Genetic Adaptation and Speciation in Darwin's Finches and Atlantic Herring

FAN HAN

UPPSALA UNIVERSITET

Dissertation presented at Uppsala University to be publicly examined in Room C8:305, BMC, Husargatan 3, Uppsala, Friday, 28 February 2020 at 13:15 for the degree of Doctor of Philosophy (Faculty of Medicine). The examination will be conducted in English. Faculty examiner: Professor Kjetill Sigurd Jakobsen (Centre for Ecological and Evolutionary Synthesis (CEES), Department of Biosciences, University of Oslo, Norway).

Abstract
Han, F. 2020. Genetic Adaptation and Speciation in Darwin's Finches and Atlantic Herring. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine* 1620. 49 pp. Uppsala: Acta Universitatis Upsaliensis. ISBN 978-91-513-0826-5.

Natural selection acts on existing genetic variation to drive genetic adaptation of organisms to various ecological niches. Interaction between closely related populations, through processes such as competition and hybridization, may either lead to their divergence or population fusion, which has consequences for adaptation and the formation of species. This thesis aims to use two natural populations, Darwin's finches and Atlantic herring, as models to explore the genetic mechanisms underlying ecological adaptation and speciation.

The ecological adaptation of Darwin's finches across the Galápagos Islands is primarily reflected by variation in beak morphology. Using whole-genome re-sequencing of all Darwin's finch species, we discover that a locus, *HMGA2*, is highly associated with variation in beak size. Data collected before and after a severe drought show that this locus plays a critical role for ecological character displacement in large ground finches *Geospiza magnirostris* and medium ground finches *G. fortis*.

Genomic islands of divergence refer to genomic regions of elevated divergence when comparing the genomes of closely related taxa. Establishment of these genomic islands can reflect a role in reproductive isolation or be related to ecological adaptation or background selection. Investigating their properties can shed light on how new species evolve. We study the landscape of genomic islands in Darwin's finches, and find that the most pronounced genomic islands are likely ancient balanced polymorphisms, which govern adaptive variation in beak morphology.

Hybridization is increasingly recognized as an important evolutionary process which may lead to speciation. We study two cases of hybridization in Darwin's finches. In the first case, a new lineage of Darwin's finches was founded through hybridization between a resident medium ground finch *G. fortis* and an immigrant Española cactus finch *G. conirostris*. In the second case, female-biased introgression occurred predominantly from medium ground finches *G. fortis* to common cactus finches *G. scandens*. Our genetic analysis on the mosaic genomes of hybrid finches show that non-random mating and natural selection primarily determine the outcome of hybridization.

We generate a chromosome-level assembly of the Atlantic herring with a total size of 726 Mb, which coincides with a high-resolution linkage map and an LD-based recombination map. This facilitates the identification of an ~8Mb inversion, which is likely to be associated with ecological adaptation in herring to differences in water temperature. The contiguity of the assembly sorts placement of loci under selection that were identified based on a previous, highly fragmented draft assembly of the herring genome.

Keywords: Darwin's finches, Atlantic herring, Population genetics, Evolution, Ecological adaptation, Speciation

Fan Han, Department of Medical Biochemistry and Microbiology, Box 582, Uppsala University, SE-75123 Uppsala, Sweden.

To the future

List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.

I Lamichhaney S, **Han F,** Berglund J, Wang C, Sällman Almen, M, Webster MT, Grant BR, Grant PR, Andersson L. (2016). A beak size locus in Darwin's finches facilitated character displacement during a drought. *Science*, 352(6284), 470–474.

II **Han F,** Lamichhaney S, Grant BR, Grant PR, Andersson L, Webster MT. (2017). Gene flow, ancient polymorphism, and ecological adaptation shape the genomic landscape of divergence among Darwin's finches. *Genome Research*, 27(6), 1004–1015.

III Lamichhaney S, **Han F,** Webster MT, Andersson L, Grant BR, Grant PR. (2018). Rapid hybrid speciation in Darwin's finches. *Science*, 359(6372), 224–228.

IV Lamichhaney S*, **Han F*,** Webster MT, Grant BR, Grant PR, Andersson L. (2019). Female-biased gene flow between two species of Darwin's finches. *Submitted manuscript.*

V Pettersson ME, Rochus CM*, **Han F*,** Chen J, Hill J, Wallerman O, Fan G, Hong X, Xu Q, Zhang H, Liu S, Liu X, Haggerty L, Hunt T, Martin FJ, Flicek P, Bunikis I, Folkvord A, Andersson L. (2019). A chromosome-level assembly of the Atlantic herring – detection of a supergene and other signals of selection. *Genome Research*, 29(11), 1919-1928.

* These authors contributed equally

Reprints were made with permission from the respective publishers.

Contents

Abbreviations

ALX1	Aristaless-like homeobox 1
BDMIs	Bateson-Dobzhansky-Muller incompatibilities
bp	Base pair
BMP4	Bone morphogenetic protein 4
CALM1	Calmodulin 1
CNV	Copy number variation
DNA	Deoxyribonucleic acid
DKK3	Dickkopf-related protein 3
F_{ST}	Fixation index
HMGA2	High mobility group AT-hook 2
INDEL	Insertion or deletion
kb	Kilobase
Mb	Megabase
MHC	Major histocompatibility complex
mtDNA	Mitochondrial DNA
MY	Million years
PCA	Principal Component Analysis
ppt	Parts per thousand
SNP	Single nucleotide polymorphism
SV	Structural variant

Introduction

Forms and consequences of genetic variation

"The power of Selection, whether exercised by man, or brought into play under nature, … absolutely depends on the variability of organic beings." Charles Darwin in his book "The variation of animals and plants under domestication" emphasized the fundamental role of variability in natural selection (1). Today we know that the variability refers to genetic variation. Genetic variation exists in the form of mutations, which are essentially nucleotide sequences that vary between the genomes of individuals. The majority of mutations, if they happen under natural circumstances, are selectively neutral or weakly deleterious. This means that most mutations are not harmful or only slightly harmful to the fitness of organism, and will therefore not be acted upon by selection. Other mutations introduce positive or negative effects to fitness and natural selection will favor or purge such mutations, respectively, by changing their frequencies in the population. Defining whether a particular mutation is harmful or beneficial to fitness, however, is meaningless without the context of space and time. A trait that is favored by contemporary conditions might be an impediment to adaptation under other spatial or temporal conditions, and vice versa. A well-documented example is the "industrial melanism" of peppered moths *Biston betularia* in England (2). In order to hide from predators, the majority of these moths have light-colored wings, which camouflage them against the color of the trees they rest on. A minority of them are dark-colored. During the industrial revolution, many of the trees were darkened by carbon-heavy air pollution, which caused most light-colored moths to die off from predation as now contrasted to the trees. This made a black coloration become a predominant trait in peppered moths and thus the number of dark-colored moths substantially increased during that period. When air pollution declined, the frequency of dark-colored phenotype in the population dropped and light-colored phenotype rose again. Therefore, a population that possesses more genetic variation is more versatile to environmental variability and changes.

Genetic variation in a population arises from three sources, spontaneous mutations, gene flow, and genetic recombination. Among these three, spontaneous mutation is the fundamental novel source of all variation. It is the random and permanent change caused by errors during DNA replication, DNA damage or the insertion of transposable elements. Modeling of species

divergence often assumes a constant (clock-like) mutation rate for computational convenience, but it is in fact variable across genome, and even in different types of cells and across different organisms. In the human genome, the average genomic mutation rate is about one error for every 10^8 nucleotides per generation (3) with the male germline having a higher rate than the female (4). To date, the lowest mutation rate among vertebrates was found in Atlantic herring, which is 2.0×10^{-9} per base per generation (5). Mutations that occur in somatic cells can only affect the individual that carries the mutations. They are not inheritable, meaning natural selection has no effect on them and the coming generations. Only mutations that occur in the germline can be passed onto progeny through the reproductive process, and such mutations are the genetic variation that can contribute to the evolution of a population.

Another way of bringing in new genetic variation into a population is through hybridization, which is also referred to as gene flow. When individuals from two separate populations hybridize, their progeny will carry alleles from both populations. In order for this process to occur, several prerequisites need to be met. First, the habitats of the two populations should be in the same or overlapping geographic areas. Second, they cannot have established reproductive isolation, which would render hybrids infertile. Third, the environmental conditions must allow the hybrids to survive and reproduce.

Compared to spontaneous mutation and hybridization, genetic recombination is the fastest way to generate genetic variation. A crossover event during meiosis reshuffles the existing allelic variants on homologous chromosomes to form new haplotype combinations, and transfers the recombinants to the next generation.

Genetic variation is not restricted to a single form in the genome. There are several types of genetic variants on the sequence level according to their sizes and structures, and they sometimes hybrid to shape a chimeric complex. The most common type is the single-nucleotide polymorphism (SNP), where a single base substitution differs between alleles at a specific locus. The density of SNPs is often used to quantify the degree of genetic diversity locally (along a genome) or globally (within a population) (6). It enables comparisons of the heterogeneity across a genome and the magnitude of diversity between populations.

Genetic variation can occur in the form of small insertions and deletions (INDELs), which usually range from several to hundreds of base pairs in length (7–9). They are the second most common form of genetic variations in the human genome (9), but they have received less attention in genetic studies (10). One of the biggest challenges when analyzing INDELs is that the nature of INDELs is highly variable, and most of the current algorithms and statistical models are designed for studies of biallelic mutations. Accurately genotyping INDELs is difficult due to technical limitations (11).

A structural variant (SV) is the type of variation that changes a large section of the genome. The term is usually applied to the genomic changes above 1

kb in size, and some could reach several megabases. The typical structural variant includes copy number variation (CNV), deletion, insertion, inversion, translocation, and complex variant that involves multiple structural alterations (Figure 1). Although this type of variation can change a relatively long stretch of DNA, it is not necessarily related to diseases and can be neutral or near neutral. For example, in a study of 26 human populations, about 13% of the human genome exhibited structural variation (12). The recent advent of long-read sequencing technique, which can read up to several kilobases of nucleotide sequences, has enhanced the study on structural variants. Although recent studies have been able to associate structural variation with selection in livestock and natural populations (13–17), it is still challenging to accurately characterize such variants. SVs, like inversions and translocations, are often surrounded by long stretches of repetitive sequences that even long read sequencing has difficulties to bypass, and they can have variable breakpoints among individuals. Detection of CNVs requires high sequencing depth and a uniformity of coverage across the genome. Besides, the algorithms for SV detection are often inadequate to generate consistent calls; every software introduces individual bias and false positives (18).
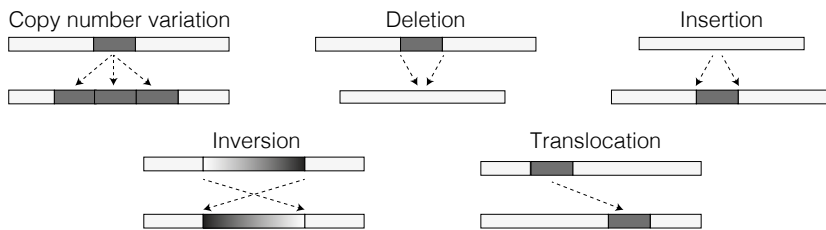

Figure 1. Illustration of example structural variants on genome.

Not all genetic variation has an impact on phenotype. It depends on where in the genome it happens and whether it, by any means, leads to the changes on the expression or translation level. Missense mutations causing phenotypic changes in morphology, physiology, behavior and even psychology, form the focus of genetic studies. Silent mutations are thought to be mainly functionally inactive and selectively neutral, but they may also have an influential function on processes like transcriptional regulation, RNA degradation or non-coding RNA regulation (19–22). In addition to their impact on phenotype, some scientists also argue that these mutations exist to pave the path for later evolutionary innovation by influencing functional mutations via neutral networks (23,24).

Natural selection, adaptation and speciation

Natural selection does not create genetic variation. It only sieves out the individuals that carry beneficial mutations and eliminates variants that are deleterious to fitness. It is unlikely that an individual bears only beneficial alleles; after all, mutations occur by chance and the force of selection is not constant. Natural selection maximizes the mean fitness of a population through changes in allele frequencies (25). Another process, genetic drift, also changes allele frequency of a population. Unlike selection, it arises from stochastic sampling that leads to random shifts in allele frequencies over generations (26). Fitness change from genetic drift is thus directionless. The influence of genetic drift and the efficacy of selection are different in populations of varied sizes. Compared to a small population, the allele frequency of a large population is less likely to fluctuate due to genetic drift, increasing the relative effect of selection. On the other hand, it takes longer time for the selective favored alleles to become common in the population due to the relatively low starting frequency of a new mutation.

As Niles Eldredge said, adaptation is the heart and soul of evolution (27). It represents the process in which natural selection adjusts the genetic composition of an organism in the gene pool, so that the organism can better survive and reproduce in current ecological niche. In most cases, adaptation relates to selection in response to the surrounding ecological environment, such as food sources, predators and climate. Interspecies competition can also lead to shift of traits when multiple species share the similar niches (28–30).

Adaptation is often mentioned together with speciation, because both processes contribute to biodiversity. The essential difference between these two processes is that speciation involves the establishment of reproductive isolation while adaptation does not have to, although adaptation sometimes also directly happens at the loci that are associated with reproductive isolation (31). Adaptation is therefore often a path to speciation. How long genetic adaptation or speciation takes is dependent on the biological characteristics of the organism, the strength of selection, and genetic drift. It could progress over millions of years, but there are also examples where rapid adaptation and/or speciation have occurred in a few generations (32,33).

Speciation is fundamentally a process where genetic exchange is stopped or restricted between populations until complete reproductive isolation is established (34). Roughly, it comprises three phases, although the borderlines between these phases are blurry. Speciation could initialize from the restriction of dispersal, such as geographic movement, human intervention and different mating seasons, which divides individuals into subpopulations and the breeding is restricted to within each subpopulation (Allopatric/Peripatric speciation; Speciation without gene flow; Figure 2) (35). This spatial or temporal separation was previously thought as a prerequisite to speciation. However, recent studies have shown that, speciation can take place spontaneously

in a shared or overlapped niche (Sympatric/Parapatric speciation; Speciation with gene flow; Figure 2) (36,37). Introgressive hybridization, competition for resources, exploitation of unoccupied niches, and sexual selection can all lead to speciation without separating the populations (38,39). A reduction of gene flow arises at the beginning of speciation. Prezygotic isolation may or may not establish at this phase. The second phase of speciation is the accumulation of new genetic variation independently in two subpopulations. At this stage, prezygotic isolation has been established. Interbreeding might still happen occasionally if contact is allowed, but the individuals prefer to mate with the members from their own subpopulations. When the two subpopulations are divergent enough, speciation enters the last phase. That is the formation of the postzygotic isolation. The genomes of the two subpopulations become so reproductively incompatible that their hybrids are unable to develop into a fertile adult. At this stage, speciation is complete.
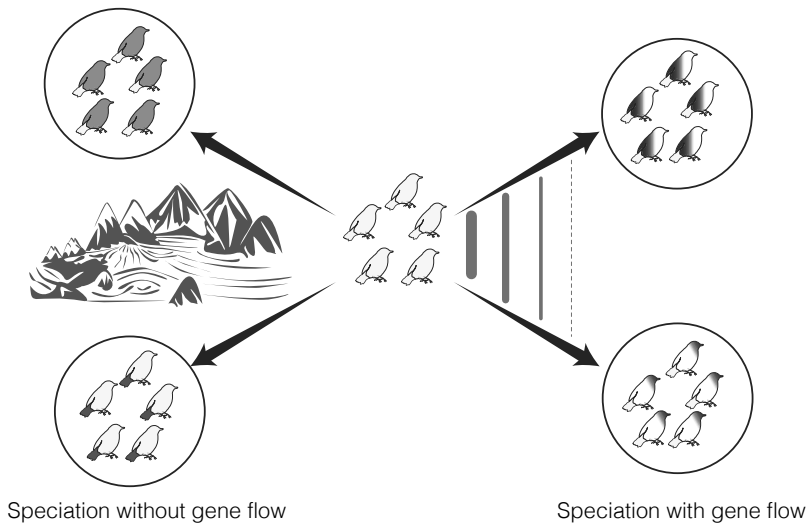

Speciation without gene flow

Speciation with gene flow

Figure 2. Illustration of two modes of speciation with respect of gene flow.

In nature, the entire process of speciation is not always a smooth transition, particularly during the initial and second phases. In some scenarios, populations could fuse into one again by secondary contact, and thus speciation collapses. This is especially likely when the divergence time between the populations is short or gene flow has been ongoing throughout the period. In other scenarios, incomplete reproductive isolation could be reinforced because the hybrids are less fit than parental individuals (40). Moreover, in some extreme cases, like the changes of ploidy in plants, where the number of chromosomes in progeny is different from their parental genomes, can rapidly result in speciation (41). Whether speciation will proceed to completion heavily depends

on the strength of selection, the magnitude of gene flow, and the intrinsic characteristics of the organisms.

Gene flow: Obstacle or aid to speciation?

Gene flow introduces genetic variation from one population to another before the establishment of reproductive isolation. It first generates F1 where each parent contributes 50% of its genome through hybridization. The F1 hybrids backcross with individuals from one, or both of the parental populations. The progeny of the onward generations then repeats this pattern. This process creates generations of mosaic hybrid genomes which comprises genetic materials from both parental populations. Meanwhile, it counteracts the loss of genetic diversity, increases effective population size (N_e), and reduces the genetic differentiation between the parental populations (Figure 3, left).

Under this mode, speciation and gene flow sound like the opposite ends of the rope in a tug of war; one promotes heterogeneity and the other promotes homogeneity. However, natural selection and genetic drift make this complex. Hypothetically, the hybrids possess phenotypes intermediate between parental phenotypes, or more similar to one of them, depending on how they breed and which loci are introgressed or retained. Most of the time, these phenotypes are not favored by selection because the parental populations may be better adapted when competing with the hybrids for food, mates and other resources (42). In addition, diverging populations accumulate genetic incompatibilities, called Bateson-Dobzhansky-Muller incompatibilities (BDMIs), by which hybrids suffer from a decline in fitness (43–45). Hybrids have lower survival and reproduction rate compared to their parental populations and sometimes even display phenotypic abnormalities. The parental populations may have diverged for a certain period and the genetic variations they have accumulated separately are fixed by selection in each population. When these diverged fixed alleles are combined into the hybrid genome, the epistatic interaction between multiple loci can introduce negative effects on fitness and the new combination of alleles can be incompatible. Therefore, hybrids are not always favored by selection, and gene flow in such conditions can reinforce the speciation of the incipient populations (Figure 3, middle).
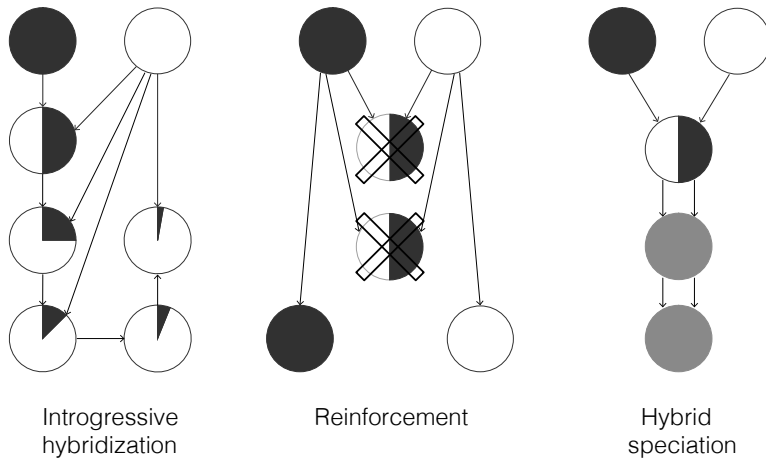
Figure 3. Consequences of gene flow concerning speciation.

Under particular circumstances, hybrids can outcompete their parental populations. In order for this to happen, there must be environmental resources that favor the hybrid lineage. For instance, unexpected ecological changes can provide opportunities that either separate hybrids from the incipient populations or create new habitat for hybrids. Some extreme phenotypes of hybrids may facilitate their exploitation of unoccupied niches or increase their competitiveness, which is referred to as hybrid vigor (46–48). These conditions are typically unlikely, but not impossible, and they lead to inbreeding becoming more common than backcrossing with parental populations in hybrids. Once the reproductive isolation is built up, hybrid speciation has completed. Hybrid speciation is more commonly observed in plants than in animals due to their differences in sex determination mechanisms and nature of genomes (49). Two forms of hybrid speciation exist: homoploid hybrid speciation, where no increase in ploidy happens via hybridization, and polyploid hybrid speciation, where hybrids carry more sets of chromosomes than parents. Both homoploid and polyploid speciation have been well documented in plants while only few cases of homoploid speciation have been recorded in animals (50,51). In hybrid speciation, gene flow facilitates the birth of a new lineage and thus it provides possibilities for speciation (Figure 3, right).

The stage of speciation at which gene flow continues to occur is also critical to the outcome. Accumulation of mutations and adaptation to independent niches takes time. Before the pre-divergence period (Figure 4, left), gene flow can easily break down isolation, especially when populations share the same niche. This is the most fragile phase of speciation. Prezygotic barriers can prevent the gene flow, but not necessarily. Gene flow could also occur after the populations have diverged for a certain time (Figure 4, middle), which can occur in secondary contact, if postzygotic barriers have not formed. In this

15

scenario, how reproductively isolated the populations are determines their future. Sometimes, gene flow can recur throughout the process of speciation (Figure 4, right), but genetic differentiation between the populations still exist. This is particularly common in species like aquatic organisms, because they are not restricted by geographical isolation. Such condition gives the populations more flexibility to either diverge or fuse into one, which is, nevertheless, still dependent on the force of selection.
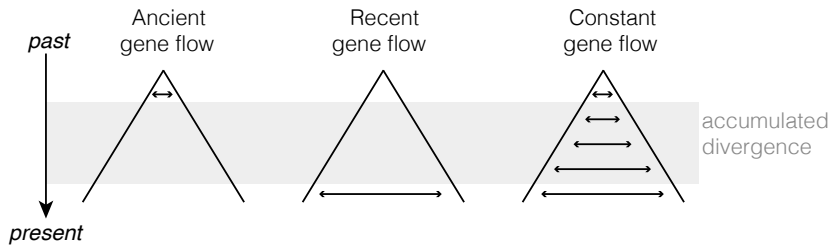

Figure 4. Modes of gene flow during speciation.

To sum it up, gene flow can lead to three consequences in respect to speciation (Figure 3). (1) If the hybrids merge with either or both incipient populations, the speciation will be prevented and the divergent populations may be fused into one, depending on frequency of the gene flow. (2) If selection is against the hybrids, then divergence of the incipient populations may be reinforced. (3) If selection strongly favors the hybrids, it can promote the speciation of the hybrid population. Any of these chains of events could transit to another or be terminated halfway through, and gene flow can be either obstacle or aid in the process.

Darwin's finches: A classic model of adaptive radiation

About two million years ago, a small number of birds arrived at the Galápagos archipelago. When Charles Darwin landed on the Galápagos Islands during his voyage of the Beagle in 1831, these birds had diversified into more than 10 different species. However, the significance of the finches was not realized until the taxonomist John Gould examined the specimens collected by Darwin. Later in Darwin's book, he explicitly described the details about their nearly perfect gradation of morphological diversification, primarily in the form of beak, which helped to further develop his theory of natural selection (52). To credit their inspiration for Darwin's idea, David Lack named them Darwin's finches (53), and they are well known as a textbook example for the study of evolutionary biology.

Although there is a lack of fossils, geographic proximity combined with molecular analysis indicates that the ancestor of Darwin's finches was most likely a member of the grassquit genus *Tiaris*, either from South America or the Caribbean (54,55). The polymorphism of MHC class II genes in modern finches indicates that the number of birds that first colonized Galápagos was at least 30 (56). The Galápagos islands were born from fire; they are the products of volcanic eruptions that initially happened about 5 million years ago (57). Since their formation, forces such as plate movements, fluctuating sea level and climate changes have periodically altered the number, height-profile and shape of the islands, giving the islands possibilities to develop into different ecoregions. Therefore, Galápagos islands are highly diversified in ecology, ranging from arid ground to humid forest. Adaptation to the different ecological conditions has led to a diversity in the finches' biological features, including body size, wing length, beak type and song. One of the classic morphological characteristics is their beak. The beak morphology reflects the food resource they feed on, which is a result of long-term selection. For instance, species that catch insects have sharp and slender beaks, those that eat seeds on the lowland have large and blunt beaks, and those that probe into cactuses have long and pointed beaks (Figure 5).
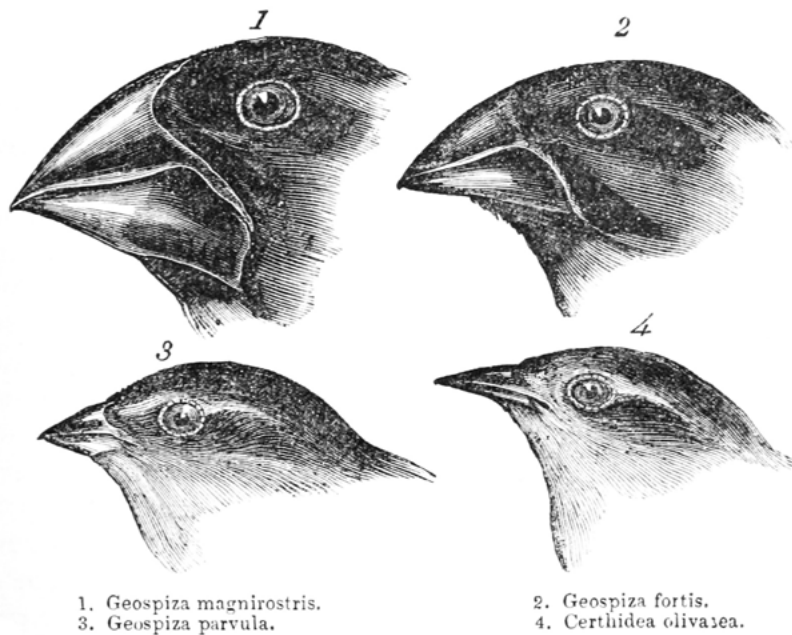

1. Geospiza magnirostris.   2. Geospiza fortis.
3. Geospiza parvula.   4. Certhidea olivacea.

Figure 5. Examples of beak variation in Darwin's finches. @ darwin-online.org.uk

What genetic changes facilitated this rapid evolution is important question to understand how this monophyletic group radiated into so many species.

During the last decades, field observation and molecular studies on microsatellite DNA, mtDNA and whole-genome DNA have revealed the taxonomy of extant species of Darwin's finches (15,58–60). Till today, 18 species in total have been recognized, 17 of which are distributed on different islands of the Galápagos and one on Cocos Island. The molecular dating indicated that these species evolved from a common ancestor about 1.5 million year ago (15). Deciphering the genetic code for beak variation is of course the first step to understand their adaptive radiation. A critical locus, *ALX1*, has been identified to be significantly associated with variation in beak shape (15). Expression analyses identified several candidate genes, such as *BMP4*, *CALM1*, and *DKK3*, that affect beak growth during embryonic development (61–63). Additionally, hybridization was rarely observed (64), but genomic data revealed that interspecies hybridization is extensive in the adaptive radiation of the finches (15).

Atlantic herring: A new model for genetic studies of ecological adaptation

The Atlantic herring (*Clupea harengus*) is a pelagic species, which inhabits both sides of the Atlantic Ocean, including the Baltic Sea. They often move in large schools near to the coast, and one single school can consist of billions of individuals with a density of 1.5-16 fish/m^3 (65). As one of the most abundant vertebrates on the planet, herring has historically been a crucial food resource in North Europe. In 2015, it was ranked as the world fourth largest commercial fishery (66), and the top species caught in the North East Atlantic (Figure 6).
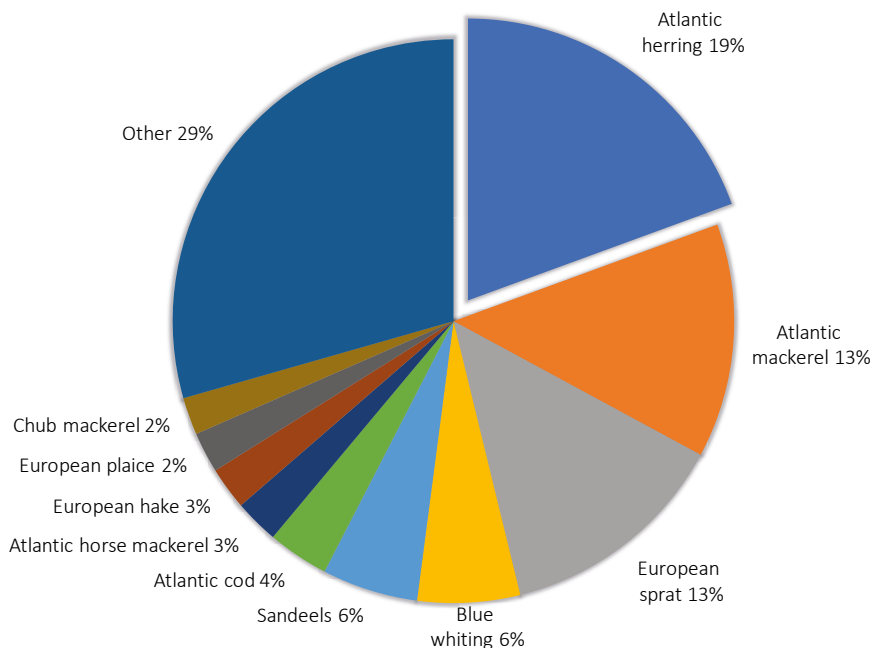
Figure 6. Top species caught in the North East Atlantic in 2015. (Data source: Euro-stat)

Herring often migrate thousands of kilometers across the open ocean for feed-ing and spawning. They primarily feed on plankton, krill, and small fish, and meanwhile they are the prey of large fish, marine mammals and seabirds. The population structure of the Atlantic herring is complex. The population dy-namics is primarily maintained by natal homing behavior, where adult fish migrate back to their birthplace to spawn (67,68). Populations from the same location may spawn in different seasons. Thus, they can be further classified into discrete spawning types. Herring, having a maximum age of 25 years, can thus spawn up to 20 times. At spawn, each female can release from 30,000 to 70,000 eggs depending on the populations and seasons (69,70). Unlike most marine fish, which only occupy marine water with above 30 ppt salinity, her-ring can survive and reproduce throughout the brackish Baltic Sea, where the salinity is in the range of 3-12 ppt. Compared to the Atlantic herring, Baltic herring are smaller in size and contain less fat.

Herring shows a considerable plasticity and adaptability to the environment and a large number of populations have been classified based on their differ-ences in morphology, life history, spawning location and spawning season (67). It has been challenging to accurately distinguish herring populations. Traditional methods measured scales and otoliths to differentiate stocks (71–73). But these methods are affected by the way of measuring, fish age, fishing season and locality. Previous studies using limited molecular markers have

revealed only small genetic differentiation between the populations (74–76). The prevalence of modern sequencing technology now enables high through-put screening of genetic markers, allowing herring to display its advantage as a model in genetic studies. Whole-genome screens showed the level of genetic differentiation in the herring is extremely low at selectively neutral loci (77), due to inter-population gene flow and/or minute genetic drift. This gives prominence to the genetic signatures that have been under natural selection. Hundreds of genomic loci have been identified associated with variations be-tween populations adapted to the Baltic Sea and the Atlantic Ocean as well as the loci associated with the adaptation to different spawning seasons (77). A pedigree study revealed that a low mutation rate may contribute to the moder-ate nucleotide diversity considering the huge population size (5). All these features make herring an excellent model species for genetic studies of eco-logical adaptation.

Summary of the papers

Paper I: A major locus associated with beak size in Darwin's finches

The evolutionary history of Darwin's finches on the Galápagos Islands has been widely observed and well documented in the last decades. Their most pronounced phenotypic variation is their highly diverged beak morphology. After the rapid radiation about 1 million year ago, the beaks of the descendent species developed into different shapes and sizes, depending on the local environment and food sources of the islands. Some species that live on distant islands but where the niches are similar carry similar beak morphology. An example is sharp-beaked ground finches *Geospiza acutirostris* (formerly *G. difficilis* on Genovesa), which is very similar to small ground finches *G. fuliginosa* on Española in beak and body size (78). However, this similarity can be minimized or lost when these species meet in the same geographical area. This phenomenon is called ecological character displacement (79,80), which is caused by competition between species. When resources are limited, the species with overlapping traits adjust their respective niches so that they can coexist in the same geographic area. The theory was proposed as an important model of speciation, but evidence for this mechanism has rarely been demonstrated in natural populations (81). Peter and Rosemary Grant observed a character displacement event in Darwin's finches (82). In late 1982, a novel breeding population was established on Daphne Island by a few immigrant large ground finches *G. magnirostris*. Their diet of large shelled seeds overlapped with the local medium ground finches *G. fortis*. A severe drought in 2004 made food scarce on the island, during which time the large ground finches outcompeted the relatively large-beaked medium ground finches for large seeds and thus caused lower survival among the medium ground finches. Instead, other medium ground finches with small beaks that could feed on small seeds had higher survival. As a consequence, a dramatic shift of beak size was observed in the medium ground finches during 2004-2005. In this study, we carried out a whole-genome study of 60 Darwin's finches and investigated the genetic basis underlying the event.
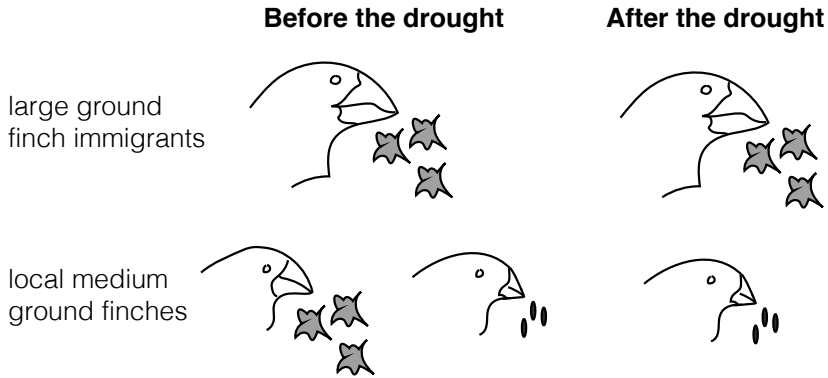
Before the drought   **After the drought**

large ground
finch immigrants

local medium
ground finches

Figure 7. A character displacement occurred between large ground finches (*G. magnirostris*) and medium ground finches (*G. fortis*) during a drought. The large ground finches were three times more efficient than the medium ground finches with large beaks when cracking the hard shell of the seeds.

Like other studies of selection on size-related traits, beak size showed a strong correlation with body size in Darwin's finches. A combination of multiple regression and selection differential analysis on the birds in the 2004-2005 drought event indicated that the association between survival and beak size was much stronger than between survival and body size (82). Taking these into account, we sequenced 10 birds from each of the six species of Darwin's finches that primarily differ in size-related traits: the small (*G. fuliginosa*), medium (*G. fortis*) and large (*G. magnirostris*) ground finches, and the small (*Camarhynchus parvulus*), medium (*C. pauper*) and large (*C. psittacula*) tree finches (Paper I: Table 1). The sequence of each individual reached ~10 X genome coverage. All 2 x 125 bp paired-end reads were aligned to the reference genome GeoFor_1.0 (83), and variant calling was performed based on stringent criteria (Paper I: supplementary materials).

  To minimize the influence of population structure, we grouped different species of similar size and performed pairwise genome-wide screens by comparing the groups in different sizes. After intersecting the outliers from each contrast, we identified seven independent genomic regions related to beak size (Paper I: Fig. 2A). Among them, one ~525 kb locus exhibited the most significant differentiation in all three contrasts. This region contains four genes, one of which is the high mobility AT-hook 2 gene (*HMGA2*). HMGA2 is an important protein for growth and development. According to previous studies, loss-of-function mutations in *HMGA2* lead to dwarfism in mice and rabbits (84,85). In human, *HMGA2* has been associated with variation in height and craniofacial distances (86). Therefore, we believe this locus is very likely to be a candidate gene associated with variation in beak size among Darwin's finches, and we refer to this entire 525 kb region as the *HMGA2* locus.

To validate whether and how much *HMGA2* locus affects beak morphology, we combined the data with sequences from 120 birds of Darwin's finches, and investigated the haplotypes in total 180 individuals. A maximum-likelihood tree was constructed based on all the SNPs from the 525 kb *HMGA2* locus, and it clearly showed that all the species except the outgroup species can be divided into two major haplogroups (paper I: Fig. 1D): One group included 98% of the small individuals with body weight less than 16 g, and the other group included 82% of the large individuals with body weight greater than 17 g. We dated the split between the two haplogroups and suggested that it occurred about 1 million years ago, which was before the radiation of Darwin's finches. The genotypes of the 17 conserved SNPs that were highly divergent in our contrasts showed that small birds were fixed for one haplotype (*S*) and large birds were fixed for the other haplotype (*L*). The birds with intermediate size showed varied degrees of segregation at this locus.

Because body and beak size are strongly correlated, the *HMGA2* locus could potentially affect either body size or beak size, or both. To investigate this, we genotyped a diagnostic SNP from *HMGA2* locus in another 133 medium ground finches, and examined the correlations between genotype and phenotypic measurements. The results suggested that this locus is highly associated with beak size rather than body size or beak shape (Paper I: Fig. 2E).

We next asked how much the *HMGA2* locus contributed to the shift of beak type in medium ground finches in the character displacement event in the drought during 2004-2005. We genotyped 71 medium ground finches that experienced the event, and summarized the survival rate of different genotypes. Individuals homozygous for the *S* haplotype have a significantly higher survival rate (73.7%) than the ones homozygous for the *L* haplotype (30.0%), which confirmed that *HMGA2* locus contributed to the dramatic change of the beak morphology in the short period.

This study identified a major locus contributing to the beak size in Darwin's finches and explained a considerable part of the genetic changes underlying the rapid shift of beak morphology in medium ground finches during the drought in 2004-2005.

Paper II: Genomic landscape of divergence among Darwin's finches

Comparing the genomes of two closely related species often reveals a heterogeneous landscape of genomic divergence; most of the genomes display a consistent degree of divergence, while some regions exhibit exceptionally high divergence, resembling volcanic islands in the ocean. At first, these regions of elevated divergence were thought to be related to reproductive isolation that contribute to the initial stages of speciation. Thus they were called "genomic

islands of speciation" (87). With further studies having carried out in various species, scientists revealed that the locally elevated divergence was not necessarily related to speciation. Other evolutionary processes, such as background selection or selective sweeps coupled with variation in recombination rate, could cause a very similar pattern. The statistical methods used for measuring the divergence also produce uncertainties in determining whether these islands emerge due to speciation. The name "genomic islands of divergence" is therefore used more often when describing these regions. In this study, we aimed to differentiate the models that can result in different patterns of genomic islands of divergence.

Traditionally, relative measures of divergence, such as F_{ST}, were widely used to identify genomic islands of divergence. These measures, however, can be inflated if either of the compared populations has low intra-population diversity (88). Thus, F_{ST} peaks may not truly reflect differentiation between populations. An absolute measure of divergence, d_{XY}, was recommended because it is less affected by within-population variation.

Based on the measures described above, we discussed how the landscape of divergence is affected in four different evolutionary circumstances. In model 1 where gene exchange is allowed for incipient species, the genomic islands are predicted to be resistant to gene flow because they contain genes favoring local adaptation and/or speciation (Figure 8A). We therefore expect to observe both F_{ST} and d_{XY} elevated at genomic islands. In model 2, a similar pattern could also be expected without gene flow, which is the result of ancestral polymorphism (Figure 8B). If highly diverged haplotypes have existed before speciation at a locus, the process of incomplete lineage sorting can maintain the high divergence in the descendant species. In model 3 where speciation occurs in allopatry, no gene exchange is possible due to geographic barriers (Figure 8C). Each of the populations independently adapts to the local environment, and genomic islands arise at the loci that are involved in ecological adaptation. In this scenario, only F_{ST} is expected to increase because recent selection does not influence absolute divergence. In model 4, the genomic islands are caused by recurrent selective sweeps and/or background selection regardless of gene flow (Figure 8D). We expect to observe elevated F_{ST} but reduced d_{XY} at such genomic islands because both processes can reduce the level of local genetic variation. This pattern is often expected at regions of low recombination rate.
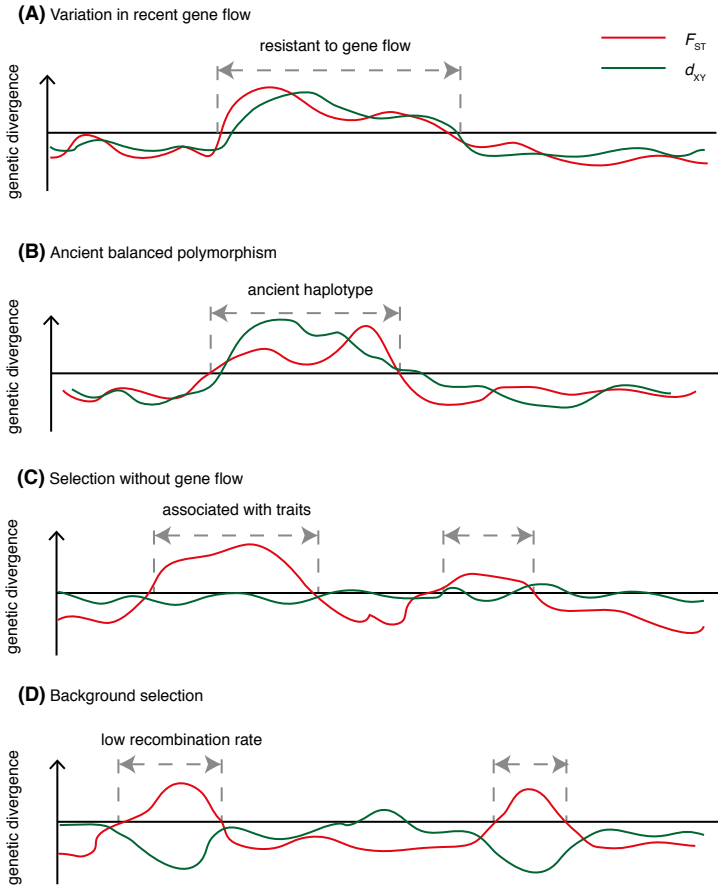
(A) Variation in recent gene flow

(B) Ancient balanced polymorphism

(C) Selection without gene flow

(D) Background selection

Figure 8. Evolutionary causes of genomic islands of divergence and the expected patterns along the genome.

For this study, we used sequencing data of 11 populations of Darwin's finches from six islands of the Galápagos archipelago and one population from Cocos Island (Paper II: Figure 1A; Table 1). According to the geographic distributions and beak morphology, we made independent contrast in 12 different species pairs and scanned the genome of each pair using 50 kb windows. In each window, we estimated genetic parameters including nucleotide diversity, Tajima's D, number of fixed differences, F_{ST} and d_{XY}. In all the species pairs, we identified genomic islands that showed elevated F_{ST} ($ZF_{ST} >= 4$) (Paper II: Figure 2A). We discovered that in all sympatric and allopatric species pairs, absolute divergence d_{XY} was elevated (Paper II: Figure 3A) and the numbers

of genomic islands were similar (Paper II: Figure 2B). This indicated that recent gene flow was not a major factor in the formation of the genomic islands.

Among the most pronounced genomic islands, two contained the *ALX1* and *HMGA2* loci, which are associated with beak shape and size, respectively. The absolute divergence d_{XY} were elevated at these loci in species pairs with different beak morphology (Paper II: Figure 3B). From previous studies, we know that the haplotypes at both loci formed before the radiation of Darwin's finches. This suggests that some of the genomic islands were involved in ecological adaptation in the early stage of speciation, and they were formed from ancient haplotypes.

Furthermore, we investigated how recombination rate behaved in the genomic islands. We utilized the population-scaled recombination map in zebra finch (89), and sorted the medium ground finch reference genome based on pairwise alignment between the two assemblies. Interestingly, the average recombination rates in the genomic islands of divergence were significantly lower than the genomic background in all species pairs, including the regions containing the *ALX1* and *HMGA2* loci (Paper II: Figure 4), indicating that the genomic islands involved in ecological adaptation or speciation tend to be related to low recombination rates.

In this study, we identified genomic islands through 12 pairwise comparisons of Darwin's finches, and discovered that two of the most pronounced genomic islands contained *ALX1* and *HMGA2* loci, which are associated with variation in beak shape and beak size, respectively. We summarized four models of genomic islands caused by different evolutionary processes, and compared them with the observation in Darwin's finches. We concluded that genomic islands of divergence in Darwin's finches were shaped by gene flow, ancient polymorphism, and ecological adaptation.

Paper III: Witnessing hybrid speciation in Darwin's finches

In 1981, when Peter and Rosemary Grant were performing field studies on Daphne Island, an immigrant bird caught their attention (90). It resembled the local medium ground finch (*G. fortis*), except that it was 70% heavier and sang a unique song. They named it 5110 and recorded the trajectory of this male ground finch on Daphne Island (Figure 9). It first mated with a local resident medium ground finch and one of their offspring bred with another female medium ground finch. From generation 2 onwards, all the descendants only mated within this lineage. Backcrossing with parental populations was not observed any more. Individuals of generation 4 – 6 were all derived from a sister-brother mating in generation 3. The fitness within the lineage was not severely

compromised by the close inbreeding, and the lineage behaved like a new species compared to the local residents. Because members of the new lineage are relatively large in size, they were referred to as the Big Bird lineage. In this work, we sequenced the parental generation as well as the members of the Big Bird lineage, with a main aim of identifying the source species of the founder male 5110 and searching for genetic clues to the success of the lineage.
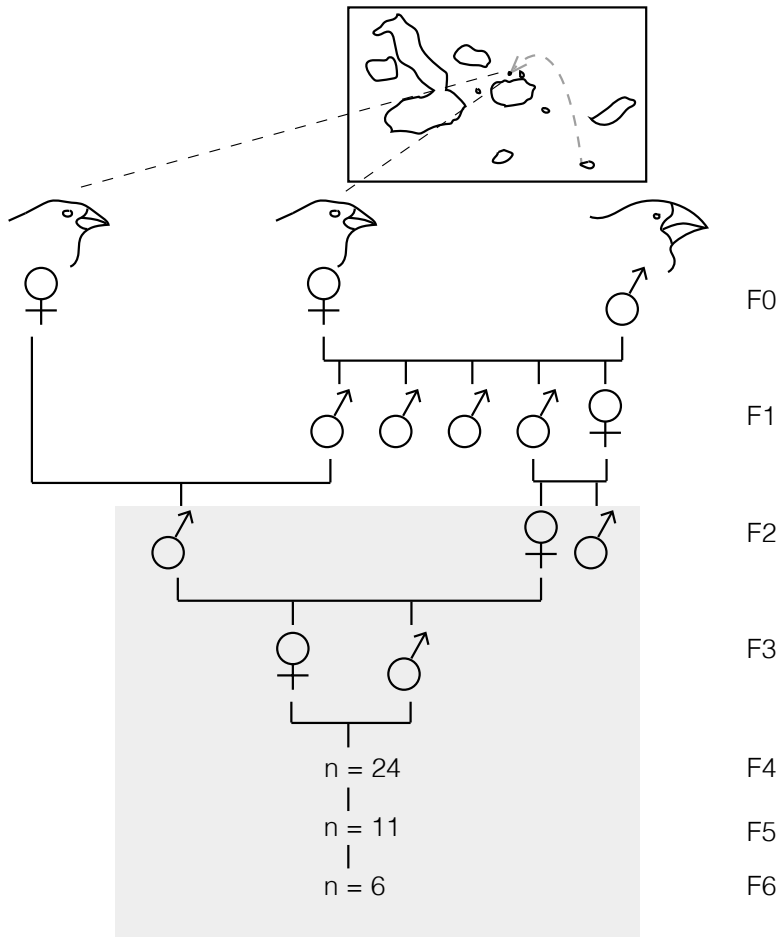

Figure 9. Generations of the Big Bird lineage. The top map depicts the resident island Daphne Island of medium ground finches *G. fortis* and migrant route of the foreign Española cactus finch *G. conirostris*. From the second and onward generations, the hybrids stopped outbreeding or backcrossing.

Morphological measurement showed that 5110 was not a purebred member of any of the species residing on Daphne Island. This determined 5110 was an immigrant or a hybrid. It was believed to be a medium ground finch *G. fortis* × common cactus finch *G. scandens* hybrid from the neighboring island

according to microsatellite assignment tests (91). However, our whole-genome phylogeny clearly placed 5110 in the cluster of *G. conirostris* (Paper III: Fig. 2A), which is surprising as *G. conirostris* only resides on Española and its satellite Gardner, which are more than 100 km away from Daphne Island (Paper III: Fig. 2B). To explore if 5110 was a hybrid, we performed ADMIX-TURE analysis, which classified 5110 as a pure *G. conirostris* (Paper III: Fig. 2C). Thus, we concluded that the immigrant 5110 was a *G. conirostris* from Española/Gardner.

We next studied morphological and genetic characteristics of the Big Bird lineage and their beaks. The average body and beak size of the Big Bird lineage are intermediate between the parental species (Paper III: Fig. 3A). A detailed look found that their body weight is closer to *G. fortis* but their beak size is closer to *G. conirostris*. This suggested natural selection may have occurred. A more indicative evidence of selection is that the beak depth showed a continuously increasing trend across six generations (Paper III: Fig. 3B). We investigated the *ALX1* and *HMGA2* loci, which contribute to beak shape and size respectively, in members of the Big Bird. The members from generation 3 onwards were homozygous for *ALX1* blunt alleles, with two sub-haplotypes that segregated among them (Paper III: fig. S5; Table S5). At the *HMGA2* locus, the allele frequency of the large allele was 60.8% in generation 4-6 (Paper III: fig. S4; Table S5). The haplotypes at these two loci matched the phenotypic measurement.

By witnessing and studying this unique hybridization event, we found that beak type and song are two key factors that determine reproductive isolation in Darwin's finches. Moreover, we learned that the initial steps that may result in speciation in nature could happen much faster than previously thought.

Paper IV: Female-biased hybridization in Darwin's finches

Introgressive hybridization could, on one hand, lead to the birth of a new species, as the case in Paper III. On the other hand, it could fuse two distinct populations into one, or blur the boundary of genetic isolation. How the mosaic genome of hybridization develops under selection, to a great degree, determines the consequence of the hybridization. In this paper, we performed a genetic analysis on the mosaic genomes resulting from the interbreeding of medium ground finches *G. fortis* and common cactus finches *G. scandens* on Daphne Island. Combining genomic data with the long-term field observation and phenotypic changes, we aimed to understand the formation of the hybrid genomes and how this process is affected by patterns of breeding, natural selection, and recombination.

Previous research documented that medium ground finch *G. fortis* and common cactus finch *G. scandens* rarely but recurrently hybridize at least since 1976 on Daphne Island. In order to compare the effect of introgression, we split birds into those born before (1975 – 1983; defined as early group) and after (1998 – 2011; defined as late group) hybridization occurred. We selected samples from each period and further separated them into groups that can best represent the beak variation based on Principal Component Analysis (PCA) of beak measurements (Paper IV: Table 1). Each group comprised 30 individuals, and was sequenced as a pool. We identified 10.3 million SNPs from autosomes and 0.5 million from Z-linked scaffolds among the pools.

Phylogenetic trees based on autosomal and Z-linked markers showed consistency for the placement of all groups except the *G. scandens* late blunt group. The Z chromosome tree placed *G. scandens* late blunt group close to the other two *G. scandens* groups, but the autosomal tree placed it near the midpoint of the internal branch between *G. scandens* and *G. fortis*. This indicated extensive gene flow from *G. fortis* to *G. scandens* on autosomes but not on the Z chromosome. No clear sign of introgression was observed from *G. scandens* to *G. fortis*. The phylogeny based on mtDNA confirmed the asymmetric gene exchange between *G. scandens* and *G. fortis*, and indicated that the introgression mainly involved *G. fortis* females mating with *G. scandens* males, which is corroborated by field observations.

A window-based tree screening was performed across the genome in order to identify genomic regions under introgression. As much as 33.8% of the autosomal regions showed topologies that indicated introgression from *G. fortis* to *G. scandens* while only 1.7% of the Z-linked regions supported such topology. This implies that selection may suppress gene flow on Z chromosome, but the field observations suggest a different explanation: The hybrid females preferentially mated with *G. scandens* males, because the hybrid males were at a disadvantage in competition with the parental *G. scandens* population in respect of body size (Figure 10).
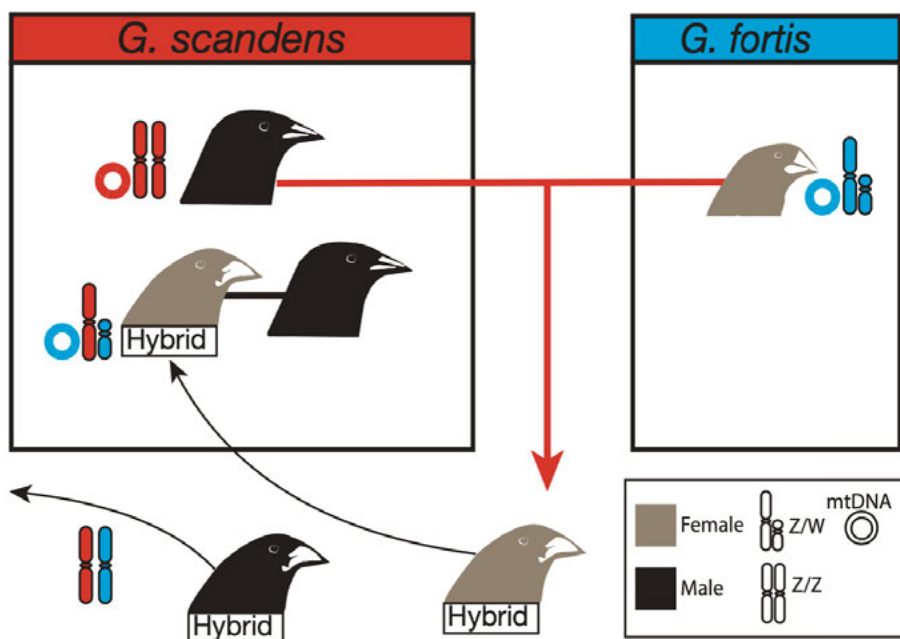
Figure 10. Illustration of female-biased introgression from *G. fortis* to *G. scandens*. The hybrid males were purged because they were disadvantageous in body size for mating preference when competing with the parental *G. scandens* (Graph credit: Erik Enbody)

If genetic incompatibilities have formed between two hybridizing populations, one expects to observe a higher recombination rate at the regions showing more introgression. High recombination can more rapidly break down the linkage between incompatible and compatible genes, and thus expose smaller blocks of deleterious genes to selective elimination. We examined this hypothesis in our model and found that recombination rate was not or only weakly correlated with the magnitude of introgression. The lack of the expected association could be due to the short divergence time and intermittent hybridization of *G. fortis* and *G. scandens*. As a consequence, no or only a few genomic regions showing genetic incompatibilities may exist between *G. fortis* and *G. scandens*. Furthermore, our study period only involved six to seven generations, which may not be sufficient to observe the effect of purifying selection.

By inspecting the pairwise genetic differentiation among the six groups, we identified two loci that are resistant to gene flow, one of which being the *HMGA2* locus with its well-characterized association with beak size. The other locus is a ~4 Mb region on chromosome 5, which is very likely to be under selection during introgression.

In this study, we showed another consequence of introgressive hybridization. Non-random mating and natural selection together contributed to the heterogeneous landscape of gene flow in the mosaic genome of the hybrids. Only

time will show whether it will lead to fusion of *G. fortis* and *G. scandens* into a single population or not.

Paper V: A chromosome-level assembly of Atlantic herring

Population studies based on the draft assembly of the Atlantic herring (v1.2) revealed extremely low differentiation at genetically neutral sites. This substantially enhances the ability to characterize genomic loci under selection. Several hundreds of loci have been identified that are associated with adaptation to different salinities and spawning seasons among herring populations (77). The number of independent selective loci, however, was difficult to estimate, because the draft assembly comprises more than 140,000 scaffolds. The fragmented assembly also hindered study of the interplay between adjacent genes under selection as well as the detection of potential structural variations. In this paper, we generated a chromosome-level genome assembly of the Atlantic herring by combining long-read sequencing and Hi-C techniques, and revisited the genetic signals associated with ecological adaptation.

High molecular weight genomic DNA was extracted from an Atlantic herring for generating long subreads using PacBio as well as a Hi-C contact map (Figure 11). These two datasets were combined to a primary assembly, followed by a series of scaffolding, polished and error-correction, which eventually delivered a refined version 2 herring assembly of 726 Mb. Comparing to the draft assembly, this assembly shows an integrated structure of herring genome, which comprises 26 chromosomes. This linear order of the genomic segments is confirmed by the presence of 26 linkage groups in the comprehensive linkage map we constructed from two-full sib families.

Extract DNA of
high molecular weight

PacBio initial
assembly graph

FALCON-Unzip
assembly

primary contig

haplotigs

Scaffolding with
Hi-C reads

Chr 1
Chr 2

Short-reads polishing
and duplicates removal

Chr 1
Chr 2
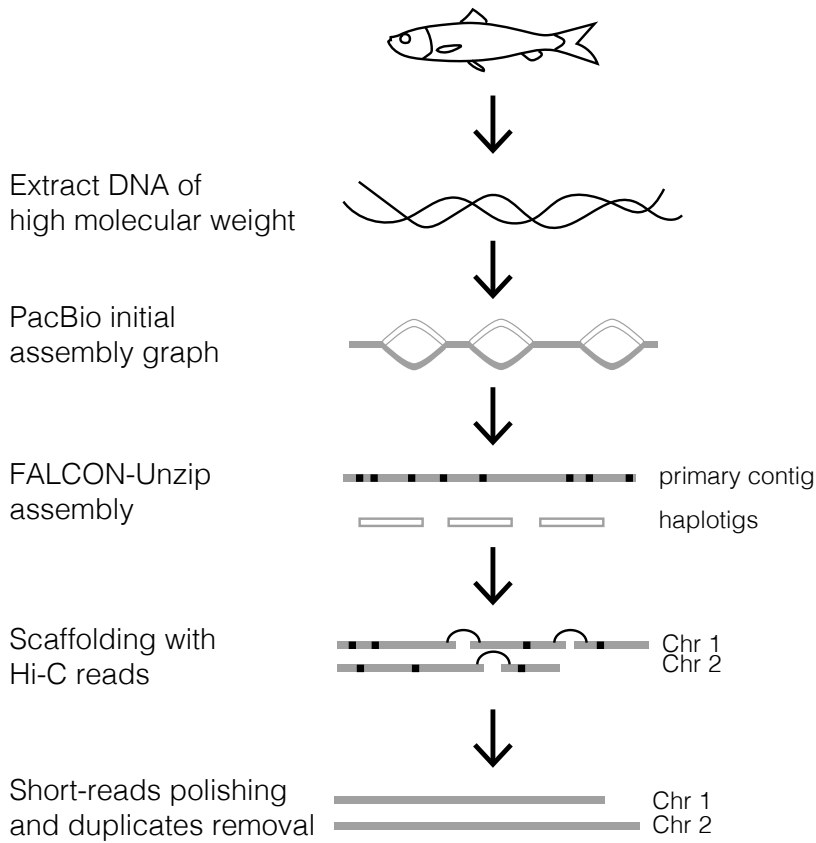
Figure 11. An overview of assembly pipeline.

The chromosome-level assembly enabled us to investigate recombination variation across the genome using population data. We used SNP data from 14 Baltic herring that were individually sequenced and inferred a fine-scale recombination map with LDhat program (92). The result shows a mean recombination rate of 2.1 cM/Mb, which is comparable to the rate in zebrafish (1.6 cM/Mb). The majority of the chromosomes display an exceptionally high recombination rate at one end over the other, which form an L-shape profile. This is in agreement with meiotic recombination data observed in pedigree-based linkage analysis.

We reordered the previously identified genetic signals that are associated with ecological adaptation along the v2 assembly, and estimated the minimum number of independent loci contributing to adaptation. We identified 125 loci associated with adaptation to salinity and 22 loci associated with adaptation to spawning seasons. By taking advantage of the new assembly and linkage map, we additionally identified a 7.8 Mb inversion on chromosome 12. By

examining the mismatch pattern of short reads aligned close to the ends of the inversion block, we identified a putative breakpoint, and found that the inverted repeats at the ends most likely facilitated the occurrence of the inversion. The genotype data show two major haplotypes (S and N) at the inversion, and the haplotype frequencies of the individuals are in accordance with their geographic distribution. That is, populations from the West Atlantic, the East Atlantic and the Baltic Sea, that spawn most northerly, have a very high frequency of the N haplotype, while populations spawning more southerly carry more S haplotypes. This trend indicates that this inversion may be maintained by selection related to temperature at spawning.

Concluding remarks and discussion

In this thesis, I uncovered genetic variations associated with ecological adaptation and examined several scenarios of hybridization. Bioinformatic analysis based on whole-genome sequencing was the primary strategy and two organisms, Darwin's finches and Atlantic herring, were used as models in the population studies. In Paper I, we discovered a major locus associated with variation in beak size, and showed how it facilitated an ecological character displacement event in large and medium ground finches. In Paper II, we explored genomic islands of divergence in multiple species pairs of Darwin's finches and related their formation to gene flow, ancient polymorphism, and ecological adaptation. Paper III described the hybrid speciation of a new finch lineage and investigated the underlying genetic causes of its formation. Paper IV studied female-biased introgressive hybridization between two closely related finch species. In Paper V, we delivered a chromosome-level assembly of the Atlantic herring accompanied by a high-resolution linkage map and an LD-based recombination map. We demonstrated the importance of a fine-structured reference genome to the interpretation of selective signatures. In the following section, I attempt to generalize these findings towards the fundamental questions of contemporary interest in evolutionary biology.

Speciation in Darwin's finches

Universality

Since the oceanic islands are remote from human activities and lack of the relatively complexity compared to continental environment, all evolutionary processes in the Darwin's finches happened largely in an undisturbed environment. This serves as an ideal laboratory for observing natural selection. Like many other organisms, speciation in Darwin's finches started with restricted space and limited food source. Rewinding the clock to the time when the first group of birds arrived at Galápagos, the islands were covered by various flora but sparse fauna. The ancestors of Darwin's finches exploited the unoccupied territories as the first colonists. According to their divergence time, 18 extant species can be separated into early and late sets. The warbler finch was the earliest diverging group and the sharp-beaked finch and

vegetarian finch diverged after a long time (about 900 MY). Together with the Cocos finch, they constitute the early set of species of the radiation. The late set is the relatively recent diverged tree finches and ground finches. The species of the early set differ from the late set in three ways: only one species in each genus, they are morphologically different from the late set and from each other, and speciation appears to have progressed at a low rate. The first two differences could be explained by the limitation of the habitats and food source on Galápagos during the early phase, because there were fewer plants and arthropods in the arid environment. For the low speciation rate in the early set, there are two possible explanations: on the one hand, the fluctuating climate changes allow the formation or adaptation of plants, which sped up the speciation of Darwin's finches in the late set; on the other hand, if we assume a steady speciation rate, the time gap between the species in the early set could be caused by the extinction of early-formed species. This acceleration of speciation-extinction rate is, nevertheless, pervasive in modern avian species, especially within young and temperate radiations (93).

The time it takes for two populations to generate genetic incompatibility primarily relies on the intrinsic characteristics of the populations and the strength of selection. Price and Bouvier (94) proposed the post-zygotic incompatibility begins to build up after populations have separated for 2-3 MY. A study in sunfish indicated it would take 10 - 15 MY before fitness of hybrids is below parental species (95). Prager and Wilson (96) inferred that the potential of producing hybrids decays 10 times slower in birds than in mammals. Considering the estimated divergence time and observations in Darwin's finches, these species are still at a very early stage of the establishment of genetic incompatibilities. Ground finch species are observed interbreeding on both Genovesa and Daphne Island, rarely but persistently. Interbreeding was also documented between tree finches (90).

In addition to genomic incompatibility, mate choice can also play a role in whether two populations interbreed or not. In principle, hybridization occurs if choosing-sex is indiscriminate or when two populations have similar courtship signals and response. If the members of a species prefer mating with the individuals sharing morphological traits that are genetically inherited, like plumage, body size, and beak size, the divergence of these traits caused by adaptive selection could reduce the frequency of interbreeding. Interbreeding could occasionally happen when the traits of two populations are overlapping. For example, the largest individual of small ground finches (*G. fuliginosa*) could breed with the smallest individual of medium ground finches (*G. fortis*) since their body sizes are closer than to the smallest ones in their own populations (90). However, if mate choice is dependent on culturally inherited traits, like song type, interbreeding could happen by learning, or mis-imprinting. Song learning happens in a short sensitive period in young finches, which can be vulnerable to interruption. This means a bird could copy the song from another population. As illustrated in Paper II and Paper IV, song determines

the fate of hybridizations in Darwin's finches. Hybridization due to mating preference could break down the speciation process, especially when genetic incompatibility is weak and unstable.

Uniqueness

What is special about Darwin's finches? First, the ancestral finches must have possessed a high potential of diversification. This is indicated by the radiation of Darwin's finch relative species in the Caribbean, whose evolution of beak diversity is also rapid and extensive (59). Second, their feeding habits are flexible and learnable. During their juvenile development, they have a period of trial-and-error learning for new feeding behaviors, sometimes even with tools. This might not be a heritable ability, but it provides possibility for the afterward adaptive selection to enhance these behaviors. Third, the occurrence of sympatry facilitates two species to hybridize, which effectively generates new variations at numerous loci.

Another Galápagos species, the mockingbird may have arrived at Galápagos about at the same time as Darwin's finches (97). They have dispersed over the archipelago, but only evolved into four allopatric species. One reason for the difference between mockingbirds and Darwin's finches could be that mockingbirds have a generalized food resource, and species of Darwin's finches are more constrained to specific food types. Additionally, mockingbirds learn songs throughout life without song differentiation among populations. Both food choice and song learning reduce the potential of diversification in this species. The similar arrival time indicated that a long time since colonization cannot solely guarantee an adaptive radiation.

David Lack said, "the peculiarities of Darwin's finches are primarily due to an unusual combination of geographical and ecological factors." (78). The many islands in the archipelago likely helped Darwin's finches to establish the radiation, but it is not a strict necessity. Madagascan vangas, for instance, are a group of endemic species that radiated on only a single island but also possess highly diversified body size and beak morphology. This bird family originated from a common ancestor about 25 million years ago and after two peaks of rapid radiation it now contains 22 species (98,99). In addition to the many individual and ecologically variable islands, fluctuation of oceanic ecology plays a more important role in the radiation of Darwin's finches. Among the first colonists on the Galápagos, they had the priority to exploit their niches, and they had sufficient time to adapt to different habitats without competitors and predators. The number and size of islands in Galápagos have varied, due to changes in sea levels and volcanic activity, and the ecological habitats have changed as a function of climate. This allows the directions and magnitudes of selection change with time, and accordingly shapes the diversified species.

Hybridization and speciation

Paper III and IV displayed two opposite outcomes of hybridization in Darwin's finches. In Paper III, we observed an immigrant Española cactus finch *G. conirostris* hybridized with a resident medium ground finch *G. fortis*. Apart from one male in the first generation that backcrossed with another medium ground finch, all the descendants of the hybrids kept breeding within the lineage without an apparent reduction in fitness. If the environment continues favoring this new lineage, it has the potential to eventually become a new species. In Paper IV, we recorded another case. Two closely related species, medium ground finches *G. fortis* and common cactus finches *G. scandens*, hybridize rarely, but continuously. A strong El Niño event triggered the breeding of F1 hybrids and the frequent backcrossing in the next 21 years. The backcrossing was bidirectional but more predominantly into *G. scandens*. If this merging process keeps going, these two species may fuse into one, which will bear a unique combination of the genetic materials from the contemporary *G. fortis* and *G. scandens*.

Both these examples of hybridization take place on Daphne Island, and the divergence time of the focal species are comparable. The hybridization case from Paper III may represent the initial stage of a rapid homoploid hybrid speciation, which is rarely documented in animals (100–102). The success of the Big Bird lineage is most likely due to their large beak and body size, as well as unique song. Their beak morphology is advantageous for feeding in competition with other species on the island while their song promotes inbreeding within this lineage. In principle, the Big Bird should carry more genetical contributions from *G. fortis* than *G. conirostris*, as two founder *G. fortis* females contributed to the pedigree but only one *G. conirostris* male. Their beak size along with beak bluntness should be highly correlated with body size like other species of Darwin's finches. Unexpectedly, the morphology of the Big Bird resembles *G. fortis* more in body size but resemble *G. conirostris* more in beak size. The beak shape of the Big Bird increasingly changes towards bluntness across generations but body size remains unchanged. This is most likely to be promoted by natural selection. The male Big Birds mimic the unusual song of the immigrant *G. conirostris* which cannot be recognized by the resident females. The combination of these phenotypes gives the Big Birds a morphology and a song beyond the scope of their parental populations. Thus, it provides the Big Bird an unprecedented opportunity to exploit unoccupied niches on this small island and populate by mating within the lineage.

In contrast, the gene flow between *G. fortis* and *G. scandens* on Daphne Island is the result of introgressive hybridization. Hybridization between two non-sister species, after which the hybrids predominantly backcrossed with one parental population, has yielded unequal exchange of autosomal, Z chromosomal and mitochondrial DNA. This asymmetric hybridization arises from the disadvantage of the hybrid males in body size when competing with the

parental population, as well as the song that hybrid females preferentially go after.

What factors lead to completely different outcomes in two hybridizations? Natural selection undoubtedly comes first. The Big Bird lineage has novel large blunt beak, which under the circumstances is favored by selection. In contrast, the *G. scandens*-*G. fortis* hybrid males are small in body size which is selected against. Song plays a critical role in both cases. The song type in Darwin's finches is culturally, not genetically inherited. Young males imprint their father's song and females preferentially mate with the males that sing like their fathers. Mis-imprinting of song can alter the mating preference, especially in small populations or in a new niche. If the founder *G. conirostris* of the Big Birds had been a female and hybridized with a resident *G. fortis* male, the offspring would have bred with *G. fortis* and subsequently become part of the *G. fortis*.

Another key to the events is chance and timing. Migration between the oceanic islands is fairly common in Darwin's finches during their adaptive radiation. However, in the Big Birds event, it was rare that a bird flew from a distant island more than 100 km away. More coincidentally, it is a male with a distinctive song. The trigger in the *G. scandens*-*G. fortis* case was the extreme weather from an El Niño event. The unexpected rains changed the vegetation that the finches fed on, creating conditions for the breeding of the F1 hybrids. If the environment was too harsh for the hybrids, introgression would not be promoted.

Conclusively, these cases clearly exemplify how natural selection, song learning and hybridization at the right time largely determine the consequence of hybridization in Darwin's finches. Meanwhile, these novel results bring up new questions for researchers. How often does hybrid speciation happen during the rapid radiation of Darwin's finches? How many of the contemporary species arose from hybrid speciation?

Atlantic herring and a highly contiguous assembly

Large populations are more genetically stable than small populations, because they possess more genetic variability and are less affected by genetic drift. For species with small body size, maintaining a large number is another essential strategy to defend predation risk (103). Atlantic herring is one of the most abundant pelagic species in the North Atlantic marine ecosystem. The census population size of the herring is more than 10^{12} while the effective population size has been estimated to be 4.0×10^5 (5). For an organism with such a huge population size, genetic drift remains minute. Multiple stocks have been recognized by fishery biologists in terms of spawning location, spawning time and morphological features. Herring disperse widely following plankton movements during the feeding period and the stocks mix together most of the

time in a year. During spawning seasons, they isolate from each other and migrate back to specific locations to reproduce. Their exact migration route is still unclear and it can be altered due to changes in climate and habitat, but the seasonal merge-disperse dynamics between stocks may give rise to gene exchange. The between-population gene flow and minute genetic drift together result in extremely low genetic differentiation at selectively neutral loci between the populations (F_{ST} = 0.026) (104). This is beneficial for the identification of selective signatures.

A high-quality assembly can improve the contiguity of existing reference and greatly facilitate genomic studies on various levels. First, it sorts the placement of selective signatures along the genome. Whole-genome screens based on the previous draft assembly of Atlantic herring suggested over 400 independent loci associated with adaptation to salinity as well as at least 125 independent loci associated with adaptation to timing of reproduction (77). Using the same population contrasts and the same cutoff, the numbers based on the chromosome-level assembly are reduced to 125 and 22, respectively. This is of great help for interpreting the loci under selection. It provides a basis to determine whether adjacent loci are linked or whether gene interaction is involved in selection at particular loci. Second, it increases the accuracy of gene annotation. Modern pipelines of gene annotation involve prediction of gene structure from empirical sequence patterns. Improvement in the completeness of genome enables comprehensive annotation. Third, a highly contiguous assembly increases the chances to characterize structural variation. Compared to genetic variants in other species, the presence of structural variants occurs quite frequently in the herring genome. The Chr 12 supergene was in fact among the identified genetic signals in the draft assembly, but split apart across different scaffolds. It may get noticed by careful inspection, but confirming its complete structure and breakpoints would be difficult. The new assembly saves the effort to build up the structure from pieces. Fourth, a chromosome-level assembly provides new insights into the profile of genomic parameters. Both the pedigree-based linkage map and the LD-based recombination map of the herring show that the recombination rate is very high towards one end of the chromosome and low in the remaining part. It displays an "L" shape across the majority of the chromosomes. This is in concordance with the positions of the centromeres observed from the karyotype of the sister species, the Pacific herring, which is not detectable in the draft assembly. It is also worthwhile to confirm if other genomic parameters, such as GC content and density of tandem repeats, are in line with this pattern.

The advent of long-read sequencing and Hi-C technology enhances the uptake of genome assembly. Unlike assembling from mate paired short reads, the new techniques enable assembling a genome in a time-efficient and cost-efficient manner. The herring assembly is one step to the ensembled genomic resources of various organisms. A global project call Earth BioGenome

Project (EBP) is recently launched, which aims to sequence the genomes of all living eukaryotes on earth in ten years (105).

Comparison of Darwin's finches and Atlantic herring as models for evolutionary studies

No model is perfect, but many are useful. An organism is chosen as a model because it often possesses some representative properties that can be generalized to interpret some universal biological characteristics. Meanwhile, the organism cannot be too complex to explain or be used for experimental work. Therefore, an ideal model should be representative as well as simple and tractable. In studies of population genetics, several additional merits are appreciated in an ideal model. An infinite or enormous population size eliminates the influence of genetic drift and thus accentuates the effect of selection. Short divergence time of closely related populations elevates the chance of tracing back the origin of adaptation or speciation and facilitates the detection of gene flow.

However, every organism has its particular biological traits and unique demographic dynamics, which is inescapable in the generalization of genetic analysis. As models in the studies of evolutionary genetics, Darwin's finches and Atlantic herring have their respective benefits and drawbacks. The majority of Darwin's finches have been restricted to an isolated environment for the past million year, and none of the species has become extinct due to human influence, although some studies find that human activities on Galápagos islands caused urbanization in several species of Darwin's finches (106). The ecosystem of the oceanic islands is completely maintained by local flora and fauna that coevolve and depend on each other. This creates an optimal system to observe how species arise and multiply in natural conditions. The volcanic islands of Galápagos initially formed about five million years ago, and the divergence time of Darwin's finches is no earlier than one million years. The ecological environment and the species are both relatively young, which makes it possible to trace selection backward and look for reasonable causes. Despite these advantages, it is still challenging to identify the loci associated with ecological adaptation and speciation. All the species originated from a common ancestor which consisted of only a small number of birds. They rapidly occupied different islands and the population size dramatically expanded. In this process, the origin of new species often initialized with a small subset of incipient population occupying new niches. In some conditions, a population could be founded by merely a pair of individuals (107). As a consequence, the populations suffer from founder effects and genetic drift, which decrease their nucleotide diversity. When performing whole-genome screens between even the closely related populations, genetic drift inevitably introduces noise

that obscures the true selective signals. Another inconvenience of using those finches as a model is about sampling and functional experiment. As a conserved natural site, the ecosystem of the Galápagos is fairly fragile and deserves to be well protected. Thus, further functional experiment can only be conducted on other model organisms as a replacement.

In contrast to the small population size of Darwin's finches, Atlantic herring has an enormous population size. The advantage of the huge population size of Atlantic herring in evolutionary studies has been emphasized many times in the preceding pages. It provides the advantage to discover the genetic variation associated with ecological adaptation. Apart from that, it is convenient to sample and dissect herring for differential expression analysis and molecular experiment. Breeding herring in laboratory conditions can be challenging but manageable. However, herring also has some downsides as a biological model. Although it is easy to sample from a spawning school, herring are too difficult to keep in the laboratory to construct transgenic fish for downstream functional validation. The minimum generation time of herring is about six years, which is too long to breed a pedigree. The marine environment is a large aquatic habitat, which comprises selections from multiple layers. Temperature, light absorbance, predator, food resource, microbiota, potential pathogens and other ecological factors together shape the genetic signatures we observe in the herring genome. These factors are also dependent and confounding variables. To make things more complicated, the influence of selection varies in different life stages of herring. The most critical stages are believed to be egg and larvae. Disentangling the driving variables associated with a given locus seems difficult without proper measurement of ecological metrics. Additionally, human pollution and overfishing certainly bring substantial changes into the natural ecosystem, which forces herring to adapt. These factors should be taken into account when explaining the genetic variation.

Nevertheless, all the models serve for unveiling new knowledge. Darwin's finches as model provide insights into rapid radiation and speciation. Atlantic herring as model are valuable for the characterization of specific loci associated with ecological adaptation. These two models complement each other for answering key evolutionary questions in different aspects.

Acknowledgements

References

1.   Charles D. The variation of animals and plants under domestication. Vol. 2. John Murray; 1868.

2.   Majerus MEN. Industrial Melanism in the Peppered Moth, Biston betularia: An Excellent Teaching Example of Darwinian Evolution in Action. Evol Educ Outreach. 2008;2(1):63–74.

3.   Roach JC, Glusman G, Smit AFA, Huff CD, Hubley R, Shannon PT, et al. Analysis of genetic inheritance in a family quartet by whole-genome sequencing. Science. 2010;328(5978):636–9.

4.   Crow JF. The origins, patterns and implications of human spontaneous mutation. Nat Rev Genet. 2000;1(1):40–7.

5.   Feng C, Pettersson M, Lamichhaney S, Rubin CJ, Rafati N, Casini M, et al. Moderate nucleotide diversity in the Atlantic herring is associated with a low mutation rate. Elife. 2017;6:e23907.

6.   Nei M, Li WH. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci U S A. 1979;76(10):5269–73.

7.   Gregory TR. Insertion-deletion biases and the evolution of genome size. Gene. 2004;324(1–2):15–34.

8.   Mills RE, Luttig CT, Larkins CE, Beauchamp A, Tsui C, Pittard WS, et al. An initial map of insertion and deletion (INDEL) variation in the human genome. Genome Res. 2006;16(9):1182–90.

9.   Mullaney JM, Mills RE, Stephen Pittard W, Devine SE. Small insertions and deletions (INDELs) in human genomes. Hum Mol Genet. 2010;19(R2):R131-6.

10.   Jiang Y, Turinsky AL, Brudno M. The missing indels: An estimate of indel variation in a human genome and analysis of factors that impede detection. Nucleic Acids Res. 2015;43(15):7217–28.

11.   Fang H, Wu Y, Narzisi G, O'Rawe JA, Barrón LTJ, Rosenbaum J, et al. Reducing INDEL calling errors in whole genome and exome sequencing data. Genome Med. 2014;6(10):89.

12.   Sudmant PH, Rausch T, Gardner EJ, Handsaker RE, Abyzov A, Huddleston J, et al. An integrated map of structural variation in 2,504 human genomes. Nature. 2015;526(7571):75–81.

13.   Rubin CJ, Megens HJ, Barrio AM, Maqbool K, Sayyab S, Schwochow D, et al. Strong signatures of selection in the domestic pig genome. Proc Natl Acad Sci U S A. 2012;109(48):19529–36.

14.   Fontanesi L, Beretti F, Riggio V, Gómez González E, Dall'Olio S, Davoli R, et al. Copy number variation and missense mutations of the agouti signaling protein (ASIP) gene in goat breeds with different coat colors. Cytogenet Genome Res. 2010;126(4):333–47.

15.   Lamichhaney S, Berglund J, Almén MS, Maqbool K, Grabherr M, Martinez-Barrio A, et al. Evolution of Darwin's finches and their beaks revealed by genome sequencing. Nature. 2015;518(7539):371–5.

16. Christmas MJ, Wallberg A, Bunikis I, Olsson A, Wallerman O, Webster MT. Chromosomal inversions associated with environmental adaptation in honeybees. Mol Ecol. 2019;28(6):1358–74.

17. Barth JMI, Villegas-Ríos D, Freitas C, Moland E, Star B, André C, et al. Disentangling structural genomic and behavioural barriers in a sea of connectivity. Mol Ecol. 2019;28(6):1394–411.

18. Kosugi S, Momozawa Y, Liu X, Terao C, Kubo M, Kamatani Y. Comprehensive evaluation of structural variation detection algorithms for whole genome sequencing. Genome Biol. 2019;20(1):117.

19. Buckland PR. The importance and identification of regulatory polymorphisms and their mechanisms of action. Biochim Biophys Acta - Mol Basis Dis. 2006;1762(1):17–28.

20. Hrdlickova B, de Almeida RC, Borek Z, Withoff S. Genetic variation in the non-coding genome: Involvement of micro-RNAs and long non-coding RNAs in disease. Biochim Biophys Acta - Mol Basis Dis. 2014;1842(10):1910–22.

21. Goymer P. Genetic variation: Synonymous mutations break their silence. Nat Rev Genet. 2007;8(2):92.

22. van Bakel H, Nislow C, Blencowe BJ, Hughes TR. Most "dark matter" transcripts are associated with known genes. PLoS Biol. 2010;8(5):e1000371.

23. Wagner A. Neutralism and selectionism: A network-based reconciliation. Nat Rev Genet. 2008;9(12):965–74.

24. Wagner A. Robustness and evolvability in living systems. Princeton University Press; 2013.

25. Birch J. Natural selection and the maximization of fitness. Biol Rev Camb Philos Soc. 2016;91(3):712–27.

26. Masel J. Genetic drift. Curr Biol. 2011;21(20):R837-8.

27. Eldredge N. Reinventing Darwin - the Great Debate at the High Table of Evolutionary Theory. New York: Wiley; 1995.

28. Goldberg DE, Barton AM. Patterns and consequences of interspecific competition in natural communities: a review of field experiments with plants. Am Nat. 1992;139(4):771–801.

29. Losos JB. Ecological character displacement and the study of adaptation. Proc Natl Acad Sci U S A. 2000;97(11):5693–5.

30. Schluter D. Ecological character displacement in adaptive radiation. Am Nat. 2000;156:4–16.

31. Schluter D. Evidence for ecological speciation and its alternative. Science. 2009;323(5915):737–41.

32. Hendry AP, Nosil P, Rieseberg LH. The speed of ecological speciation. Funct Ecol. 2007;21(3):455–64.

33. Uyeda JC, Hansen TF, Arnold SJ, Pienaar J. The million-year wait for macroevolutionary bursts. Proc Natl Acad Sci U S A. 2011;108(38):15908–13.

34. Cook OF. Factors of species-formation. Science. 1906 Mar 30;23(587).

35. Tregenza T, Butlin RK. Speciation without isolation. Nature. 1999;400(6742):311–2.

36. Coyne JA. Sympatric speciation. Curr Biol. 2007;17(18).

37. Malinsky M, Challis RJ, Tyers AM, Schiffels S, Terai Y, Ngatunga BP, et al. Genomic islands of speciation separate cichlid ecomorphs in an East African crater lake. Science. 2015;350(6267):1493–8.

38. Nosil P. Speciation with gene flow could be common. Mol Ecol. 2008;17(9):2103–6.

39. Feder JL, Egan SP, Nosil P. The genomics of speciation-with-gene-flow. Trends Genet. 2012;28(7):342–50.
40. Hoskin CJ, Higgie M, McDonald KR, Moritz C. Reinforcement drives rapid allopatric speciation. Nature. 2005;437(7063):1353–6.
41. Adams KL, Wendel JF. Polyploidy and genome evolution in plants. Curr Opin Plant Biol. 2005;8(2):135–41.
42. Burke JM, Arnold ML. Genetics and the Fitness of Hybrids. Annu Rev Genet. 2001;35(1):31–52.
43. Wright S. Genetics and the origin of species. J Hered. 1942;33(8):283–4.
44. Seehausen O, Butlin RK, Keller I, Wagner CE, Boughman JW, Hohenlohe PA, et al. Genomics and the origin of species. Nat Rev Genet. 2014;15(3):176–92.
45. Schaeffer SW. Muller "elements" in Drosophila: How the search for the genetic basis for speciation led to the birth of comparative genomics. Genetics. 2018;210(1):3–13.
46. Livesay EA. An Experimental Study of Hybrid Vigor or Heterosis in Rats. Genetics. 1930;15(1):17–54.
47. Baack EJ, Rieseberg LH. A genomic view of introgression and hybrid speciation. Curr Opin Genet Dev. 2007;17(6):513–8.
48. Chen ZJ. Molecular mechanisms of polyploidy and hybrid vigor. Trends Plant Sci. 2010;15(2):57–71.
49. Otto SP, Whitton J. Polyploid Incidence and Evolution. Annu Rev Genet. 2000;34(1):401–37.
50. Vallejo-Marín M, Hiscock SJ. Hybridization and hybrid speciation under global change. New Phytol. 2016;211(4):1170–87.
51. Nieto Feliner G, Álvarez I, Fuertes-Aguilar J, Heuertz M, Marques I, Moharrek F, et al. Is homoploid hybrid speciation that rare? An empiricist's view. Heredity (Edinb). 2017;118(6):513–6.
52. Darwin C. Journal of researches into the natural history and geology of the countries visited during the voyage round the world of H.M.S. "Beagle" under the command of Captain Fitz Roy, R.N. John Murray; 1845.
53. Lack D. Darwin's Finches. London: Cambridge University Press; 1947.
54. Sato A, Tichy H, O'Huigin C, Grant B PR, Grant R, Klein J. On the origin of Darwin's finches. Mol Biol Evol. 2001;18(3):299–311.
55. Funk ER, Burns KJ. Biogeographic origins of Darwin's finches (Thraupidae: Coerebinae). Auk. 2018;135(3):561–71.
56. Vincek V, O'Huigin C, Satta Y, Takahata N, Boag PT, Grant PR, et al. How large was the founding population of Darwin's finches? Proc R Soc B Biol Sci. 1997;264(1378):111–8.
57. Parent CE, Caccone A, Petren K. Colonization and diversification of Galápagos terrestrial fauna: A phylogenetic and biogeographical synthesis. Philos Trans R Soc B Biol Sci. 2008;363(1508):3347–61.
58. Sato A, O'huigin C, Figueroa F, Grant PR, Grant BR, Tichy H, et al. Phylogeny of Darwin's finches as revealed by mtDNA sequences. Proc Natl Acad Sci U S A. 1999;96(9):5101–6.
59. Burns KJ, Hackett SJ, Klein NK. Phylogenetic relationships and morphological diversity in Darwin's finches and their relatives. Evolution (N Y). 2002;56(6):1240–52.
60. Petren K, Grant PR, Grant BR, Keller LF. Comparative landscape genetics and the adaptive radiation of Darwin's finches: The role of peripheral isolation. Mol Ecol. 2005;14(10):2943–57.

61. Abzhanov A, Protas M, Grant BR, Grant PR, Tabin CJ. Bmp4 and morphological variation of beaks in Darwin's finches. Science. 2004;305(5689):1462–5.

62. Abzhanov A, Kuo WP, Hartmann C, Grant BR, Grant PR, Tabin CJ. The calmodulin pathway and evolution of elongated beak morphology in Darwin's finches. Nature. 2006;442(7102):563–7.

63. Mallarino R, Grant PR, Grant BR, Herrel A, Kuo WP, Abzhanov A. Two developmental modules establish 3D beak-shape variation in Darwin's finches. Proc Natl Acad Sci U S A. 2011;108(10):4057–62.

64. Grant PR. Hybridization of Darwin's finches on Isla Daphne Major, Galapagos. Philos Trans - R Soc London, B. 1993;340(1291):127–39.

65. Rieucau G, De Robertis A, Boswell KM, Handegard NO. School density affects the strength of collective avoidance responses in wild-caught Atlantic herring Clupea harengus: A simulated predator encounter experiment. J Fish Biol. 2014;85(5):1650–64.

66. Food and Agriculture Organization of the United Nations Statistics Division (FAO). FAO Yearbook: Fishery and Aquaculture Statistics of 2015. FAO Annual Yearbook. 2017.

67. McQuinn IH. Metapopulations and the Atlantic herring. Rev Fish Biol Fish. 1997;7(3):297–329.

68. Ruzzante DE, Mariani S, Bekkevold D, André C, Mosegaard H, Clausen LAW, et al. Biocomplexity in a highly migratory pelagic marine fish, Atlantic herring. Proc R Soc B Biol Sci. 2006;273(1593):1459–64.

69. Kennedy J, Nash RDM, Slotte A, Kjesbu OS. The role of fecundity regulation and abortive maturation in the reproductive strategy of Norwegian spring-spawning herring (Clupea harengus). Mar Biol. 2011;158(6):1287–99.

70. Bucholtz RH, Tomkiewicz J, Nyengaard JR, Andersen JB. Oogenesis, fecundity and condition of Baltic herring (Clupea harengus L.): A stereological study. Fish Res. 2013;145:100–13.

71. Parrish BB, Saville A, Craig RE, Baxter IG, Priestley R. Observations on herring spawning and larval distribution in the firth of clyde in 1958. J Mar Biol Assoc United Kingdom. 1959;38(3):445–53.

72. Postuma KH. The nucleus of the herring otolith as a racial character. ICES J Mar Sci. 1974;35(2):121–9.

73. Folkvord A, Høie H, Johannessen A, Solbakken T. Effects of prey concentration, light regime, and parental origin on growth and survival of herring larvae under controlled experimental conditions. ICES J Mar Sci. 2009;66(8):1702–9.

74. L A, N R, R R, G S. Genetic variability in Atlantic herring (Clupea harengus harengus): description of protein loci and population data. Hereditas. 1981;95(1):69–78.

75. Ryman N, Lagercrantz U, Andersson L, Chakraborty R, Andersson L. Lack of correspondence between genetic and morphologic variability patterns in atlantic herring (Clupea harengus). Heredity (Edinb). 1984;53(3):687–704.

76. Larsson LC, Laikre L, Palm S, André C, Carvalho GR, Ryman N. Concordance of allozyme and microsatellite differentiation in a marine fish, but evidence of selection at a microsatellite locus. Mol Ecol. 2007;16(6):1135–47.

77. Barrio AM, Lamichhaney S, Fan G, Rafati N, Pettersson M, Zhang H, et al. The genetic basis for ecological adaptation of the Atlantic herring revealed by genome sequencing. Elife. 2016;5:e12081.

78. Lack D. Darwin's finches. Vol. 15, Scientific American. 1947. 614–615 p.

79. Brown WL, Wilson EO. Character displacement. Syst Zool. 1956;5(2):49–64.

80.   GRANT PR. Convergent and divergent character displacement. Biol J Linn Soc. 1972;4(1):39–68.

81.   Stuart YE, Losos JB. Ecological character displacement: Glass half full or half empty? Trends Ecol Evol. 2013;28(7):402–8.

82.   Grant PR, Grant BR. Evolution of character displacement in Darwin's finches. Science. 2006;313(5784):224–6.

83.   Zhang G, Li C, Li Q, Li B, Larkin DM, Lee C, et al. Comparative genomics reveals insights into avian genome evolution and adaptation. Science. 2014;346(6215):1311–20.

84.   Zhou X, Benson KF, Ashar HR, Chada K. Mutation responsible for the mouse pygmy phenotype in the developmentally regulated factor HMGI-C. Nature. 1995;376(6543):771–4.

85.   Carneiro M, Hu D, Archer J, Feng C, Afonso S, Chen C, et al. Dwarfism and altered craniofacial development in rabbits is caused by a 12.1 kb deletion at the HMGA2 locus. Genetics. 2017;205(2):955–65.

86.   Fatemifar G, Hoggart CJ, Paternoster L, Kemp JP, Prokopenko I, Horikoshi M, et al. Genome-wide association study of primary tooth eruption identifies pleiotropic loci associated with height and craniofacial distances. Hum Mol Genet. 2013;22(18):3807–17.

87.   Turner TL, Hahn MW, Nuzhdin S V. Genomic islands of speciation in Anopheles gambiae. PLoS Biol. 2005;3(9):1572–8.

88.   Cruickshank TE, Hahn MW. Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. Mol Ecol. 2014;23(13):3133–57.

89.   Singhal S, Leffler EM, Sannareddy K, Turner I, Venn O, Hooper DM, et al. Stable recombination hotspots in birds. Science. 2015;350(6263):928–32.

90.   Grant PR, Grant BR. 40 years of evolution: Darwin's Finches on Daphne Major Island. Princeton University Press; 2014.

91.   Grant PR, Grant BR. The secondary contact phase of allopatric speciation in Darwin's finches. Proc Natl Acad Sci U S A. 2009 Dec 1;106(48):20141–8.

92.   Auton A, McVean G. Recombination rate estimation in the presence of hotspots. Genome Res. 2007;17(8):1219–27.

93.   Jetz W, Thomas GH, Joy JB, Hartmann K, Mooers AO. The global diversity of birds in space and time. Nature. 2012;491(7424):444–8.

94.   Price TD, Bouvier MM. the Evolution of F1 Postzygotic Incompatibilities in Birds. Evolution (N Y). 2002;56(10):2083.

95.   Bolnick DI, Near TJ. Tempo of hybrid inviability in centrarchid fishes (Teleostei: Centrarchidae). Evolution (N Y). 2005;59(8):1754–67.

96.   Prager EM, Wilson AC. Slow evolutionary loss of the potential for interspecific hybridization in birds: a manifestation of slow regulatory evolution. Proc Natl Acad Sci U S A. 1975;72(1):200–4.

97.   Arbogast BS, Drovetski S V., Anderson DJ, Curry RL, Boag PT, Seutin G, et al. The origin and diversification of Galapagos mockingbirds. Evolution (N Y). 2006;60(2):370–82.

98.   Reddy S, Driskell A, Rabosky DL, Hackett SJ, Schulenberg TS. Diversification and the adaptive radiation of the vangas of Madagascar. Proc R Soc B Biol Sci. 2012;279(1735):2062–71.

99.   Jønsson KA, Fabre PH, Fritz SA, Etienne RS, Ricklefs RE, Jørgensen TB, et al. Ecological and evolutionary determinants for the adaptive radiation of the Madagascan vangas. Proc Natl Acad Sci U S A. 2012;109(17):6620–5.

100. Mavárez J, Salazar CA, Bermingham E, Salcedo C, Jiggins CD, Linares M. Speciation by hybridization in Heliconius butterflies. Nature. 2006;441(7095):868–71.

101. Larsen PA, Marchán-Rivadeneira MR, Baker RJ. Natural hybridization generates mammalian lineage with species characteristics. Proc Natl Acad Sci U S A. 2010;107(25):11447–52.

102. Hermansen JS, Sæther SA, Elgvin TO, Borge T, Hjelle E, Sætre GP. Hybrid speciation in sparrows I: Phenotypic intermediacy, genetic admixture and barriers to gene flow. Mol Ecol. 2011;20(18):3812–22.

103. Magurran AE, Henderson PA. How selection structures species abundance distributions. Proc R Soc B Biol Sci. 2012;279(1743):3722–6.

104. Lamichhaney S, Fuentes-Pardo AP, Rafati N, Ryman N, McCracken GR, Bourne C, et al. Parallel adaptive evolution of geographically distant herring populations on both sides of the North Atlantic Ocean. Proc Natl Acad Sci U S A. 2017;114(17):E3452–61.

105. Lewin HA, Robinson GE, Kress WJ, Baker WJ, Coddington J, Crandall KA, et al. Earth BioGenome Project: Sequencing life for the future of life. Proc Natl Acad Sci U S A. 2018;115(17):4325–33.

106. De León LF, Sharpe DMT, Gotanda KM, Raeymaekers JAM, Chaves JA, Hendry AP, et al. Urbanization erodes niche segregation in Darwin's finches. Evolutionary Applications. 2019;1329– 1343.

107. Grant PR, Grant BR, Petren K. A population founded by a single pair of individuals: Establishment, expansion, and evolution. Genetica. 2001;112–113:359–82.

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine 1620

Editor: The Dean of the Faculty of Medicine

A doctoral dissertation from the Faculty of Medicine, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine. (Prior to January, 2005, the series was published under the title "Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine".)