Genomic analyses of migratory divides in the willow warbler

Lundberg, Max

2014

I. Lundberg M, Åkesson S, Bensch S. 2011. Characterization of a divergent chromosome region in the willow warbler Phylloscopus trochilus using avian genomic resources. Journal of Evolutionary Biology. 24: 1241-1243

II. Larson KW, Liedvogel M, Addison B, Kleven O, Laskemoen T, Lifjeld JT, Lundberg M, Åkesson S, Bensch S. 2014. Allelic Variation in a Willow Warbler Genomic Region is associated with Climate Clines. Plos One. 9: e95252

III. Boss J, Liedvogel M, Lundberg M, Olsson P, Reischke N, Naurin S, Åkesson S, Hasselquist D, Wright A, Grahn M, Bensch S. Gene expression in the brain of a migratory songbird during breeding and migration. Submitted.

IV. Lundberg M, Boss J, Canbäck B, Liedvogel M, Larson KW, Grahn M, Åkesson S, Bensch S, Wright A. 2013. Characterisation of a transcriptome to find sequence differences between two differentially migrating subspecies of the willow warbler Phylloscopus trochilus. BMC Genomics. 14: 330

V. Liedvogel M, Larson KW, Lundberg M, Gurzoy A, Wassenaar LI, Hobson KA, Bensch S, Åkesson S. 2014. No evidence for assortative mating within a willow warbler migratory divide. Frontiers in Zoology. 11:52

VI. Lundberg M, Liedvogel M, Larson KW, Grahn M, Wright A, Åkesson S, Bensch A. The genomic landscape of a migratory songbird. Manuscript.

Printed by Media-Tryck, Lund University 2014

MAX LUNDBERG

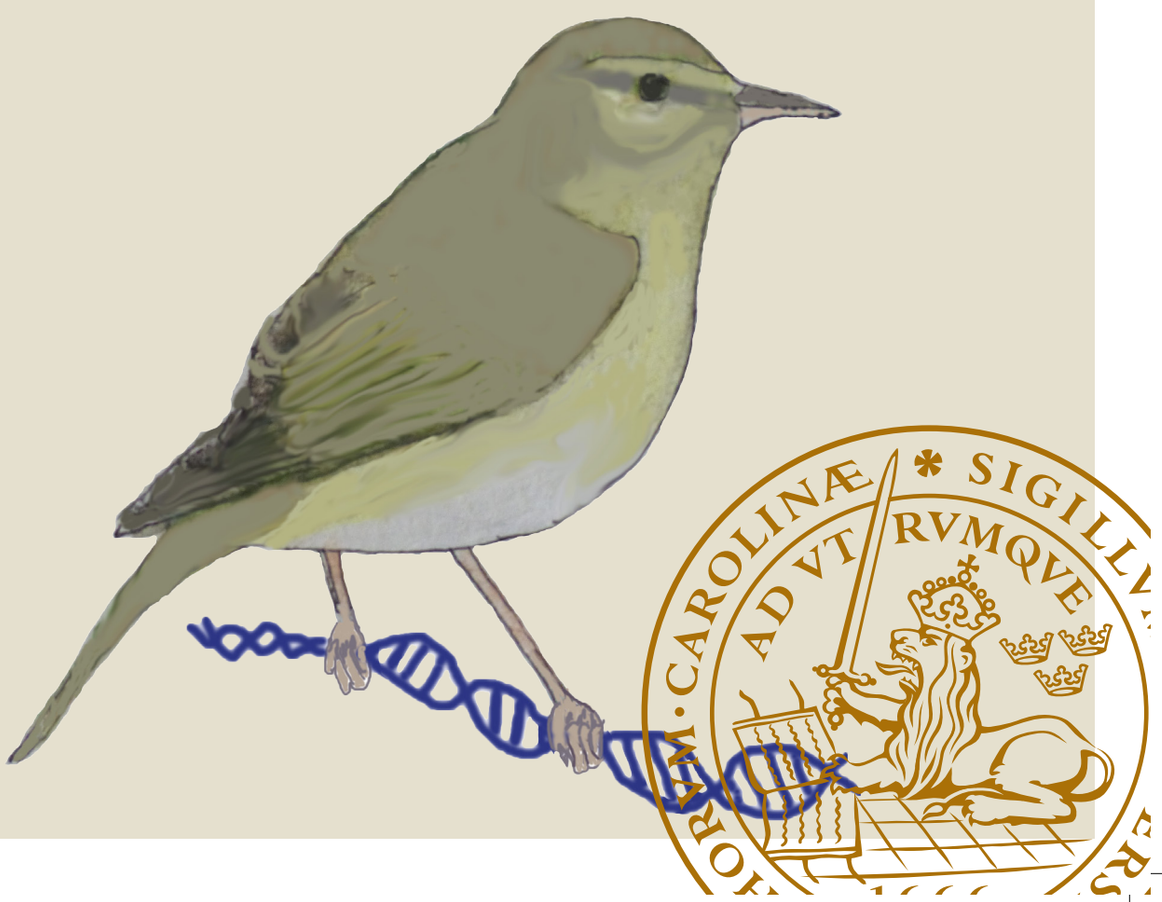
Genomic analyses of migratory divides in the willow warbler

Department of Biology
Faculty of Science

ISBN 978-91-7623-134-0

Genomic analyses of migratory divides in the willow warbler

MAX LUNDBERG
DEPARTMENT OF BIOLOGY | LUND UNIVERSITY



Genomic analyses of migratory divides in the willow warbler

Max Lundberg

LUND
UNIVERSITY

DOCTORAL DISSERTATION
By due permission of the Faculty of Science, Lund University, Sweden. To be defended in the Blue Hall, Ecology Building, Sölvegatan 37, Lund University, on Friday 7th of November 09:30 for the degree of Doctor of Philosophy, Department of Biology.

Faculty opponent

Prof. William A. Cresko
University of Oregon
Institute of Ecology and Evolution

Organization

LUND UNIVERSITY
Department of Biology
MEMEG

Ecology Building
SE – 223 62 LUND, SWEDEN

Author(s) Max Lundberg

Document name

DOCTORAL DISSERTATION

Date of issue

Sponsoring organization

Title and subtitle Genomic analyses of migratory divides

Abstract

In many species of birds juveniles migrate independently of experienced adults and thus have to rely on innate information about migratory routes and wintering area. This information is encoded as a set of inherited migratory directions and a timing schedule that provides information on when and how far to migrate. Apart from these remarkable behavioral adaptations, migration also requires several morphological, physiological and immunological adaptations. Although many migratory traits have been demonstrated to have strong heritable basis, virtually nothing is known about the specific genes underlying them. The lack of known migration genes limits a deeper understanding of the mechanistic processes and evolution of migration. In this thesis, I use the willow warbler *Phylloscopus trochilus* to explore the molecular genetics of migration. The willow warbler occurs with two subspecies in Europe. Ringing recoveries and stable isotopes have demonstrated that the subspecies use different migratory routes and wintering areas. An extremely low genetic differentiation and otherwise similar phenotypes, suggest that most of the genetic differences are likely to be involved in adaptations to the different migratory strategies of the subspecies. Here I use several different high-throughput molecular techniques to identify differences between the subspecies. I find that genetic differences between the willow warbler subspecies are clustered in three divergent regions on different chromosomes, which each span several million base pairs and are comprised of numerous coding genes. The regions show restricted recombination between the warbler subspecies and could be maintained by inversions. A region on chromosome 3 is not specifically associated with the subspecies, but appears to be involved in adaptations to high altitude and latitude environments. The two other chromosome regions contain subspecies-specific variation. In a region on chromosome 1, genetic differentiation peaks close to the gene *FOXO1*, which is important in the formation of fat cells and gluconeogenesis. In the other region, which is located on chromosome 5, there is an enrichment of gene functions associated with fatty acid genes. The functions of these genes suggest that the two chromosome regions could be associated with adaptations to fuelling. Future studies on genetics of migration would benefit from assembled genomes that could be used as a basis for several high-resolution molecular techniques that could detect more localized differences and provide higher resolution of divergent chromosome regions. There is also a need to associate particular genes or chromosome regions with particular phenotypes. This would be aided by efficient and precise high-throughput phenotyping for example by new tracking technologies or controlled behavioral experiments in laboratories.

Key words migratory divide, hybrid zone, willow warbler, *Phylloscopus trochilus*, genomics, transcriptome, bead array, 454 sequencing

Classification system and/or index terms (if any)

Supplementary bibliographical information

Language English

ISSN and key title

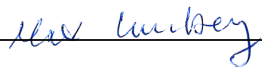
ISBN 978-91-7623-134-0

Recipient's notes

Number of pages 194

Price

Security classification

Signature  Date: 2014-09-30

Genomic analyses of migratory divides in the willow warbler

Max Lundberg

LUND
UNIVERSITY

Faculty of Science, Department of Biology
ISBN (tryck) 978-91-7623-134-0
ISBN (pdf) 978-91-7623-135-7

Printed in Sweden by Media-Tryck, Lund University
Lund 2014

Contents

List of papers

I. Lundberg M, Åkesson S, Bensch S. 2011. Characterization of a divergent chromosome region in the willow warbler *Phylloscopus trochilus* using avian genomic resources. Journal of Evolutionary Biology. 24: 1241-1243

II. Larson KW, Liedvogel M, Addison B, Kleven O, Laskemoen T, Lifjeld JT, Lundberg M, Åkesson S, Bensch S. 2014. Allelic Variation in a Willow Warbler Genomic Region is associated with Climate Clines. Plos One. 9: e95252

III. Boss J, Liedvogel M, Lundberg M, Olsson P, Reischke N, Naurin S, Åkesson S, Hasselquist D, Wright A, Grahn M, Bensch S. Gene expression in the brain of a migratory songbird during breeding and migration. Submitted.

IV. Lundberg M, Boss J, Canbäck B, Liedvogel M, Larson KW, Grahn M, Åkesson S, Bensch S, Wright A. 2013. Characterisation of a transcriptome to find sequence differences between two differentially migrating subspecies of the willow warbler *Phylloscopus trochilus*. BMC Genomics. 14: 330

V. Liedvogel M, Larson KW, Lundberg M, Gurzoy A, Wassenaar LI, Hobson KA, Bensch S, Åkesson S. 2014. No evidence for assortative mating within a willow warbler migratory divide. Frontiers in Zoology. 11:52

VI. Lundberg M, Liedvogel M, Larson KW, Grahn M, Wright A, Åkesson S, Bensch A. The genomic landscape of a migratory songbird. Manuscript.

List of contributions

I. MLu and SB conceived the study. MLu analyzed the data and wrote the manuscript with input from SB and SÅ.

II. KWL, MLi, MLu, SÅ and SB conceived the study. KWL, MLi, MLu, BA, OK, TL, JTL analysed the data or provided samples. KWL wrote the paper with input from the other authors.

III. SB, MG and AW conceived the study. NR, SÅ, SB and KWL collected the samples. SN, DH and SB developed the songbird microarray. JB, MLu, MLi, PO and AW performed analyses of the data. JB wrote the manuscript with input from all co-authors.

IV. SB, MG and AW conceived the study. MLu analyzed the data with the assistance from JB and BC. MLu drafted the manuscript with input from all co-authors.

V. SB, SÅ, KWL and MLu participated in study design. KWL, MLu, MLi and AG carried out field work. MLi, MLu, KWL, LIW and KH performed data analyses. MLi wrote the manuscript with input from all co-authors.

VI. MLu and SB conceived the study. MLu, MLi, KWL, SÅ and SB collected the samples. MLu performed lab work. MLu and MLi carried out analyses of the data. MLu wrote the manuscript with input from all co-authors.

Abstract

In many species of birds juveniles migrate independently of experienced adults and thus have to rely on innate information about migratory routes and wintering area. This information is encoded as a set of inherited migratory directions and a timing schedule that provides information on when and how far to migrate. Apart from these remarkable behavioral adaptations, migration also requires several morphological, physiological and immunological adaptations. Although many migratory traits have been demonstrated to have strong heritable basis, virtually nothing is known about the specific genes underlying them. The lack of known migration genes limits a deeper understanding of the mechanistic processes and evolution of migration. In this thesis, I use the willow warbler *Phylloscopus trochilus* to explore the molecular genetics of migration. The willow warbler occurs with two subspecies in Europe. Ringing recoveries and stable isotopes have demonstrated that the subspecies use different migratory routes and wintering areas. An extremely low genetic differentiation and otherwise similar phenotypes, suggest that most of the genetic differences are likely to be involved in adaptations to the different migratory strategies of the subspecies. Here I use several different high-throughput molecular techniques to identify differences between the subspecies. I find that genetic differences between the willow warbler subspecies are clustered in three divergent regions on different chromosomes, which each span several million base pairs and are comprised of numerous coding genes. The regions show restricted recombination between the warbler subspecies and could be maintained by inversions. A region on chromosome 3 is not specifically associated with the subspecies, but appears to be involved in adaptations to high altitude and latitude environments. The two other chromosome regions contain subspecies-specific variation. In a region on chromosome 1, genetic differentiation peaks close to the gene *FOXO1*, which is important in the formation of fat cells and gluconeogenesis. In the other region, which is located on chromosome 5, there is an enrichment of gene functions associated with fatty acid genes. The functions of these genes suggest that the two chromosome regions could be associated with adaptations to fuelling. Future studies on genetics of migration would benefit from assembled genomes that could be used as a basis for several high-resolution molecular techniques that could detect more localized differences and provide higher resolution of divergent chromosome regions. There is also a need to associate particular

genes or chromosome regions with particular phenotypes. This would be aided by efficient and precise high-throughput phenotyping for example by new tracking technologies or controlled behavioral experiments in laboratories.

Introduction

Bird migration involves some of the most spectacular behavioral adaptations in Nature. An intuitive example is the Eurasian Cuckoo *Cuculus canorus*. Young cuckoos do not interact with their parents (Seel 1977) but still possess the knowledge that they should migrate to sub-Saharan Africa. The strategy of the cuckoo may be extreme, but in songbirds it is a widespread phenomenon that juveniles migrate independently of experienced adults (Newton 2008). More long-lived species, such as white storks *Ciconia ciconia*, travel in flocks, and in this case juveniles could be guided by experienced individuals. However, even in this species, juveniles appear to be equipped with an innate, albeit rough, migratory program (Chernetsov et al. 2004). In addition to these remarkable behavioral adaptations, migration also requires adaptations in morphological, physiological and immunological traits (Winkler and Leisler 1992, Åkesson and Hedenström 2007, Hasselquist 2007).

Most migratory traits could be regarded as quantitative traits. Quantitative traits could be measured on a continuous scale and typically follows a normal distribution that arises from the combined effects of several to many genetic variants and the effect of the environment (Falconer and Mackay 1996). Most of our understanding of the genetics of migration comes from the use of quantitative genetics methods, which generally aim at partitioning the phenotypic variance of a trait into its genetic and environmental components. A particular challenge for studying the genetics of migration is that behavioral traits are difficult to measure in migrating birds. Most studies have overcome this problem by measuring migratory restlessness of birds held in captivity. This restlessness behavior consists of rapid wing fluttering or jumping in the cages and is expressed at night (Berthold 1990a). The timing and directional component of the restlessness behavior have been shown to correspond well to the migratory behavior seen in wild populations (Berthold 1973, Helbig et al. 1989). Quantitative genetic studies on captive birds have demonstrated a strong genetic basis of several migratory traits and suggest that innate migratory information is encoded as a heritable set of preferred directions and a timing program that provides information on when and how far to migrate (Berthold 1990b). Another major finding is that

the propensity to migrate could be described by a threshold model (Pulido 2011). In this case, the decision to migrate or to be resident, depends on whether a continuous migratory propensity variable is above or below a certain threshold.

Migratory direction

The use of an innate migratory direction in young birds was first demonstrated in a large-scale displacement experiment of ringed European Starlings *Sturnus vulgaris* (Perdeck 1958). In this experiment, starlings were caught during migration in the Netherlands, ringed and moved to Switzerland, where they were released. After being released, adults were demonstrated to be able to compensate for the displacement and found their way to their normal wintering area. Juveniles, on the other hand, continued in the same direction as they would have in the Netherlands to reach wintering areas in France and in Southern England and ended up in the wrong wintering area in Southern France and on the Iberian Peninsula. Most information on the genetics of migratory directions has come from studies of captive populations of the blackcap *Sylvia atricapilla*. Helbig (1991) crossbred blackcaps from two differentially migrating populations in Europe and demonstrated that the offspring (F1 generation) expressed intermediate directions relative to the parents (Figure 1). Birds in a second generation of crosses (F2 generation) showed more spread in migratory directions than was seen in the first generation and appeared to have the parental directions segregating (Helbig 1996). The large spread was interpreted as reflecting that migratory direction in blackcaps could be modulated by a small number of genes with large effect.

Timing program

Several physiological processes in birds, such as breeding, migration and moult, are controlled by an endogenous circannual rhythm (Gwinner 1996). This rhythm has a period of approximately one year and arises from molecular oscillations. For a bird not to be physiologically mismatched with the natural year, the clock has to be synchronized with external cues, of which the most important is day length (Gwinner 1996). During migration this rhythm controls the temporal expression of different migratory intensities and directions (Berthold 1990a). Berthold and Querner (1981) showed that large differences in circannual rhythms with regard to the duration and intensity of migratory activity were evident between blackcaps from the Canary Islands and Southern Germany. They further demonstrated that crosses between individuals from each popula-

tion expressed an intermediate duration and intensity of the migratory activity pattern.
Migratory behavior is also regulated on a daily basis by circadian rhythms. These rhythms, which are similar to circannual rhythms, have a period that is approximately 24 h. Many songbirds migrate at night and it has been hypothesized that these birds could have two activity peaks during migration, one during day and one at night, because of splitting two circadian rhythms that are otherwise in complete phase with each other (Bartell and Gwinner 2005).

To migrate or not to migrate

In birds it is common to find closely related species with different migratory behavior (Zink 2011, Rolland et al. 2014). In some species there are also differences in migratory behavior between different populations of the same species and even between individuals of the same population (Newton 2008). The evolutionary potential of migratory behavior has been explained by modelling it as a threshold trait (Pulido 2011). A threshold trait is a special case of a quantitative trait that only has two or a few phenotypic states. The phenotypic state depends on an underlying migration propensity variable that, like other quantitative traits, has a normal distribution in a population. When the propensity of an individual is below the threshold, the individual is resident, and when it is above, the individual is migratory (Figure 1). An evolutionary consequence of threshold traits is that variation for the alternate phenotypic state erodes much slower than for a continuous trait (Roff 1998). Migrating populations in this case are likely to maintain variation associated with residency and vice versa for resident populations (Pulido 2011). This means that a fast evolutionary response is possible when selection would favor the alternate phenotype. The threshold model has been extended to account for a differential sensitivity to environmental conditions of the threshold (Pulido 2011). For example, migration in an obligate migrant such as the willow warbler *Phylloscopus trochilus* appears to be hardwired and largely insensitive to the present environmental conditions. In more facultative migrants, such as the blue tit *Cyanistes caeruleus*, the decision to migrate may be largely dependent on environmental factors. These factors could for example be weather, food availability or social interactions (Chapman et al. 2011).

Other adaptations

Long-distance migratory species tend to have more pointed wings than closely related short-distance migrants and resident species, which is thought to reflect an adaptation to long-distance flight (Winkler and Leisler 1992). Tarka et al. (2010) showed that wing length in the Great Reed Warbler

Acrocephalus arundinaceus has a strong heritable component and were also able to identify a specific chromosome region that was associated with the trait. Long-distance migration also requires efficient fat deposition. An increased fat load is achieved by an increased food intake, hyperphagia, and a more efficient fat assimilation (Bairlein and Gwinner 1994). These processes are regulated by the circannual program but otherwise very little is known about the genetics of fuelling adaptations (Berthold 1990a). Adaptations to a broader range of pathogens may also be important to migrating birds. For example, different sets of pathogens in the wintering area or along the migration route may select for a different immune system

Molecular genetics

Quantitative genetic studies on migration have demonstrated the relative effect of environment and genes for migratory traits, how well they respond to selection and provided some ideas on their rough genetic architecture. However, without knowledge of the underlying genes, it is impossible to get a detailed molecular understanding of migration. For example, to what extent do similar behavioral adaptations in distantly related organisms such as birds and butterflies share the same genetic architecture? What are the effect sizes of migration genes? Are there large-effect variants controlling migratory direction as suggested by studies of the blackcaps?

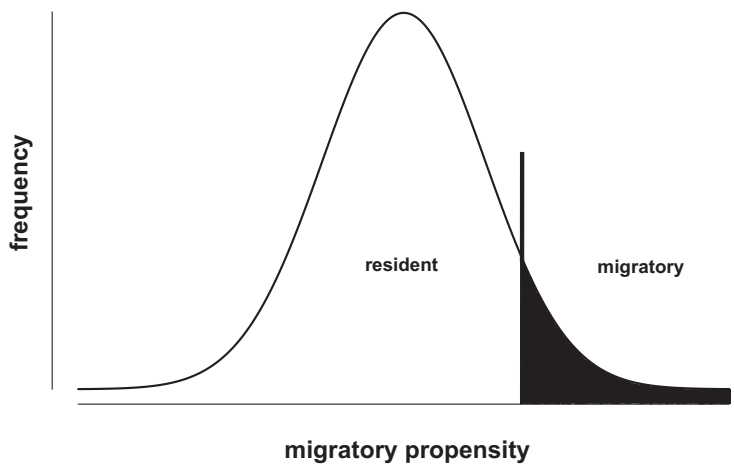


Figure 1 Threshold model of migration. The x axis represent an underlying migratory propensity variable that has a normal distribution within a population. If an individual has migratory propensity below a certain threshold (vertical line), it is resident. If the value is above the threshold the individual is migratory.

There have been very few molecular genetic studies on migration. Mueller et al. (2011) found a significant association between alleles in the *ADCYAP1* gene and the amount of migratory activity both within and between blackcap populations in Europe. However, another study on juncos *Junco spp.* in North America (Peterson et al. 2013) found that the relationship between this polymorphism and migratory behavior is more complicated and cannot be easily generalized across species or populations.

Aims of thesis

The main aim of the thesis was to identify genetic differences between two subspecies of the willow warbler *Phylloscopus trochilus* and relate these differences to the subspecies' different migratory strategies. Another aim was to understand the processes maintaining migratory divides between the subspecies, with a specific emphasis of migratory behavior as an isolating mechanism.

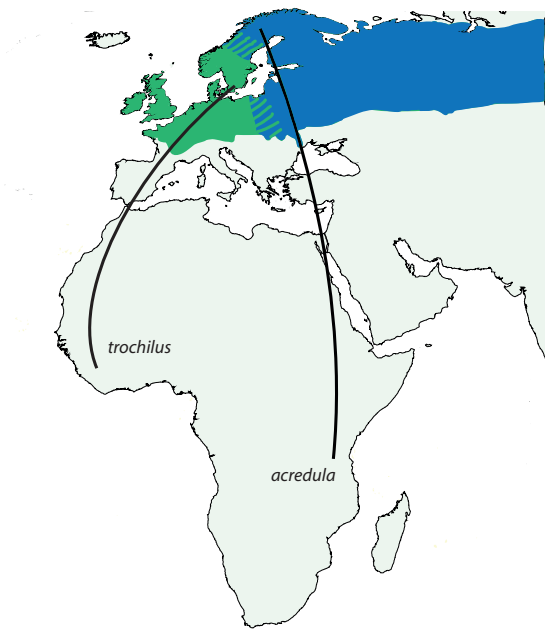


Figure 2 Approximate breeding distribution, migratory routes and wintering areas of the two willow warbler subspecies in Northern Europe. Green is the breeding distribution of *trochilus* and blue the distribution of *acredula*. Striped areas represent hybrid zones.

Methodology

Study species

The willow warbler is a small songbird that is commonly found from Europe to Far Eastern Russia (Cramp 1992). Two subspecies breed in Europe: *P.t.trochilus* in the Western Europe including Southern Scandinavia and *P.t.acredula* in Northern and Eastern Europe. Ring recoveries and analyses of stable isotope data have demonstrated that the two subspecies use different migratory routes and wintering areas in Africa (Figure 2). While ssp. *trochilus* migrates southwest to wintering areas in Western Africa, ssp. *acredula* migrates southeast to wintering areas in Eastern and Southern Africa (Hedenström and Pettersson 1987, Bensch et al. 2006). A third subspecies, *P.t.yakutensis*, breeds in Eastern Asia and winters in Southern Africa (Cramp 1992).

Apart from different migratory behaviors, the subspecies are phenotypically and ecologically very similar (Figure 3). Genetically, the subspecies have been found to be extremely similar (Bensch et al. 1999, Bensch et al. 2002, Bensch et al. 2009). This is likely the result of a very recent divergence with large population sizes that have been little affected by genetic drift. Hybrid zones between the subspecies exist in central Scandinavia and in Poland/Lithuania.

Figure 3 Two extreme color phenotypes in Scandinavian willow warblers. The left bird (grey) is more commonly found in Northern Scandinavia, whereas the right bird (yellow) is typical of Southern Scandinavia.

Field methods

Willow warblers were caught in mist nets by assistance of playback of conspecific song. Once caught, birds were ringed, measured and weighed. Wing feathers were collected for measuring stable isotopes of carbon and nitrogen. These measurements could be compared to those obtained from feathers collected in Africa and be used to infer the wintering grounds of individual birds (Bensch et al. 2006). Blood for DNA extraction was collected through puncture of the brachial vain. For transcriptome-based analyses (**Paper III and IV**), birds were sacrificed through decapitation and had several organs immediately stored in an appropriate buffer. To study pair formation and individual behavior in the field (**Paper V**), we ringed each bird with a unique color combination.

Conserved primer design

In **Paper I**, we made use of the genomes of the chicken *Gallus gallus* and the zebra finch *Taeniopygia guttata*, which at the time were the only available bird genomes. By aligning the genomes, it is possible to detect stretches of sequence, for example in exons (i.e., parts of coding genes), which are conserved between the species. Since the regions are conserved between as distantly related bird species as the zebra finch and the chicken, they are also likely to be conserved in the willow warbler. The regions are therefore suitable for design of conserved primers that could amplify more variable sequence, for example introns, located in between.

Microarray

In **Paper III**, we used a microarray to quantify differences in gene expression between breeding and migrating individuals of both subspecies. The microarray had been designed from zebra finch genes expressed in brain tissue and contained probe sets from an estimated 15,800 genes (Naurin et al. 2008). Each probe set is made up of eleven 25 basepair long nucleotide sequences that are hybridized to complementary DNA. When hybridization occurs, a light signal will be emitted that is proportional to the DNA abundance. From the willow warbler samples, we extracted messenger RNA (mRNA) from whole brain, converted it to a more stable complementary DNA (cDNA) and hybridized the cDNA to the array. In the analyses, we only included signal reports from probes (i.e., exon regions) that showed a very high sequence similarity in an earlier study hybridizing willow warbler DNA to the array.

Transcriptome sequencing

In **Paper IV**, we performed transcriptome sequencing to identify sequence differences between the subspecies. For this purpose we used a subset of the brain tissue-derived samples used for the microarray. As for the microarray, we extracted mRNA and converted it to cDNA. Most cDNA is predicted to originate from a set of relatively few conserved genes that are expressed at a high level in most cells. If cDNA is sequenced without taking this in consideration, most sequence reads will originate from these genes. The cDNA was therefore normalized by removing most of those originating from highly-expressed genes. In order to reduce costs, cDNA from individuals from each subspecies was pooled together before normalization. The cDNA was then fragmented and sequenced using 454 technology. The resulting sequence reads were assembled by mapping them to the genome of the zebra finch. When the sequences had been aligned to the genome we identified single nucleotide polymorphisms (SNPs) and compared how their alleles differed in frequency between the two subspecies pools.

RAD sequencing

As a complement to the transcriptome sequence data we also used Restriction-associated DNA (RAD) sequencing (**Paper VI**) to detect genetic differences between the subspecies. In this method, genomic DNA is randomly digested into fragments onto which adaptors are ligated. Since the adaptor sequence is known, it is possible to design primers that could amplify the fragments. The ends of these amplified fragments are then sequenced with high-throughput Illumina technology. As with the transcriptome sequence data, we used the zebra finch genome to assemble the resulting sequence reads and identified SNPs from the alignments.

SNP array

Both the transcriptome sequencing and the RAD sequencing have been useful to detect genetic variation in the willow warbler. However, in both cases, only a small number of samples were used. To genotype a larger number of samples, we designed a bead array ("SNP chip") from 6000 SNPs discovered in the transcriptome data and to lesser extent in the RAD data (**Paper VI**). Each targeted SNP has a specific probe on the array that will hybridize to the sequence next to the SNP in the genome. Each probe come in two different types that will specifically bind to either of the two alleles of the SNP and result in different fluorescence. By analyzing the fluorescence for a SNP it is therefore

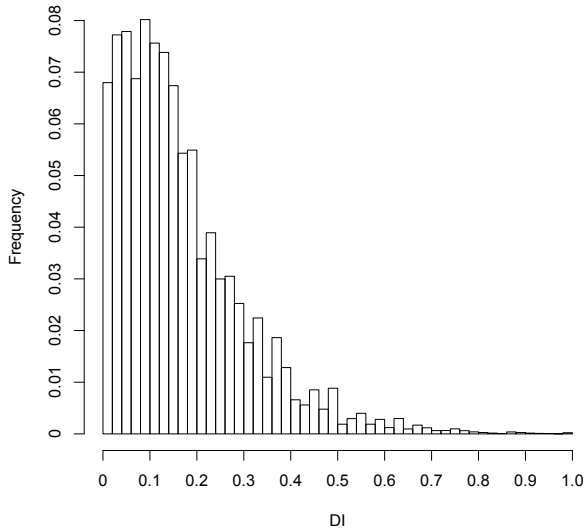


Figure 4 Frequency distribution of genetic differences found between a northern and southern subspecies pool of samples in transcriptome data set. The x axis represents a differentiation index that quantifies the differences in allele frequency between a northern and southern subspecies pool. It ranges from 0 (no differentiation) to 1 (different alleles fixed). For further details see **Paper IV**.

possible to determine whether an individual is heterozygous or homozygous for either of the two alleles.

Results

A chromosome region associated with climatic variables

Bensch et al. (2002) detected an Amplified Length Polymorphism (AFLP) marker, denoted AFLP-ww1, which was highly differentiated between the willow warbler subspecies. A closer inspection of its allelic distribution demonstrated that its two alleles were divided between the Scandinavian mountains and the rest of Fennoscandia, rather than the distribution of the subspecies. In particular, the southern allele of the marker is found in high frequency in Southern Finland where, based on the migratory behavior, we would have expected a high frequency of northern alleles. This marker was found to be located in an intergenic region on chromosome 3 in the zebra finch genome. It was not known whether divergent selection on the specific marker or some nearby gene had caused the observed geographic pattern.

In **Paper I**, we used the genomes of the zebra finch and the chicken to design primers that could amplify and sequence genes in the willow warbler that

were adjacent to the marker. The sequence data revealed that the marker is part of a divergent chromosome region that spans several million basepairs and contains numerous coding genes. Several of the genes showed comparably high genetic differentiation which makes it difficult to associate a specific gene as the target of selection. Although this marker is not associated with migratory differences between the two subspecies, it provides important insights in what genomic footprints that could be expect from divergent selection in general in the willow warbler. The size of the region will have an important impact on the likelihood to identify migration genes that differ between the two subspecies. A large region would require fewer molecular markers in order to be detected, but once identified, it might be difficult to determine which of the genes that are driving the pattern.

Spatial models offer the possibility to explicitly test the association between genetic markers and geographical or environmental factors. In **Paper II**, we genotyped the AFLP-ww1 marker in 2,355 willow warbler samples from 128 sites across Fennoscandia. We first used the frequency of the northern allele as a response variable in a generalized additive model (GAM) with latitude, longitude, their interaction and altitude as predictor variables. We also used a similar GAM, but with several environmental factors as predictor variables. The effect of altitude, latitude and the interaction between latitude and longitude were found to be significant. The final environmental model included as significant predictors maximum temperature in May and potential evapotranspiration (PET), which is the atmosphere's ability to remove water through the transpiration of plants and from soil and water bodies. Together these models show that the northern allele of AFLP-ww1 is associated with alpine breeding in the willow warbler.

Gene expression in the willow warbler

Differences in gene expression could be useful to associate genes with particular phenotypes or environmental conditions. We used a microarray designed from zebra finch genes (Naurin et al. 2008) to study gene expression in the brains of migrating and breeding willow warblers (**Paper III**). We analysed seven migrating males of *trochilus*, seven migrating males of *acredula*, eight breeding males of *trochilus* and four breeding males of *acredula*. The array included in total 22,109 probe sets, of which 11,898 (51.8%) could be reliably and uniquely matched to a total of 6,758 zebra finch genes. Due to the limited sample sizes for some comparisons, we focused on differences between samples pooled by season and by subspecies. In the season comparison 3,045 probe sets were found to be differentially expressed. In the subspecies comparison only 14 probe sets were

found to be differentially expressed, which confirms the low level of divergence found in other genetic studies (Bensch et al. 1999, Bensch et al. 2002, Bensch et al. 2009). Only one gene, *PPP3CA*, was significantly different in both the seasonal and subspecies comparison. For each of the two comparisons, we extracted the 600 annotated genes with most significant expression differences and performed gene ontology (GO) enrichment analysis. For the seasonal comparison there was an enrichment of GO term "calcium ion transport", while for the subspecies comparison there was a significant enrichment of the terms "synaptosome" and "post-synaptic membrane". These enrichments could be associated with the proliferation of neurons and the creation of new memories needed during migration.

Sequencing of a willow warbler transcriptome

Previous studies using microsatellites and AFLP had demonstrated an extreme genetic similarity between the willow warbler subspecies (Bensch et al. 1999, Bensch et al. 2002, Bensch et al. 2009). In order to find more localized genetic differences between the subspecies we sequenced a transcriptome originating from brain tissue. The tissue samples constituted a subset of what was used for the microarray gene expression study and were derived from eight individuals of each subspecies. The samples from each subspecies were pooled together and sequenced using 454 technology. From alignments of the resulting sequence reads to the genome of the zebra finch, we were able to detect 85,000 SNPs. The vast majority of SNPs did not show any difference in allele frequency between the two subspecies-specific pools (Figure 4). An extremely small proportion of SNPs were highly differentiated between the two willow warbler subspecies. These clustered in the earlier identified region on chromosome 3 (**Paper I**) and in two regions on chromosome 1 and 5, respectively. Variation within the two newly identified chromosome regions was demonstrated to be subspecies-specific.

Assortative mating in the willow warbler

In blackcaps assortative mating arises as a consequence of different arrival times between populations (Bearhop et al. 2005). In **Paper V,** we explored if assortative mating exists between the two willow warbler subspecies in the hybrid zone and if it could explain the observed level of hybridization. We studied 40 breeding pairs between 2011 and 2012, and captured an additional 35 males holding territories. The genetic ancestry of samples were determined by using newly developed subspecies-specific markers from **Paper IV**.

For each of the 40 pairs, we compared their genetic ancestry, morphometrics and stable isotope values, which serve as a proxy for wintering area. There was a significant correlation between arrival times of males and their genetic ancestry. However, we found no significant correlation for any of the measured characters between mates. In addition, most individuals were admixed. This suggests that arrival time is likely to play a limited role in modulating the hybridization between the subspecies of willow warbler in our study area.

High-throughput genotyping of willow warblers in Northern Europe

In **Paper VI** we designed a customized bead array ("SNP chip") for high-throughput genotyping of willow warblers. On the array we included 5,843 SNPs detected in the transcriptome sequencing data (**Paper IV**). The remainder of the in total 6,000 SNPs mainly included SNPs that were found to be highly differentiated between the subspecies in a RAD sequencing data set. The SNPs were genotyped in more

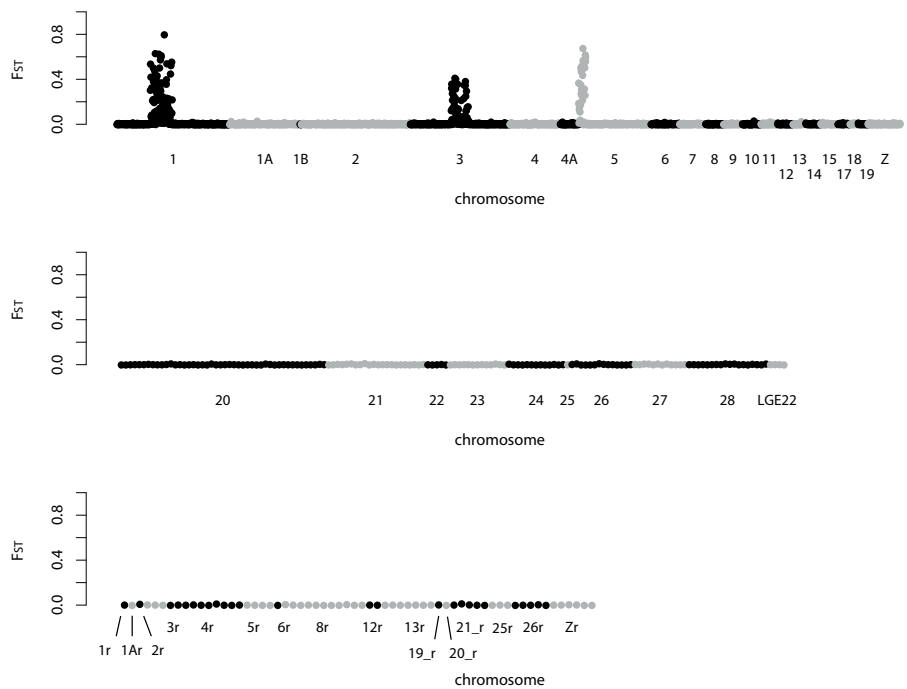


Figure 5 Genetic differentiation (F_{ST}) between *trochilus* and *acredula* across different chromosomes. Random chromosomes are denoted with "r". For further details see **Paper VI**.

than 1100 willow warblers in Fennoscandia. We also included a small number of samples from Scotland, Yekaterinburg (Russia) and from Anadyr in far-east Russia.

Out of an original 6,000 SNPs, 4,094 were retained after quality filtering. The data set confirmed previous findings of extremely similar genomes between the subspecies. It further showed that isolation by distance was minimal even across the most distantly located sites in Europe and East Asia. Genetic differences between the subspecies were limited to the previously identified divergent regions on chromosome 1, 3 and 5 (Figures 5, 6). These regions were larger than previously shown, containing 146, 126 and 62 genes respectively, and showed limited introgression across two hybrid zones. The chromosome regions showed limited recombination between northern and southern alleles and could therefore be maintained by inversions. The most differentiated SNP on chromosome 1 was located close to the *FOXO1* gene, which has an important role in adipogenesis and gluconeogenesis (Nakae et al. 2003, Gross et al. 2008). For chromosome 5 there was an enrichment of gene functions associated with fatty acid desaturase (FAD) genes. This could suggest that the subspecies-specific variation found within the regions on chromosome 1 and 5 are driven by adaptations that has to do with migratory fuelling.

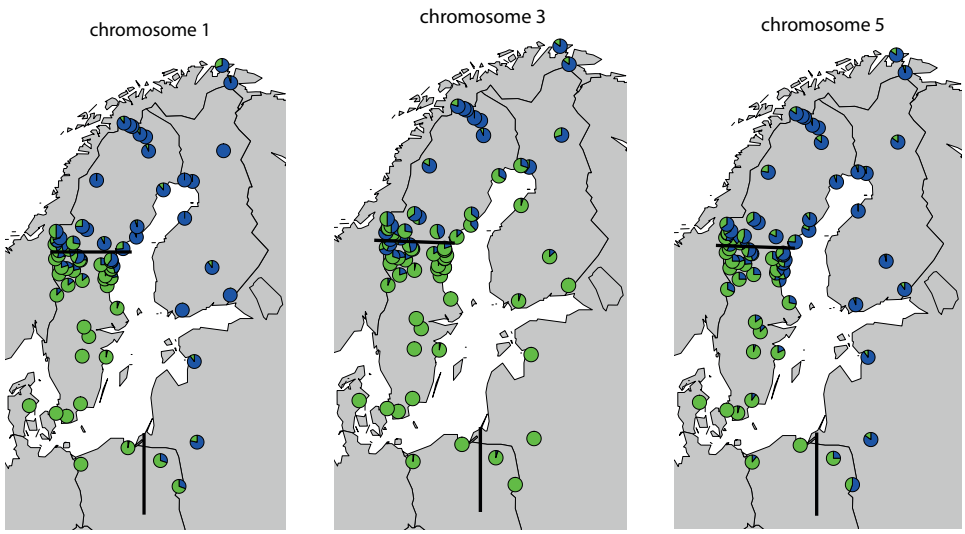


Figure 6 Allele frequency distribution of three divergent chromosome regions in Northern Europe. Green represents southern alleles, while blue represents northern alleles. For further details see **Paper VI**.

Conclusions and future studies

Why is so little known about migration genes? One reason is that none of the traditional model species have expressed clear migratory phenotypes (Liedvogel et al. 2011). In this case there has been no possibility to associate particular genes with a migratory phenotype. In addition, some of the genomes of these species may not even express or contain the genes relevant to migratory species. With the advancement of molecular technology that has occurred over the past few years (Koboldt et al. 2013), this will be less of a problem. Indeed, central to this thesis is the use of modern high-throughput methods to explore the genetic architecture of migration. In **Paper III** we used a large-scale microarray analysis to detect expression differences between birds caught during migration and breeding. In **Paper IV** we used transcriptome sequencing to identify genetic differences potentially associated with migratory traits. Finally, in **Paper VI** we used high-throughput genotyping to explore how these genetic differences are distributed on a larger scale.

An essential step to obtain further understanding of the genetics of migration is having access to a reference genomes for a migratory species. A reference genome enables the use of several high-resolution molecular techniques which could be used to detect localized differences and improve the resolution in divergent chromosome regions. It will be particularly fruitful to combine genetic data sets from different populations and species and look for shared genes and gene networks. The next step is to associate genetic variation with particular migratory phenotypes. These analyses would benefit from efficient and precise phenotyping, for example through modern tracking technology and controlled laboratory experiments.

References

Åkesson, S., and A. Hedenström. 2007. How Migrants Get There: Migratory Performance and Orientation. BioScience **57**:123.

Bairlein, F., and E. Gwinner. 1994. Nutritional mechanisms and temporal control of migratory energy accumulation in birds. Annu Rev Nutr **14**:187-215.

Bartell, P. A., and E. Gwinner. 2005. A separate circadian oscillator controls nocturnal migratory restlessness in the songbird Sylvia borin. J Biol Rhythms **20**:538-549.

Bearhop, S., W. Fiedler, R. W. Furness, S. C. Votier, S. Waldron, J. Newton, G. J. Bowen, P. Berthold, and K. Farnsworth. 2005. Assortative mating as a mechanism for

rapid evolution of a migratory divide. Science **310**:502-504.
Bensch, S., S. Åkesson, and D. E. Irwin. 2002. The use of AFLP to find an informative SNP: genetic differences across a migratory divide in willow warblers. Mol Ecol **11**:2359-2366.
Bensch, S., T. Andersson, and S. Åkesson. 1999. Morphological and molecular variation across a migratory divide in willow warblers Phylloscopus trochilus. Evolution **53**:1925-1935.
Bensch, S., G. Bengtsson, and S. Åkesson. 2006. Patterns of stable isotope signatures in willow warbler Phylloscopus trochilus feathers collected in Africa. Journal of Avian Biology **37**:323-330.
Bensch, S., M. Grahn, N. Müller, L. Gay, and S. Åkesson. 2009. Genetic, morphological, and feather isotope variation of migratory willow warblers show gradual divergence in a ring. Mol Ecol **18**:3087-3096.
Berthold, P. 1973. Relationships between Migratory Restlessness and Migration Distance in six Sylvia species. Ibis **115**:594-599.
Berthold, P. 1990a. Control of Bird Migration. Chapman & Hall, London.
Berthold, P. 1990b. Spatiotemporal programs and the genetics of migration. Experentia **46**:363-371.
Berthold, P., and U. Querner. 1981. Genetic Basis of Migratory Behavior in European Warblers. Science **212**:77-79.
Chapman, B. B., C. Brönmark, J.-Å. Nilsson, and L.-A. Hansson. 2011. The ecology and evolution of partial migration. Oikos **120**:1764-1775.
Chernetsov, N., P. Berthold, and U. Querner. 2004. Migratory orientation of first-year white storks (Ciconia ciconia): inherited information and social interactions. Journal of Experimental Biology **207**:937-943.
Cramp, S., ed. . 1992. The birds of the Western Palearctic. Oxford University Press.
Falconer, D. S., and T. F. Mackay. 1996. Inrtroduction to Quantitative Genetics. 4th edition. Longmans Green, Harlow, Essex, UK.
Gross, D. N., A. P. van den Heuvel, and M. J. Birnbaum. 2008. The role of FoxO in the regulation of metabolism. Oncogene **27**:2320-2336.
Gwinner, E. 1996. Circadian and circannual programmes in avian migration. Journal of Experimental Biology **199**:39-48.

Hasselquist, D. 2007. Comparative immunoecology in birds: hypotheses and tests. Journal of Ornithology **148**:571-582.

Hedenström, A., and J. Pettersson. 1987. Migration Routes and wintering areas of willow warblers. Ornis Fennica **64**:137-143.

Helbig, A. J. 1991. Inheritance of Migratory Direction in a Bird Species - a Cross-Breeding Experiment with Se-Migrating and Sw-Migrating Blackcaps (Sylvia-Atricapilla). Behavioral Ecology and Sociobiology **28**:9-12.

Helbig, A. J. 1996. Genetic basis, mode of inheritance and evolutionary changes of migratory directions in palearctic warblers (Aves: Sylviidae). J Exp Biol **199**:49-55.

Helbig, A. J., P. Berthold, and W. Wiltschko. 1989. Migratory Orientation of Blackcaps (Sylvia a tricapilla): Population-specific Shifts of Direction during the Autumn. Ethology **82**:307-315.

Koboldt, D. C., K. M. Steinberg, D. E. Larson, R. K. Wilson, and E. R. Mardis. 2013. The next-generation sequencing revolution and its impact on genomics. Cell **155**:27-38.

Liedvogel, M., S. Åkesson, and S. Bensch. 2011. The genetics of migration on the move. Trends Ecol Evol **26**:561-569.

Mueller, J. C., F. Pulido, and B. Kempenaers. 2011. Identification of a gene associated with avian migratory behaviour. Proceedings of the Royal Society B: Biological Sciences **278**:2848-2856.

Nakae, J., T. Kitamura, Y. Kitamura, W. H. Biggs, C. K. Arden, and D. Accili. 2003. The Forkhead Transcription Factor Foxo1 Regulates Adipocyte Differentiation. Dev Cell **4**:119-129.

Naurin, S., S. Bensch, B. Hansson, T. Johansson, D. F. Clayton, A. S. Albrekt, V. O. N. S. T, and D. Hasselquist. 2008. TECHNICAL ADVANCES: A microarray for large-scale genomic and transcriptional analyses of the zebra finch (Taeniopygia guttata) and other passerines. Mol Ecol Resour **8**:275-281.

Newton, I. 2008. The Migration Ecology of Birds. Academic Press, London.

Perdeck, A. C. 1958. Two types of orientation in migrating starlings, Sturnus vulgaris L., and chaffinches, Fringilla coelebs, L. as revealed by displacement experiments. Ardea **46**:1-37.

Peterson, M. P., M. Abolins-Abols, J. W. Atwell, R. J. Rice, B. Mila, and E. D. Ketterson. 2013. Variation in candidate genes CLOCK and ADCYAP1 does not consistently predict differences in migratory behavior in the songbird genus Junco [v1; ref status: indexed, http://f1000r.es/11p]. F1000Research **2**:115.

Pulido, F. 2011. Evolutionary genetics of partial migration - the threshold model of migration revis(it)ed. Oikos **120**:1776-1783.

Roff, D. A. 1998. Evolution of threshold traits: the balance between directional selection, drift and mutation. Heredity (Edinb) **80**:25-32.

Rolland, J., F. Jiguet, K. A. Jonsson, F. L. Condamine, and H. Morlon. 2014. Settling down of seasonal migrants promotes bird diversification. Proc Biol Sci **281**:20140473.

Seel, D. C. 1977. Migration of the northwestern European population of the Cuckoo Cuclus canorus, as shown by ringing. Ibis **119**:309-322.

Tarka, M., M. Åkesson, D. Beraldi, J. Hernandez-Sanchez, D. Hasselquist, S. Bensch, and B. Hansson. 2010. A strong quantitative trait locus for wing length on chromosome 2 in a wild population of great reed warblers. Proceedings of the Royal Society B: Biological Sciences **277**:2361-2369.

Winkler, H., and B. Leisler. 1992. On the ecomorphology of migrants. Ibis **134**:21-28.

Zink, R. 2011. The evolution of avian migration. Biological Journal of the Linnean Society **104**:237-250.

Populärvetenskaplig sammanfattning

Fågelflyttning omfattar några av de mest fascinerande beteendemässiga anpassningarna vi finner i naturen. Hos många fågelarter flyttar unga fåglar obereoende av erfarna vuxna individer och kräver således medfödd information om flyttningsvägar, rastning och övervintringsområden. Denna information har visat sig bestå av ett nedärvt program som innehåller flyttningsriktningar och ett tidschema som innehåller information när och hur länge de ska flytta. Utöver dessa beteendemässiga anpassningar finns också flera morfologiska, fysiologiska och immunologiska anpassningar. Trots att flera anpassningar har visat sig vara genetiskt betingade, är idag infomationen mycket bristfällig över vilka gener som ligger till grund för flyttningsbeteendet. Den bristande kännedomen om särskilda flyttningsgener begränsar en detaljerad förståelse av flyttningens mekanismer och evolution. Exempelvis, har liknande anpassningar hos flyttande fjärilar och fåglar samma genetiska uppbyggnad?

I denna avhandling använder jag mig av lövsångaren som ett studiesystem. Lövsångaren förekommer med två olika underarter i Europa: *trochilus* i sydväst och *acredula* i norr och i öst. Ringåterfynd och analyser av stabila isotoper i fjädrar har visat att underarterna använder sig av olika flyttningsvägar och övervintrar i olika delar av Afrika. Medan *trochilus* flyttar sydväst till övervintringsområden i västra Afrika, flyttar *acredula* syd och sydöst till övervintringsområden i östra och södra Afrika. Genetiskt sett är de båda underarterna ytterst lika. Den omfattande likheten kan förväntas göra det lättare att hitta gener under divergerande selektion mellan underarterna, eftersom dessa kommer att vara betydligt mer skilda åt än genomsnittsgenen. Med tanke på att underarterna i övrigt till utseende och ekologi är mycket lika, förväntas merparten av de gener som varierar mellan underarterna att ha att göra med anpassningar till deras olika flyttningsstrategier. I denna avhandling använder jag mig av flera olika storskaliga molekylära metoder för att hitta genetiska skillnader mellan två underarter av lövsångare.

Jag har funnit att genetiska skillnader mellan underarterna är begränsade till regioner på tre olika kromosomer som var och en innefattar många gener. Dessa kromosomregioner uppvisar mycket begränsad rekombination mellan underarterna vilket tyder på att de skulle kunna hållas samman av inversioner. En region på kromosom 3 är inte specifikt kopplad till de olika underarterna utan tycks vara inblandad i anpass-

ningar till häckning i alpina miljöer. De övriga två kromosomregionerna innehåller genetisk variation som är unik till var och en av underarterna. I en av dessa regioner, på kromosom 1, återfinns den största skillnaden mellan underarterna nära genen *FOXO1*. Denna gen har visat sig ha en betydande roll vid bildande av fettceller och glukos. I den andra regionen på kromosom 5 finns ett överskott av gener som är inblandade i regleringen av omättade fettsyror. Detta skulle kunna betyda att de båda kromosomregionerna har att göra med anpassningar till omsättning av fett i samband med flyttningen.

Framtida studier av genetiken hos flyttning skulle gynnas av en fullständig kartläggning av arvsmassan hos flyttande fåglar. Med en kartlagd arvsmassa skulle flera molekylära tekniker med mycket hög upplösning kunna användas. Dessa skulle bidra till att man mer precist kan lokalisera var skillnader i arvsmassan finns och förbättra upplösningen i differentierade kromosomregioner (dvs. kunna identifiera vilka specifika gener som är av betydelse). Ytterligare ett steg vore att koppla samman genetisk variation med särskilda fenotypiska egenskaper. För detta ändamål kommer effektiv och noggrann mätning av flyttfåglars beteende att vara avgörande.

Tack!

Staffan, du är en av anledningarna till att jag fortsatte inom molekylär ekologi. Du har en imponerande bred kunskap inom biologi och tycks ha lösningar till vilka problem som än dyker upp. Genom åren har du alltid varit stöttande och ofta trott mer på mig än vad jag har gjort själv. Du förtjänar ett stort tack!

Susanne, tack för att du har alltid varit stöttande och mån om min utveckling som doktorand och forskare. Ditt stöd har varit betydelsefullt för min forskning.

Dan, min vetenskapliga mentor, tack för alla våra diskussioner. Det var både kul och lärorikt att kunna diskutera de stora vetenskapliga frågorna.

Dennis, tack för att du tog dig tid för att läsa igenom kappan och min halvtidsuppsats, och för visat intresse för min forskning.

Björn, tack för all hjälp med bioinformatiken. Tack vare dig har jag snöat in mig på Perl och Linux.

Miriam. Tack för tiden här i Lund och allt det praktiska inom vetenskap som du har lärt mig. Jag hoppas vi kan fortsätta samarbeta i framtiden.

Keith, you are the quintessential field biologist. Thank you for teaching me your skills and for the time we had in the in the field. It was great fun and I really miss it!

John, tack för tiden i fält och i Stockholm. Du har ett gott humör som lätt smittar av sig på andra.

Mats och **Tony**, tack för lyckat samarbete! Det har varit mycket lärorikt.

Tack **Bengt** för intressanta diskussioner inom genomik och att jag hade möjligheten att delta på den fantastiska workshopen i Sydafrika.

Tack till **Maja**, **Martin S** och **Pablo** för resesällskap på konferenser och worskhops. Maja, tack också för du har stått ut med min frågvishet angående kvantitativ genetik.

Elin V och **Martin A**, tack för att ni har varit mina vänner och för alla de aktiviteter vi har deltagit i utanför jobbet. Det har varit allt från middagar till fågelskådning och frisbeegolf.

Tack **Jane**, för att du har stått ut med mig som kontorskamrat.

Ett stort tack till alla övriga kollegor! Ni är för många för att jag ska kunna tacka er var och en.

Ett stort tack till min familj, pappa **Dick**, mamma **Marketta**, brorsan **Frans** och tvillingsyrran **Jill**, för att ni har stöttat mig under årens gång!

Som vanligt tänkte jag att avsluta med att tacka den som har betytt mest. **Helen**, utan ditt stöd hade inte avhandlingen varit vad den är idag.