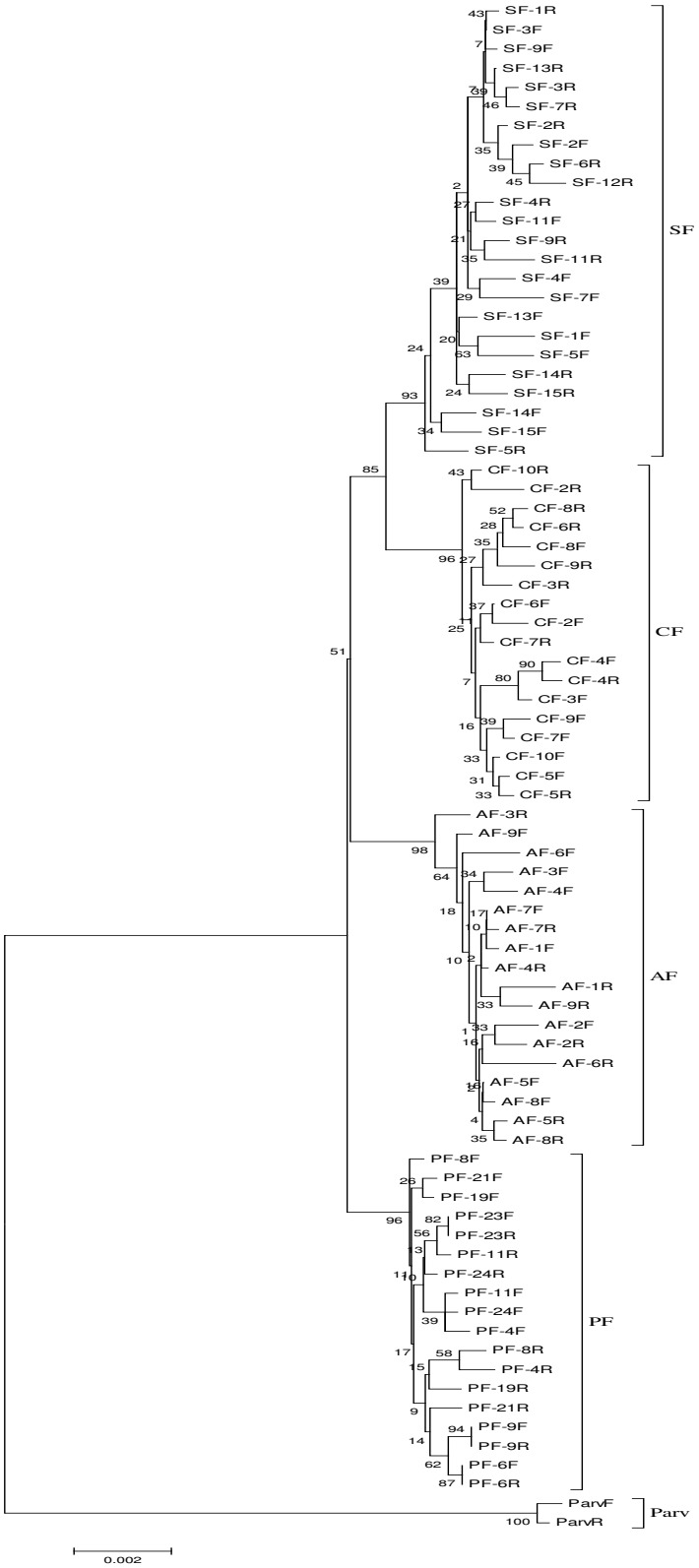


Figure 9. Neighbor joining tree for all Z-linked loci combined. AF = Atlas flycatcher, CF = Collared flycatcher, PF = Pied flycatcher, SF = Semi-collared flycatcher, Parv = Red-breasted flycatcher. Length of branches indicates number of changes per site in the aligned sequences. The scale bar shows 0.002 changes per site. Values at nodes represent the bootstrap support values.

DISCUSSION

The evolutionary history of the black-and-white flycatcher species

Tajima's *D* is often used to infer the past demographic history of a species. It is practical, because it can be used for all sequences (coding as well as none-coding) and for all species without the need of an outgroup. If $\pi > \theta_W$, then Tajima's *D* will be positive, and this might be a result of population expansion, balancing selection or an admixture of two distinctive populations. If $\pi < \theta_W$, then Tajima's *D* will be negative, and a selective sweep or a recent population bottleneck are possible explanations for this (e.g. NIELSEN 2001; RAND 1996; SIMONSEN *et al.* 1995; TAJIMA 1989). In my study, I did not find any significant Tajima's values, but all species showed slightly negative values for all Z-linked and autosomal loci combined, except the Pied and Semi-collared flycatcher. The Pied flycatcher showed a slightly positive Tajima's *D* for all Z-linked loci combined, while the Semi-collared flycatcher showed a slightly positive Tajima's *D* for all autosomal loci combined. All the negative values indicate an overall excess of rare variants. This could potentially reflect a bottleneck. However, for the Pied flycatcher, a positive Tajima's *D* is found at Z-linked loci and a negative value at autosomal loci, indicating an excess of common variants in the Z-linked loci. This might be because of a reduction in population size that will cause a larger initial increase in Tajima's *D* at loci with a smaller N_e. Z-linked loci have a N_e ¾ of autosomal loci. In addition, the sample size used here is small, so there might be a deviation from the expected values by chance. The deviations found here are all small and non-significant, so I cannot reject the neutrality hypothesis.

All the Fu and Li's *D*-values were slightly negative, except for all Z-linked loci combined in the Pied flycatcher. The slightly negative values indicate that there is an excess of recently derived haplotypes, consistent with a population expansion (FU and LI 1993). However, none of the results here are statistical significant, so neutral evolution cannot be ruled out.

I found that the mismatch distribution for each species fits with the scenario of a recent population expansion. This means that the species have not always had the same population structure as they have today. The phylogenetic position of the Semi-collared flycatcher has been unclear, but based on my results, it is clear that the Semi-collared flycatcher should be classified as a species and not a sub-species of the Collared flycatcher. The F_{ST}-values are relatively independent of mutation rate, mutation processes and of total population size (ROUSSET 1997). The F_{ST}-values show the same patterns as the two phylogenetic trees, that all

four species are highly divergent groups. In the tree based on Z-linked loci, all four groups had a bootstrap value above 93, indicating that they are all monophyletic groups. The monophyletic groups found here are consistent with those found by SÆTRE *et al.* (2001). The *Ficedula* species are thought to have diverged from a common ancestor relatively simultaneously during the Pleistocene (SÆTRE *et al.* 2001). It has been suggested that the ancestral flycatcher species became isolated in the four southern refuges where they are found today (Figure 1), during the Pleistocene, and that they were mainly limited by the presence of suitable forest habitat around the Mediterranean sea (SÆTRE *et al.* 2001). My phylogenetic results support the suggestion that the four species diverged from a common ancestor relatively simultaneously. The results from the WH model also supports that the four species split off relatively simultaneously.

Polymorphism and divergence on Z-linked and autosomal loci

In this study, I found reduced variation at Z-linked loci compared to autosomal loci in the Semi-collared flycatcher. This is consistent with previous results found in the Collared and the Pied flycatchers (BORGE *et al.* 2005) and in the Atlas flycatcher (BRUVIK 2007). There are several reasons why different levels of genetic variation in Z-linked and autosomal loci might occur. The first is that there is a higher mutation rate in males compared to females, due to many more rounds of cell divisions in the male germline (spermatogenesis) than in the female germline (oogenesis, HALDANE 1947). This may influence the Z-linked loci such that these get more nucleotide variation than autosomal loci. Another reason is that the effective population size of Z-linked and autosomal loci differ. Random genetic drift is expected to reduce variation and to fix shared polymorphism between two divergent species faster in a small than in a large population (WAKELEY and HEY 1997). The effective population size of Z-linked loci are ¾ that of autosomal ones since females have only one Z-copy. Hence, random genetic drift is therefore expected to reduce variation at Z-linked loci 25% faster than at autosomal loci.

A female-biased sex ratio will reduce the value of N_e at Z-linked loci even further. An operational female-biased sex ratio (some males mate with more than one female) has been observed for both Collared and Pied flycatchers. In a study by QVARNSTRÖM *et al.* (2003) ~ 4% of the Collared flycatcher females were mated with an already mated male, while the corresponding figure in the Pied flycatcher is around 10-15% (LUNDBERG and ALATALO 1992). No such studies have been done on the Atlas or Semi-collared flycatchers, but since

they are closely related to the Pied and Collared flycatchers, it likely that they have a similar mating system, and thus have a somewhat female-biased sex ratio. This could contribute to the observed loss of genetic variation on the Z-linked loci, as argued by BORGE *et al.* (2005). Measuring the extra pair paternity (EPP) rate is another way of measuring operational sex ratio. VEEN *et al.* (2001) found that the mean rate of EPP was 14.5% for the collared flycatcher, and a little lower for the Pied flycatcher.

With no operational sex ratio bias and the same mutation rate at Z-linked loci, the ratios Z_π:A_π and Z_θ:A_θ are predicted to be 0.75. Under the most extreme scenario, where only one male fertilizes all the females in the population, these ratios would approach 0.5. However, a male-biased mutation rate is likely to increase the mutation rate at Z-linked loci, so that the predicted ratios would be higher than 0.5. I this study, I found that the Z_π:A_π ratio was 0.56, 0.50, 0.53 and 0.37 and the Z_θ:A_θ ratio was 0.65, 0.44, 0.44 and 0.53 for the Atlas, Collared, Pied and Semi-collared flycatcher respectively. None of the ratios based on π were statistically significantly different from 0.75, but the Semi-collared ratio was almost significant. For the values based on θ_W, the Pied flycatcher ratio was significantly different from 0.75. These results show the same trends as previous studies by BORGE *et al.* (2005) and BRUVIK (2007). However, these previous studies did not attempt to find any statistical support for their results. The low Z:A-ratios found here is consistent with selective sweeps having occurred on the Z-chromosome in the four species. On the other hand, the HKA-test and the neutrality test did not yield any significant deviations from the neutral expectation. Further studies with more loci and a larger sample size are therefore needed before any firm conclusions can be drawn.

An isolation model states that as two populations diverge and become isolated from each other, shared polymorphisms are gradually lost and become fixed differences due to random drift (WAKELEY and HEY 1997; WANG *et al.* 1997). This process will occur faster when the population size is small, and the Z-linked loci will therefore be expected to have fewer shared polymorphisms and more fixed differences than the autosomal loci. However, the isolation model has a number of factors that might cause populations to deviate from this model. These include gene flow between species, selective sweeps and an unequal variance in mating success between males and females (BORGE *et al.* 2005). In this study, I found a trend that there was a higher level of shared polymorphisms on autosomal loci, and a higher level of fixed differences on Z-linked loci. This trend is consistent with a trend found by BORGE *et al.*

(2005) and BRUVIK (2007). The isolation model fitting indicates that the values found here are compatible with the isolation model of speciation, and these results suggest that patterns of polymorphism and divergence could be a result from a recent scenario without gene flow or selection. However, it is possible that historical changes in θ (perhaps from changes in N_e) would have generated a greater degree of stochastic variance in these patterns (BORGE *et al.* 2005).

When comparing the variation on Z-linked and autosomal loci between species, I found that there was a clear trend of more shared polymorphisms at autosomal loci and a higher level of fixed differences on the Z-linked loci. One explanation for this might be that there is a higher mutation rate on Z-linked loci resulting in more fixed differences after species divergence. The reduced variation on Z-linked loci found within the species might be consistent with this result if the polymorphisms found at the Z-linked loci have been fixed between the different species. Another explanation might be that different selection sweeps have occurred in the four species. This can result in a higher level of fixed differences on Z-liked loci, even if the mutation rate have been the same on Z-linked and autosomal loci.

Genetic variation within and between species

I found that the frequency of polymorphic sites in the Semi-collared flycatcher was not significantly lower than in the other flycatcher species. The Semi-collared flycatcher is much rarer and has a more fragmented population structure than the Pied and the Collared flycatchers. I therefore expected to find less genetic variation in this species than in the other ones. It is possible that the Semi-collared flycatcher had much larger population sizes and a less fragmented distribution until quite recently (supported by my mismatch results, but uncertain from the frequency spectra analyses), and that the population decline happened so recently that most of the genetic variation still persist.

However, the populations of the four species investigated here may not be representative for the overall differences in population sizes and structure of the species. The population of Semi-collared flycatchers in Kamcheya is possibly one out of very few relatively large persisting populations of this species and may thus be genetically more variable than the average. It is also possible that the local populations of the other three species investigated here are smaller than average for these species. Indeed, the Spanish population of the Pied flycatcher investigated here showed signs of a relatively small population size in a previous

study (HAAVIE *et al.* 2000). These signs were reduced allelic diversity and heterozygous deficiency. HAAVIE *et al.* (2000) used microsatellites, and compared this Pied flycatcher population from Spain with populations from the Czech Republic and Norway. Clearly, more populations of the different species should be investigated and samples from larger populations are needed.

Conservation issues

There is limited information available about the Semi-collared flycatcher. The reason for this might be that the Semi-collared flycatcher consists of small and fragmented populations (BURFIELD *et al.* 2004; CRAMP and PERRINS 1993), and that it has been suggested to be only a sub-species of the Collared flycatcher until relatively recently (SÆTRE *et al.* 2001; SÆTRE *et al.* 1997). The limited information about the Semi-collared flycatcher makes it difficult to know exactly how it should be classified in the IUCN red list. It is currently listed as "near threatened". However, more information about the species is needed to be certain about this classification. Examples of information needed are more accurate information on the different breeding sub populations; their population sizes, breeding habitat and the status of the breeding habitat. Since many of the sub populations appears to be threatened by habitat fragmentation, it is important to monitor and reduce this fragmentation. In addition, it would be helpful to check the genetic variation in other sub populations in order to find out which of the sub populations have the highest chance of further existence. One of these sub populations might be the one in Kamcheya, since it seems that this sub population has relatively high genetic diversity.

It has been suggested that the Semi-collared flycatcher has gone through severe declines in Europe during the 1970-1990 and early 2000 (CRAMP and PERRINS 1993; TUCKER and HEATH 1994). The reason for this is probably because much of the habitat that the Semi-collared flycatcher occupies has been destroyed during the last few decades (BIRDLIFE 2008; TUCKER and HEATH 1994). Previous studies on other species have shown that habitat destruction can lead to inbreeding and that this inbreeding can be severely reduced by only one unrelated migrant into the population (FRANKLIN 1980). This makes it important to conserve the remaining habitat which the Semi-collared flycatchers occupy, since they already consist of small, patchy populations where migration might be low between the sub populations. With continuing degradation of the habitat, more populations will be fragmented, resulting in smaller population sizes, and possibly more inbreeding.

CONCLUSION

Even though my analyses did not show any signs of reduced population size, the current classification of the Semi-collared flycatcher as "near threatened" is probably the right classification today. This is because the sub population studied here is considered to be one of the largest in Europe. The Semi-collared flycatcher consists of several small fragmented populations, and much of the distribution, especially in the Middle East is uncertain (CRAMP and PERRINS 1993). Previous studies on other species have shown that habitat fragmentation can be an important factor in causing inbreeding depression. Thus, it is important to keep monitoring this bird and the habitat it occupies. I suggest that more studies of the Semi-collared flycatcher should be conducted. In these studies, the genetic variation in the Semi-collared flycatcher should be compared to larger populations of the other species than those used in the present study, to see if the results found here are representative. In addition, it is important to keep monitoring different sub populations of the Semi-collared flycatcher and conserve its habitat, in order to maintain this species in the future.

REFERENCES

AKEY, J. M., M. A. EBERLE, M. J. RIEDER, C. S. CARLSON, M. D. SHRIVER *et al.*, 2004 Population history and natural selection shape patterns of genetic variation in 132 genes. PLoS Biology **2:** e286.

BAILLIE, J. E. M., C. HILTON-TAYLOR and S. N. STUART, 2004 *2004 IUCN Red List of Threatened Species™ A global species assessment*. IUCN Publications Services Unit, Gland, Switzerland and Cambridge, UK.

BIRDLIFE, 2008 Birdlife international web page (http://www.birdlife.eu/index.html). Accessed 27 March 2008.

BORGE, T., M. T. WEBSTER, G. ANDERSSON and G. P. SÆTRE, 2005 Contrasting patterns of polymorphism and divergence on the Z chromosome and autosomes in two *Ficedula* flycatcher species. Genetics **171:** 1861-1873.

BRUVIK, T., 2007 *Tracing the evolutionary history of the Atlas flycatcher (Ficedula speculigera) - A molecular genetic approach*. Unpublished MSc-thesis, University of Oslo.

BURFIELD, I., F. V. BOMMEL and U. GALLO-ORSI, 2004 *Birds in Europe: Population estimates, trends and conservation status*. BirdLife International, Cambridge.

CRAMP, S., and C. M. PERRINS, 1993 *Handbook of the birds of Europe, the Middle East and North Africa: the birds of the Western Palearctic, Vol 7: Flycatchers to Shrikes*. Oxford University Press, Oxford.

CURIO, V. E., 1959 Beobachtungen am Halbringschnäpper, *Ficedula semitorquata*, im Mazedonischen Brutgebiet. Journal für Ornitologie **100:** 176-209.

EXCOFFIER, L., G. LAVAL and D. BALDING, 2003 Gametic phase estimation over large genomic regions using an adaptive window approach. Human Genetics **1:** 7-19.

EXCOFFIER, L., G. LAVAL and S. SCHNEIDER, 2005 Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Bioinformatics Online **1:** 47-50.

FAY, J. C., and C.-I. WU, 2000 Hitchhiking under positive Darwinian selection. Genetics **155:** 1405-1413.

FRANKHAM, R., 1996 Relationship of genetic variation to population size in wildlife. Conservation Biology **10:** 1500-1508.

FRANKHAM, R., 1999 Quantitative genetics in conservation biology. Genetical Research **74:** 237-244.

FRANKHAM, R., J. D. BALLOU and B. D. A, 2004 *A primer of conservation genetics*. Cambridge University Press, Cambridge.

FRANKLIN, I. R., 1980 Evolutionary change in small populations, pp. 135-149 in *Conservation biology: an evolutionary-ecological perspective*, edited by M. E. SOULÉ and B. A. WILCOX. Sinauer Associates, INC. Publisher, Sunderland.

FU, Y. X., and W. H. LI, 1993 Statistical tests of neutrality of mutations. Genetics **133:** 693-709.

GILPIN, M. E., and M. E. SOULÉ, 1986 Minimum viable populations: processes of species extinction, pp. 19-34 in *Conservation biology: the science of scarcity and diversity*, edited by M. E. SOULÉ. Sinauer associates, INC, Sunderland.

HALDANE, J. B. S., 1947 The Mutation rate of the gene for Haemophilia, and its segregation ratios in males and females. Annals of Eugenics **13:** 262-271.

HUDSON, R. R., M. KREITMAN and M. AGUADE, 1987 A test of neutral molecular evolution based on nucleotide data. Genetics **116:** 153-159.

HAAVIE, J., G. P. SÆTRE and T. MOUM, 2000 Discrepancies in population differentiation at microsatellites, mitochondrial DNA and plumage colour in the pied flycatcher - inferring evolutionary processes. Molecular Ecology **9:** 1137-1148.

IUCN, 2007 2007 IUCN Red list of threatened species (http://www.iucnredlist.org/). Accessed 27 March 2008.

KREITMAN, M., 2000 Methods to detect selection in populations with applications to the human. Annual Review of Genomics and Human Genetics **1:** 539-559.

KUMAR, S., K. TAMURA and M. NEI, 2004 MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Briefings in Bioinformatics **5:** 150-163.

LANDE, R., 1988 Genetics and demography in biological conservation. Science **241:** 1455-1460.

LAVERY, S., C. MORITZ and D. R. FIELDER, 1996 Genetic patterns suggest exponential population growth in a declining species. Molecular Biology and Evolution **13:** 1106-1113.

LUNDBERG, A., and R. V. ALATALO, 1992 *The pied flycatcher*. T & A D Poyser, London.

NEI, M., 1987 *Molecular evolutionary genetics*. Columbia University Press, New York.

NIELSEN, R., 2001 Statistical tests of selective neutrality in the age of genomics. Heredity **86:** 641-647.

PACKER, C., 1979 Inter-troop transfer and inbreeding avoidance in *Papio anubis*. Animal Behaviour **27:** 1-36.

PRIMMER, C. R., T. BORGE, J. LINDELL and G. P. SÆTRE, 2002 Single-nucleotide polymorphism characterization in species with limited available sequence information: high nucleotide diversity revealed in the avian genome. Molecular Ecology **11:** 603-612.

QVARNSTRÖM, A., B. C. SHELDON, T. PART and L. GUSTAFSSON, 2003 Male ornamentation, timing of breeding, and cost of polygyny in the collared flycatcher. Behavioral Ecology **14:** 68-73.

RAND, D. M., 1996 Neutrality tests of molecular markers and the connection between DNA polymorphism, demography, and conservation biology. Conservation Biology **10:** 665-671.

REED, D. H., 2005 Relationship between population size and fitness. Conservation Biology **19:** 563-568.

REED, D. H., and R. FRANKHAM, 2003 Correlation between fitness and genetic diversity. Conservation Biology **17:** 230-237.

REYNOLDS, J., B. S. WEIR and C. C. COCKERHAM, 1983 Estimation of the coancestry coefficient: basis for a short-term genetic distance. Genetics **105:** 767-779.

ROGERS, A. R., and H. HARPENDING, 1992 Population growth makes waves in the distribution of pairwise genetic differences. Molecular Biology and Evolution **9:** 552-569.

ROUSSET, F., 1997 Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. Genetics **145:** 1219-1228.

ROZAS, J., and R. ROZAS, 1999 DnaSP version3: an integrated program for molecular populaiton genetics and molecular evolution analysis. Bioinformatics **15:** 174-175.

SCHLÖTTERER, C., 2003 Hitchhiking mapping - functional genomics from the population genetics perspective. Trends in Genetics **19:** 32-38.

SCHNEIDER, S., and L. EXCOFFIER, 1999 Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. Genetics **152:** 1079-1089.

SEUTIN, G., B. N. WHITE and P. T. BOAG, 1991 Preservation of avian blood and tissue samples for DNA analyses. Canadian Journal of Zoology **69:** 82-90.

SIMONSEN, K. L., G. A. CHURCHILL and C. F. AQUADRO, 1995 Properties of statistical tests of neutrality for DNA polymorphism data. Genetics **141:** 413-429.

SLATKIN, M., 1995 A measure of population subdivision based on microsatellite allele frequencies. Genetics **139:** 457-462.

SLATKIN, M., and R. R. HUDSON, 1991 Pairwise comparisons of mitochondrial-DNA sequences in stable and exponentially growing populations. Genetics **129:** 555-562.

SPIELMAN, D., B. W. BROOK and R. FRANKHAM, 2004 Most species are not driven to extinction before genetic factors impact them. Proceedings of the National Academy of Sciences USA **101:** 15261-15264.

SÆTHER, S. A., G. P. SÆTRE, T. BORGE, C. WILEY, N. SVEDIN *et al.*, 2007 Sex chromosome-linked species recognition and evolution of reproductive isolation in flycatchers. Science **318:** 95-97.

SÆTRE, G. P., T. BORGE, J. LINDELL, T. MOUM, C. R. PRIMMER *et al.*, 2001 Speciation, introgressive hybridization and nonlinear rate of molecular evolution in flycatchers. Molecular Ecology **10:** 737-749.

SÆTRE, G. P., T. MOUM, S. BURES, M. KRAL, M. ADAMJAN *et al.*, 1997 A sexually selected character displacement in flycatchers reinforces premating isolation. Nature **387:** 589-592.

TAJIMA, F., 1989 Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics **123:** 585-595.

TUCKER, M. G., and M. F. HEATH, 1994 *Birds in Europe: their conservation status*. Birdlife International, Cambridge.

VEEN, T., T. BORGE, S. C. GRIFFITH, G. P. SÆTRE, S. BURES *et al.*, 2001 Hybridization and adaptive mate choice in flycatchers. Nature **411:** 45-50.

WAKELEY, J., and J. HEY, 1997 Estimating ancestral population parameters. Genetics **145:** 847-855.

WANG, R. L., J. WAKELEY and J. HEY, 1997 Gene flow and natural selection in the origin of *Drosophila pseudoobscura* and close relatives. Genetics **147:** 1091-1106.

WATTERSON, G. A., 1975 Number of segregating sites in genetic models without recombination. Theoretical Population Biology **7:** 256-276.

WEIR, B. S., and C. C. COCKERHAM, 1984 Estimating *F*-statistics for the analysis of population structure. Evolution **38:** 1358-1370.

APPENDIX

Appendix 1. Polymorphism summary. a) Z-linked genes, b) autosomal genes.

a)

Gene	Species[a]	L[b]	K[c]	S[d]	S_{exp}[e]	s[f]	π[g]	SD	θ_w[h]	SD	Dti	D_{FL}[j]	F_{FL}[k]
Aldob-6	AF	465	0	10	0.0000	0	0.0000	±0.0000	0.0000	±0.0000	0.000	0.000	0.000
	CF	465	1	9	0.5431	0	0.0005	±0.0003	0.0006	±0.0006	-0.529	0.667	0.405
	PF	465	3	9	1.0297	1	0.0017	±0.0004	0.0019	±0.0012	-0.259	-0.085	-0.151
	SF	465	10	0	3.4322	0	0.0000	±0.0000	0.0000	±0.0000	0.000	0.000	0.000
Brm-12	AF	1554	36	15	15.3621	8	0.0017	±0.0004	0.0028	±0.0012	-1.524	-1.108	-1.420
	CF	1554	40	9	21.7241	5	0.0013	±0.0002	0.0017	±0.0008	-0.720	-1.113	-1.157
	PF	1553	37	10	12.6992	0	0.0028	±0.0002	0.0019	±0.0009	1.791	1.412*	1.754*
	SF	1549	33	27	11.3263	10	0.0034	±0.0004	0.0047	±0.0017	-1.092	-0.680	-0.940
Chdz	AF	687	14	1	5.6757	0	0.0006	±0.0001	0.0004	±0.0004	0.870	0.667	0.822
	CF	687	13	1	7.0603	1	0.0002	±0.0001	0.0004	±0.0004	-1.165	-1.499	-1.612
	PF	687	14	2	4.8051	0	0.0007	±0.0002	0.0009	±0.0006	-0.328	0.885	0.643
	SF	687	13	1	4.4619	0	0.0004	±0.0002	0.0004	±0.0004	0.139	0.623	0.565
Ghr	AF	548	9	3	3.6486	0	0.0023	±0.0003	0.0016	±0.0010	1.175	1.024	1.221
	CF	555	9	1	4.8879	1	0.0012	±0.0002	0.0011	±0.0008	0.248	-0.552	-0.388
	PF	555	10	0	3.4322	0	0.0000	±0.0000	0.0000	±0.0000	0.000	0.000	0.000
	SF	555	10	5	3.4322	3	0.0018	±0.0004	0.0024	±0.0013	-0.726	-1.340	-1.350
Vldlr	AF	712	5	14	2.0270	5	0.0041	±0.0009	0.0056	±0.0024	-1.002	-0.274	-0.560
	CF	712	4	13	2.1724	3	0.0060	±0.0006	0.0053	±0.0023	0.469	0.370	0.461
	PF	712	4	3	1.3729	0	0.0009	±0.0005	0.0012	±0.0008	-0.778	-0.778	0.617
	SF	712	7	7	2.4025	4	0.0018	±0.0003	0.0026	±0.0013	-1.041	-1.350	-1.460
Total Z	AF	3966	74	30	37.6271	13	0.0017	±0.0002	0.0022	±0.0009	-0.992	-0.668	-0.884
	CF	3973	75	26	40.7328	10	0.0018	±0.0001	0.0019	±0.0007	-0.136	-0.404	-0.378
	PF	3972	74	18	25.3983	1	0.0016	±0.0001	0.0013	±0.0005	0.769	1.272	1.305
	SF	3966	73	40	25.0551	17	0.0020	±0.0002	0.0027	±0.0010	-1.095	-0.990	-1.200

b)

Gene	Species[a]	L[b]	K[c]	S[d]	S_{exp}[e]	s[f]	π[g] ± SD		θ_w[h] ± SD		Dt[i]	D_{FL}[j]	F_{FL}[k]
Acly-16	AF	358	5	0	0.0000	0	0.0000	±0.0000	0.0000	±0.0000	0.000	0.000	0.000
	CF	358	5	1	0.6897	1	0.0003	±0.0003	0.0008	±0.0008	-1.165	-1.499	-1.612
	PF	358	5	2	1.3793	2	0.0006	±0.0004	0.0016	±0.0012	-1.508	-1.989	-2.130
	SF	358	5	1	0.6897	1	0.0002	±0.0002	0.0007	±0.0007	-1.156	-1.634	-1.727
Alas1-8	AF	290	12	5	3.4483	1	0.0047	±0.0008	0.0050	±0.0027	-0.219	0.420	0.281
	CF	290	14	7	4.8276	3	0.0055	±0.0010	0.0070	±0.0035	-0.745	-0.513	-0.666
	PF	290	15	4	2.7586	2	0.0022	±0.0008	0.0040	±0.0023	-1.347	-0.701	-1.008
	SF	288	13	8	5.5172	3	0.0071	±0.0008	0.0073	±0.0034	-0.066	-0.500	-0.432
Fas-y	AF	565	10	0	0.0000	0	0.0000	±0.0000	0.0000	±0.0000	0.000	0.000	0.000
	CF	566	11	2	1.3793	0	0.0009	±0.0003	0.0010	±0.0008	-0.328	0.885	0.643
	PF	565	10	0	0.0000	0	0.0000	±0.0000	0.0000	±0.0000	0.000	0.000	0.000
	SF	565	10	2	1.3793	0	0.0003	±0.0002	0.0009	±0.0007	-1.513	-2.204	-2.319
Rho-1	AF	375	9	4	2.7586	1	0.0032	±0.0005	0.0031	±0.0018	0.085	0.211	0.203
	CF	375	7	7	4.8276	1	0.0062	±0.0007	0.0054	±0.0027	0.508	0.701	0.746
	PF	375	7	7	4.8276	1	0.0067	±0.0007	0.0054	±0.0027	0.813	0.701	0.844
	SF	375	6	11	7.5862	3	0.0069	±0.0008	0.0077	±0.0033	-0.333	-0.001	-0.118
Rpl30-3	AF	983	17	18	12.4138	9	0.0051	±0.0004	0.0053	±0.0022	-0.187	-0.969	-0.862
	CF	983	16	19	13.1034	4	0.0053	±0.0007	0.0056	±0.0023	-0.244	0.494	0.327
	PF	983	17	11	7.5862	5	0.0026	±0.0008	0.0033	±0.0015	-0.718	-0.685	-0.802
	SF	983	21	15	10.3448	3	0.0055	±0.0004	0.0040	±0.0016	1.295	0.408	0.408
Tgfb2-5	AF	402	5	6	4.1379	2	0.0040	±0.0006	0.0043	±0.0022	-0.245	-0.104	-0.165
	CF	402	5	6	4.1379	4	0.0028	±0.0008	0.0043	±0.0022	-1.208	-1.468	-1.607
	PF	402	5	3	2.0690	1	0.0023	±0.0004	0.0022	±0.0014	0.115	-0.085	-0.035
	SF	401	5	6	4.1379	3	0.0042	±0.0004	0.0039	±0.0020	0.243	-1.023	-0.758
Total A	AF	2973	58	33	22.7586	13	0.0031	±0.0002	0.0032	±0.0012	-0.190	-0.462	-0.444
	CF	2974	58	42	28.9655	13	0.0036	±0.0003	0.0041	±0.0015	-0.465	-0.009	-0.165
	PF	2973	59	27	18.6207	11	0.0023	±0.0003	0.0026	±0.0010	-0.502	-0.525	-0.601
	SF	2970	60	43	29.6552	15	0.0040	±0.0002	0.0038	±0.0013	0.238	-0.494	-0.308

[a] AF, Atlas Flycatcher; CF, Collared Flycatcher; PF, Pied Flycatcher; SF, Semi-collared Flycatcher. [b] Number of sites surveyed. [c] Divergence with outgroup. [d] Number of segregating sites. [e] The expected number (using formula 1 or 2) of segregating sites based on N_e estimated from autosomal or Z-linked loci. [f] Number of singleton sites. [g] Average pairwise sequence differences per nucleotide (Nei 1987) and their standard deviation. [h] Expected heterozygosity per nucleotide (Watterson 1975) and their standard deviation. [i] Tajima's D (Tajima 1989). [j] Fu and Li's D (Fu and Li 1993). [k] Fu and Li's F (Fu and Li 1993).

Appendix 2. Fixed and shared polymorphisms between species. S-values are the number of variable sites that occur in or between the different species pairs (a-f), Semi-collared flycatcher (S_{SF}), Collared flycatcher (S_{CF}), the Pied flycatcher (S_{PF}) and the Atlas flycatcher (S_{AF}). S_{shared} indicates the number of shared polymorphisms between the different species; S_{fixed} indicates the number of fixed differences between the species and $S_{average}$ indicates the average pairwise differences between the species pairs. Total-Z is all Z-linked loci combined and total-A is all autosomal loci combined.

a)

Locus	S_{SF}	S_{CF}	S_{shared}	S_{fixed}	$S_{average}$	Length
ALDOB-6	0	1	0	3	3,889	465
BRM-12	26	7	2	4	11,010	1549
CHDZ	1	1	0	0	0,222	687
GHR	5	0	0	2	2,958	555
VLDLR	6	12	1	0	3,444	712
Total-Z	38	23	3	7	20,097	3966
ACLY-16	1	1	0	0	0.111	358
ALAS1-8	7	6	1	0	3.969	288
FAS-Y	2	2	0	3	3.389	807
RHO-1	4	3	4	0	4.093	375
RPL30-3	5	11	8	3	12.204	996
TGFB2-5	4	4	2	0	1.997	402
Total-A	23	27	15	6	25.763	3226

b)

Locus	S_{SF}	S_{PF}	S_{shared}	S_{fixed}	$S_{average}$	Length
ALDOB-6	0	3	0	1	1,833	465
BRM-12	26	8	2	2	10,653	1548
CHDZ	1	2	0	1	1,444	687
GHR	5	2	0	0	1,514	555
VLDLR	7	3	0	3	4,625	712
Total-Z	39	16	2	9	21,597	3965
ACLY-16	1	2	0	0	0.167	358
ALAS1-8	7	3	1	0	2.216	288
FAS-Y	2	0	0	0	0.111	807
RHO-1	7	6	1	1	5.327	375
RPL30-3	10	9	3	2	12.969	996
TGFB2-5	6	3	0	0	2.278	402
Total-A	33	23	5	3	23.068	3226

c)

Locus	S_{SF}	S_{AF}	S_{shared}	S_{fixed}	$S_{average}$	Length
ALDOB-6	0	3	0	1	1,833	465
BRM-12	26	8	2	2	10,653	1548
CHDZ	1	2	0	1	1,444	687
GHR	5	2	0	0	1,514	555
VLDLR	7	3	0	3	4,625	712
Total-Z	39	16	2	9	21,597	3965
ACLY-16	1	2	0	0	0.167	358
ALAS1-8	7	3	1	0	2.216	288
FAS-Y	2	0	0	0	0.111	807
RHO-1	7	6	1	1	5.327	375
RPL30-3	10	9	3	2	12.969	996
TGFB2-5	6	3	0	0	2.278	402
Total-A	33	23	5	3	23.068	3226

d)

Locus	S_{CF}	S_{PF}	S_{shared}	S_{fixed}	$S_{average}$	Length
ALDOB-6	1	3	0	2	3.722	465
BRM-12	8	9	1	5	11.247	1569
CHDZ	1	2	0	1	1.333	702
GHR	2	0	0	1	1.556	555
VLDLR	13	3	0	1	4.611	731
Total-Z	25	17	1	10	22.469	4022
ACLY-16	1	2	0	0	0.167	358
ALAS1-8	7	4	0	0	2.611	290
FAS-Y	2	0	0	3	3.278	807
RHO-1	6	6	1	1	4.716	375
RPL30-3	13	6	6	0	6.13	996
TGFB2-5	5	2	1	0	1.377	402
Total-A	34	20	8	4	18.279	3228

e)

Locus	S_{CF}	S_{AF}	S_{shared}	S_{fixed}	$S_{average}$	Length
ALDOB-6	1	0	0	3	3.889	465
BRM-12	8	14	1	8	12.204	1569
CHDZ	1	1	0	1	1.333	702
GHR	1	2	1	0	1.833	548
VLDLR	12	11	1	0	7.889	731
Total-Z	23	28	3	12	27.148	4015
ACLY-16	1	0	0	0	0.056	358
ALAS1-8	7	5	0	0	3.333	290
FAS-Y	2	2	0	4	4.444	807
RHO-1	5	2	2	0	2.981	375
RPL30-3	15	15	4	0	12.068	996
TGFB2-5	6	6	0	0	1.556	402
Total-A	36	30	6	4	24.438	3228

f)

Locus	S_{PF}	S_{AF}	S_{shared}	S_{fixed}	$S_{average}$	Length
ALDOB-6	3	0	0	1	1.833	465
BRM-12	8	13	2	5	10.815	1568
CHDZ	2	1	0	0	0.556	702
GHR	0	3	0	1	2.778	548
VLDLR	3	12	0	1	6.111	731
Total-Z	16	29	2	8	22.093	4014
ACLY-16	2	0	0	0	0.111	358
ALAS1-8	4	5	0	1	3.167	290
FAS-Y	0	2	0	1	1.167	807
RHO-1	7	4	0	3	6.389	375
RPL30-3	4	11	8	1	11.432	996
TGFB2-5	3	6	0	0	1.778	402
Total-A	20	28	8	6	24.044	3228

Appendix 3. HKA-results. Expected and observed number of segregating sites for each species, and the divergence (average no. of nucleotide differences) between the different species pairs (a-f).

a)

	Semi-collared flycatcher		Collared flycatcher		Divergence	
Locus	Observed	Expected	Observed	Expected	Observed	Expected
ALDOB-6	0	2.03	1	1.65	3.89	1.21
BRM-12	26	18.26	7	14.86	11.01	10.88
CHDZ	1	0.92	1	0.75	0.22	0.55
GHR	5	3.54	2	2.87	1.51	2.11
VLDLR	6	8.91	12	7.23	3.44	5.30
ACLY-16	1	0.91	1	0.73	0.09	0.46
ALAS1-8	8	8.21	7	6.57	3.89	4.11
FAS-Y	2	2.33	2	1.86	1.36	1.16
RHO-1	11	9.65	7	7.66	4.12	4.81
PRL30-3	15	20.09	19	15.95	12.05	10.01
TGFB2-5	6	6.13	6	4.88	2.06	3.05

b)

	Semi-collared flycatcher		Pied flycatcher		Divergence	
Locus	Observed	Expected	Observed	Expected	Observed	Expected
ALDOB-6	0	2.27	3	1.20	1.83	1.37
BRM-12	26	20.92	8	11.12	10.65	12.62
CHDZ	1	2.08	2	1.10	1.44	1.26
GHR	5	3.73	0	1.98	2.96	2.25
VLDLR	7	6.86	3	3.64	4.63	4.13
ACLY-16	1	1.57	2	0.81	0.15	0.77
ALAS1-8	8	7.04	4	3.68	2.21	3.49
FAS-Y	2	1.03	0	0.54	0.08	0.51
RHO-1	11	11.60	7	6.02	5.35	5.73
PRL30-3	15	19.27	11	9.99	12.77	9.51
TGFB2-5	6	5.63	3	2.93	2.33	2.78

c)

	Semi-collared flycatcher		Atlas flycatcher		Divergence	
Loci	Observed	Expected	Observed	Expected	Observed	Expected
ALDOB-6	0	0.85	0	0.63	2.00	0.52
BRM-12	27	21.45	13	15.98	10.74	13.30
CHDZ	1	1.46	1	1.08	1.44	0.90
GHR	5	4.57	3	3.35	2.74	2.81
VLDLR	7	10.82	11	8.04	7.57	6.71
ACLY-16	1	0.46	0	0.34	0.04	0.24
ALAS1-8	8	7.75	5	5.67	4.41	3.99
FAS-Y	2	0.93	0	0.67	0.08	0.48
RHO-1	11	8.65	4	6.29	4.39	4.45
PRL30-3	15	19.62	18	14.26	10.95	10.08
TGFB2-5	6	6.44	6	4.69	2.44	3.31

d)

	Collared flycathcer		Pied flycatcher		Divergence	
Locus	Observed	Expected	Observed	Expected	Observed	Expected
ALDOB-6	3	2.24	1	3.38	3.72	2.10
BRM-12	10	8.70	9	13.15	11.00	8.15
CHDZ	2	1.26	1	1.90	1.33	1.18
GHR	0	1.03	2	1.56	1.56	0.97
VLDLR	3	5.98	13	9.03	4.61	5.60
ACLY-16	2	0.96	1	1.45	0.17	0.76
ALAS1-8	4	4.12	7	6.22	2.61	3.27
FAS-Y	0	0.99	2	1.50	1.28	0.79
RHO-1	7	5.66	7	8.56	4.72	4.49
PRL30-3	11	10.92	19	16.49	6.07	8.66
TGFB2-5	3	3.14	6	4.74	1.38	2.49

e)

	Collared flycatcher		Atlas flycatcher		Divergence	
Locus	Observed	Expected	Observed	Expected	Observed	Expected
ALDOB-6	1	1.79	0	1.74	3.89	1.37
BRM-12	9	13.24	15	12.68	12.2	10.11
CHDZ	1	1.22	1	1.18	1.33	0.93
GHR	2	2.52	3	2.41	1.83	1.91
VLDLR	13	12.64	14	12.28	7.58	9.66
ACLY-16	1	0.40	0	0.39	0.06	0.26
ALAS1-8	7	5.86	5	5.69	3.33	3.78
FAS-Y	2	1.25	0	1.22	1.28	0.81
RHO-1	7	5.35	4	5.19	2.98	3.44
PRL30-3	19	18.55	18	18.01	11.51	11.95
TGFB2-5	6	5.18	6	5.03	1.56	3.34

f)

	Pied flycatcher		Atlas flycatcher		Divergence	
Locus	Observed	Expected	Observed	Expected	Observed	Expected
ALDOB-6	3	1.37	0	20.10	1.83	1.45
BRM-12	10	10.15	15	14.91	10.82	10.75
CHDZ	2	0.76	1	1.11	-0.33	0.80
GHR	0	1.65	3	2.39	2.78	1.74
VLDLR	3	6.50	14	9.53	5.91	6.88
ACLY-16	2	0.63	0	0.92	0.11	0.56
ALAS1-8	4	3.63	5	5.33	3.17	3.21
FAS-Y	0	0.00	0	0.00	0.00	0.00
RHO-1	7	5.19	4	7.61	6.39	4.59
PRL30-3	11	11.90	18	17.46	10.88	10.52
TGFB2-5	3	3.22	6	4.72	1.78	2.84