NTNU

Norwegian University of
Science and Technology

The Genetics of Basal Metabolic Rate in free-living Populations of House Sparrows (*Passer domesticus*)

Øystein Øy Tennfjord

MSc in Biology
Submission date: June 2016
Supervisor: Henrik Jensen, IBI
Co-supervisor: Bernt Rønning, IBI
Ingerid Julie Hagen, IBI

Norwegian University of Science and Technology
Department of Biology

Acknowledgements

I would like to extend my gratitude to my supervisor Associate Professor Henrik Jensen for tutoring and providing me with supervision. I am very thankful for being allowed to participate in this project. My co-supervisors Dr. Bernt Rønning and Dr. Ingerid Hagen have been valuable resources for me. All of these persons have done excellent work on this project, long before I appeared, my thesis would literally not be possible if it were not for them!

I want to thank all the people who made fieldwork a joyful experience. A special mention goes to the most valuable player in the mist-netting business Peter Ranke, who also taught me the finer arts about sipping sherry. My fellow students Tomas “the train” Taylor, and Erik “the berserker” Degnes Neerland have been a great help disturbing my daily work schedule, it has truly been a pleasure. I also want to thank my professors and lecturers I have had here at NTNU, they are quite a diverse bunch. I have learned a lot about scientific perspectives and critique, which I hope to carry with me for the rest of my life.

There is not enough room on this piece of paper to thank all the people I want to thank. have spent most of my last years in Trondheim studying at NTNU and I have met many people that has made my life here a joy. It is not possible to do anything in isolation, you always have helpers. My greatest helper is my girlfriend Elida Skøien. I would never had finished this thesis if it was not for her. Thank you so much for not making me give up.

Thank you Marit and Ove, I made it!

Øystein Øy Tennfjord

June 2016

Abstract

Animals must allocate their energy reserves between growth, reproduction and survival. Basal metabolic rate (BMR) is a quantitative trait, which can represent the "cost of living", and is a key component in the energy budget of animals. In this study, I explored the genetic basis of BMR through next-generation molecular methods, using individuals from wild populations of house sparrows (*Passer domesticus*). Birds from three genetically connected populations off the coast of mid-Norway were genotyped on a custom Affymetrix 200k SNP-chip, and were subsequently used in genome-wide association study (GWAS). BMR data from the study populations were collected during 2012-2015. One of three populations in this study was artificially founded in 2012 using a common garden approach. Founders from the two source populations had high and low BMR, respectively. Furthermore, the source populations were up-selected and down-selected for BMR 2012-2014. By using a recently developed GWA method that accounts for relatedness and repeated records from individuals, GWA analyses were performed within each island, and on a pooled dataset containing all individuals. Genetic variance components of BMR and the chromosomal contribution to variance in BMR were estimated via SNPs. In the three study populations of house sparrows, additive genetic variance for BMR was significant. Further, genomic partitioning of variance suggested an oligogenic architecture of BMR. It is plausible that the observed variation in BMR can be partly explained by rare loci of larger effect. The repeated GWAS showed mixed results. One suggestive SNP associated with BMR was detected in the pooled analysis, revealing a link between a genotype and BMR in wild house sparrows. However, within island analyses found no significant or suggestive signal for this SNP. Instead, another SNP was found to be associated with BMR within one of the islands. The implications of these findings are discussed in terms of the possible genetic architecture of BMR in these house sparrows, statistical power, biology and population structure. The present study is the very first to use SNPs to map QTLs for BMR in a wild population. Additionally, it contains the largest marker density in a wild outbred vertebrate population to date. The findings were mixed, but this study provides new knowledge about the link between genetics and phenotypic variation in BMR, which again give rise to new questions regarding the genetic architecture and eco-evolutionary dynamics of BMR.

Sammendrag

Dyr må allokere sine energireserver mellom vekst, reproduksjon og overlevelse. Basal metabolsk rate (BMR) er et kvantitativt trekk som kan representere «kostnaden av å leve». BMR er en av hovedkomponentene til energibudsjett hos dyr. I dette studiet utforsket jeg det genetiske fundamentet for BMR fra en vill populasjon gråspurv (*Passer domesticus),* ved bruk av de nyeste molekylære metodene som eksisterer innenfor feltet genetikk. Fugler fra tre genetisk nære populasjoner fra utenfor kysten av midt-Norge ble genotypet med en Affymetrix 200k SNIP-chip. BMR data fra studiepopulasjonene ble samlet inn fra 2012-2015. Den ene av de tre øy-populasjonene ble i 2012 kunstig grunnlagt ved en «common garden approach», ved bruk av utvalgte individer fra de to andre øy-populasjonene, henholdsvis med høy og lav BMR. Populasjonene i de originale øyene har videre blitt kunstig selektert for høyere og lavere BMR. Nylig utviklede genkartleggings metoder,som tar hensyn til slektskap og repeterte mål, ble brukt innen hver øy og i et «pooled» dataset med data fra samtlige individer. Genetisk varianskomponenter til BMR og det kromosomale bidraget til varians i BMR ble videre estimert via SNPer. Additiv genetisk varians for BMR var signifikant for de tre studiepopulasjonene. Videre, så indikerte genetiske partisjonsanalyser en mulig «oligogenic» arkitektur til BMR. Det er plausibelt at den observerte variasjonen I BMR kan delvis forklares av sjeldne loci med moderat til stor effekt. Genkartlegging analysen viste blandede resultater. En SNP assosiert med BMR ble funnet i pooled analysis, noes om avslørte en kobling mellom en genotype og BMR i gråspurv. I analysene innen øy ble det ikke funnet signifikante eller suggestive signal for denne SNP. I stedet ble en annen SNP funnet å være assosiert med BMR innen en av øyene. Implikasjonene av disse funnene er her diskutert med omsyn på den mulige genetiske arkitekturen av BMR i gråspurv, statistisk tyngde, biologi og populasjonsstruktur. Dette studiet er det aller første som bruker SNPs til å kartlegge det genetiske basis for BMR i frittlevende popuasjoner. I tillegg har dette studiet den største markørtetthet i en frittlevende ut-avla vertebrat populasjon. Funnene fra dette studiet var blandede, men tilfører ny kunnskap om sammenhengen mellom genotype og fenotype variasjon i BMR, som igjen leder til nye spørsmål om den genetiske arkitekturen til BMR.

Table of contents

Abbreviations

GWAS	Genome-wide association study
SNP	Single-nucleotide polymorphism
BMR	Basal metabolic rate
Kbp	Kilobase pair
QTL	Quantitative trait loci
HGLM	Hierarchical generalized linear model
LRT	Likelihood-ratio tests
rGLS	Repeated measurement GWAS model
GRM	Genomic relationship matrix
BLAST	Basic local alignment search tool
GCTA	Genome-wide complex trait analysis
LD	Linkage disequilibirum
MAF	Minor allele frequency
V_a	Additive genetic variance
V_{pe}	Permanent environment variance
MDS	Multidimensional scaling

Introduction

Why is nature so diverse? Living organisms show variation in mostly every way imaginable, from how they behave to how their physiology functions. Since the beginning of the 20th century, quantitative genetics have emerged as a field searching for the genetic underpinnings of variation of continuous traits (Roff, 2007). Research in this field has mostly dealt with interpreting distribution of phenotypes in terms of genetic variance components estimated through controlled designs and statistical methodology, without sampling the molecular underpinnings of the phenotype (Hill, 2012; Lynch and Walsh, 1998). Although traditional quantitative genetic are useful (Lynch and Walsh, 1998), they cannot give answers about the DNA sequence variation that affects phenotype variation. How many genes affect a trait? Where are they located, and what is their proportional effect? Understanding the link between genotype and phenotype requires molecular level analyses, and we must start at the simplest level before we delve into the black box of genetic architecture. Recent advancement in next-generation sequencing methods enables a highly specific analyses of the molecular level of quantitative genetics (Mardis, 2008)

Quantitative trait loci (QTLs) are loci underlying the variation in continuous traits (Mackay et al., 2009). Classic examples of continuous traits are height, blood pressure and metabolic rate. QTLs can be mapped using visible markers with known genotypes, either in linkage disequilibrium (LD) with QTLs, or indirectly linked through LD with markers that are in LD with the QTL (Hirschhorn and Daly, 2005). Individuals with different genotypes on markers will then have different phenotypic means, and markers will be associated with the quantitative trait of interest (Mackay et al., 2009). The first QTL mapping was performed by Sax (1923), who found an association between a phenotypic marker and seed weight in beans (*Phaseolus vulgaris*). Current studies of natural animal populations are more complex, demanding thousands of molecular markers and mapping methods (Schielzeth and Husby, 2014). Mapping QTLs in natural populations is advantageous in several ways. Traits studied under laboratory conditions might not exhibit the full range of natural variation, and results might not be generalized to the natural environment (Mackay et al., 2009). Mapping the genes underlying the natural phenotypic variation directly addresses the contribution of the genome on phenotype variation. Knowledge of the genetic architecture of quantitative phenotypic traits is necessary in search for deeper understanding of evolutionary dynamics

in animal populations. The genetics of natural populations are not only of interest for evolutionary biologists. Management of wildlife and vertebrate populations depend on predictive models about the evo-ecological dynamics of populations, which is of great importance in development of sustainable harvest models of animal populations (Allendorf et al., 2008).

In general, methods of gene mapping of natural populations use either experimentally manipulated linkage disequilibrium in the genome of populations (LD), or natural LD to map genes (Schielzeth and Husby, 2014). Both approaches have been successful in natural populations (Colosimo et al., 2004; Husby et al., 2015; Johnston et al., 2011; Tarka et al., 2010). Linkage analyses utilizes long reaching LD, either experimentally manipulated or naturally occurring within a pedigree (Schielzeth and Husby, 2014). Another approach is association analyses, which utilizes natural occurring short reaching LD, based on historical recombination (Mackay et al., 2009; Slate et al., 2010).

Diverse molecular techniques and markers are available, but next-generation sequencing and high-density single nucleotide polymorphism (SNP) genotyping are becoming a cost effective approach towards QTL/gene mapping in natural populations (Ekblom and Galindo, 2011; Ellegren and Sheldon, 2008; Slate, 2005; Slate et al., 2010). SNPs appear in high density throughout the genome, which make genome-wide association studies (GWAS) possible (Hirschhorn and Daly, 2005). Recent developments in genomic analyses also makes it possible to partition genetic variance across chromosomes (Yang et al., 2011b). By comparing the relationship between variance explained by chromosomes and chromosome size, one can get insight to the general genetic basis of a trait (Robinson et al., 2013; Yang et al., 2011b), e.g. a polygenic basis would be displayed as a linear relationship between the variance explained and chromosome size, assuming effect loci are evenly distributed across the genome.

In the present study, I explored the genetics of basal metabolic rate (BMR), an important trait of the energetics family, through recent developed mixed model GWAS methods (Rönnegård et al., 2016) and genome partitioning analyses, in free-living populations of house sparrows (*Passer domesticus*), typed on a custom Affymetrix Axiom 200k-SNP-array.

Animals must allocate their energy reserves between growth, reproduction and survival. BMR is measured within the thermoneutral zone, while the organism is at rest and in a postabsorptive state (Bligh and Johnson, 1973). Thus, BMR can be described as the "the cost of living" (Hulbert and Else, 2000). BMR show great variation within-species (Konarzewski and Ksiazek, 2013), and is recognized as an important factor shaping the ecology, physiology and behavior of organisms (Berteaux et al., 1996). BMR is for example associated with a slower pace of life (long lived and produce fewer offspring) in tropical birds compared to temperate birds (Wiersma et al., 2007). Furthermore, BMR is found to be associated with fitness in two of our study populations (Rønning et al., 2015). However, the genetic basis of intra-specific variation in BMR is not completely identified (Konarzewski and Ksiazek, 2013). Exploring the genetic basis of BMR will thus add to our understanding about an important energetic trait associated with the life history of animals.

Study islands are off the coast of Norway where BMR measurements have been collected periodically from 2012-2015. Additionally, one of the three populations in this study was founded using a common garden approach. Founders from the two parent populations had high and low BMR, respectively. Further, individuals in parent islands were subsequently up-selected and down-selected for BMR. In this study, I use a newly developed mixed model approach adjusting for repeated observations (Rönnegård et al., 2016), first applied in Husby et. al (Husby et al., 2015). This model controls both for relatedness between individuals, and repeated observations, which is ideal for our longitudinal dataset. A GWAS analysis was first performed on a pooled dataset including all islands since they are genetically connected via the individuals translocated to the common garden population on Lauvøya, and the larger sample size was expected to increase power to detect QTL for BMR. Second, I carried out GWAS analyses in each population separately. This was because I expected LD-blocks to be larger in the "hybrid" population on Lauvøya, which was founded by individuals at the opposite phenotypic ends of the phenotypic distribution for BMR, increasing the correlation between my SNPs and any genes coding for BMR (Mackay et al. 2009;Schielzeth and Husby, 2014). In addition, I expected that genes or alleles causing either low or high BMR were more likely to show up in the down-selected population on Leka and the up-selected population on Vega, respectively.

The aim of this study was to explore the genetic architecture of BMR in free-living house sparrows, which included four main goals: I) Estimate genetic and environmental variance components of BMR II) Find QTLs underlying variation in BMR. III) Estimate the effects of QTLs and their location. IV) Partition additive genetic variance across chromosomes to study the contribution different chromosomes has on BMR to detect the genetic architecture of BMR.

Methods

Study area

The study area include three islands along the coast of mid-Norway, with boreal climate. The island locations are; Vega (65.6°N, 11.9°E) and Leka (65°N, 11.6°E) in Nordland, and Lauvøya (63.9°N, 9.9°E) in Trøndelag.

Study species

The house sparrow (*Passer domesticus*) is a common passerine bird with a worldwide distribution. In this study system the birds inhabit islands, and only 10 % of the fledglings that recruit into the breeding population are immigrants (Pärn et al., 2012, 2009) . Each pair lay on average 1-3 clutches per season(Kvalnes et al., 2013; Ringsby et al., 2002) with an average clutch size of 5 eggs (Husby et al., 2006). Hatching date is from early May until the middle of August (Ringsby et al., 2002). A practical feature of the house sparrows is that their habitat tends to be associated with farms and human settlements, which makes sampling convenient (Jensen et al., 2013).

Fieldwork

The fieldwork for this project has been ongoing since 2012-2015. I participated in one field season in 2015. Every winter (February-March) all house sparrows were captured on the three islands using mist nets. Birds captured first time were marked with a numbered aluminum ring, and three colored plastic rings to enable individual identification. Due to the high proportion of ringed individuals in these populations, an individual without rings was defined as a recruit from the previous breeding season (see Jensen et al., 2008; Rønning et al., 2015). Nests were visited on Lauvøya during the breeding season (May - August), where the number of eggs and fledglings (nestlings in the nest 8-12 days after hatching) were counted for each nest, and the birds were marked See Ringsby et al. (2002) for procedures on gathering nestling data. On first handling, individuals were sampled for 25µL blood taken by brachial venipuncture, for the use in e.g. microsatellite- and SNP-genotyping. For details on blood sampling see Jensen et al. (2003).

Common garden experiment

In February-March 2012, BMR was collected for virtually all house sparrows on Leka (ca. 180 individuals) and Vega (ca. 170 individuals). Approximately 40% of the male and female birds on each island with lowest (Leka) and highest (Vega) levels of BMR were released back on the island, whereas 70 house sparrows from each of the islands were translocated to a common environment on Lauvøya in Sør-Trøndelag. Nearly all (97%) of the 72 indigenous house sparrows on Lauvøya, and 83% of the 18 house sparrows on the mainland closest (<2.5 km) to Lauvøya were moved across mountains and fjords and released in suitable habitat >80 km away prior to translocation. From Leka, the 70 individuals with the highest levels of BMR were translocated to Lauvøya. Similarly, from Vega the 70 individuals with the lowest BMR were selected to be translocated to Lauvøya. By introducing approximately twice as many individuals as the original population on Lauvøya, we compensated for the fact that many birds would not establish and breed on the island (see Skjelseth et al., 2007). The remaining ca. 20% individuals on Leka and Vega (with close to average levels of BMR) were removed from the island and released in suitable habitat across mountains and fjords >120 km away . BMR selection continued on Leka until 2014 and on Vega until 2013, and any responses to selection were recorded until 2015 and 2014 on Leka and Vega, respectively. After the experiment there were high BMR individuals on Vega, low BMR individuals on Leka, and a mix of high and low BMR individuals from two different populations on Lauvøya. Individual measurements on BMR have been collected from all individuals on each study island in February-March 2012-2015, with year 2015 missing for Vega (Appendix Fig. A1).

Data collection - Basal metabolic rate

In the present study, BMR was measured as oxygen consumption rates using an open flow system. A custom-made climate chamber (Appendix Fig. A2) was used to house eight respirometer chambers (metabolic chambers) holding the house sparrows. Here, the temperature was constantly held at 25-30°C, which is within the thermoneutral zone of the house sparrow (Hudson and Kimzey, 1966). These enable BMR data collection of eight birds at the same time. Of the daily catch from different locations on the three islands, 16 birds were randomly selected for BMR measurements the same day. The 16 birds were randomly divided in two batches of 8 individuals for measurements within their two resting periods. Individuals of the first batch were measured in the evening (16:00 – 23:00), whereas individuals in the other batch were measured during the night 23:00 – 08:00). The birds measured during the night were kept in separate cages, supplied with food until they were transferred to the metabolic chambers at 23:00. A small piece of bread was fed to the birds to assure well-being, not enough to disturb them from entering a postabsorptive state during measurement period. All birds were weighted (to the nearest 0.1 g) before the BMR measurements were conducted. Any remaining birds in the daily catch were moved to an aviary and kept for BMR measurements the day after. All birds went through the same experimental protocol, irrespective if they were measured on catch day or after staying one night in the aviary. For more details about the BMR measurement procedure in see Rønning et al. (2015).

Data collection - SNP Genotyping

The protocol and development of a medium density 10k SNP-array is explained in detail in Hagen et al. (2013). In the present study, an improved custom Affymetrix Axiom 200k-SNP-array was used. Variable sites were identified from a 10x whole genome sequencing of 33 individual house sparrows, at least one presumably unrelated male and female were selected from 15 populations across Norway and Finland. Genome sequences were aligned to the House sparrow reference genome (Elgvin et al. ms; Hagen et al. ms) , where 185k SNPs with little variation in flanking sequence were selected to be distributed more or less evenly, 6000bp apart, across the chromosomes (Elgvin et al. ms; Hagen et al. ms). Additionally, 15k SNPs were placed close to selected candidate genes. A total number of 186 056 SNPs with the highest Affymetrix quality score (Polymorphic High Resolution) were used

for genotyping, and were successfully typed in 715 individuals with BMR measurements. Genomic filtering was applied with Identical by state (IBS) > 0.95 to identify duplicate samples, minor allele frequency (MAF) > 0.01, genotype success rate > 0.9, and callrate > 0.95 to create map. and ped. files for the three islands Leka, Vega and Lauvøya. This filtering resulted in 184 688 markers and 711 genotyped individuals with BMR measures, where 4 indviduals within the study populations were identified as duplicates, hence excluded from the analyses. Positions are correct within scaffolds on the different chromosomes, but not within chromosomes because I did not have access to positions in a completed reference genome sequence where scaffold were ordered correctly. We also do not have any assembly for chromosome 16 probably because of many gene-duplications on this chromosome.

Genetic analyses – Quality control and population structure

If not otherwise mentioned, all statistical analyses was performed in R version 3.2.5 (R Core Team, 2016). Before GWAS and genome partitioning analyses, a second quality control (QC) was performed with the GenABEL R-package (Aulchenko et al. 2007a). First, a QC was conducted within islands to remove SNPs likely to have genotyping error. This was done by excluding SNPs for which a Hardy-Weinberg (HW) equilibrium test had a level of significance of $p = 2.7 \times 10^{-7}$ (0.05/184 688 nr. of SNPs). This was because deviations from Hardy-Weinberg equilibrium were expected to occur in the dataset due to clear population structuring and strong artificial selection. Hence, genotyping error was confounded with these factors. Performing QC within population should help mitigate these confounding effects and uncover SNPs with real genotyping errors. I then pooled the accepted markers and individuals from within islands, and adjusted for callrate > 0.95, genotype sucessrate > 0.95, minor allele MAF > 0.01, and Identical by state IBS> 0.95. After QC we were left with 183 876 markers and 697 individuals for the pooled analysis. Four individuals were excluded due to disagreement between phenotypic and genotypic sex. One individual was excluded due to too high heterozygosity. Nine individuals were excluded due to high number of missing genotypes. Population structure was visually examined by performing multidimensional scaling (MDS) (Mardia, 1978) on the distance-transformed genomic relationship matrix estimated from autosomal markers (GRM). The GRM was estimated using the IBS-function in GenABEL (Aulchenko et al. 2007a). In addition to subsequent clustering analysis via the k-means algorithm (Hartigan and Wong, 1979). For the within-

island analyses, QC was performed separately on each island population with the same control parameters as for pooled analysis, which gave 181 844 markers and 258 individuals for Lauvøya, 181 342 markers and 172 individuals from Vega, and 181 298 markers and 267 individuals from Leka. Because markers has to be adjusted QC parameters for within-islands analyses, some markers will not appear in all islands. For an overview of sample statistics, see Table 1. The Cgmisc genomic visualization software R-package was used to create linkage disequilibrium (LD) plots (Kierczak et al., 2015), where LD (r^2) between markers was calculated via the fast LDcompare method developed by Hao (2007).

Geneetic analyses – Controlling for covariates and factors

Due to having repeated observations on individuals with different degrees of relatedness, I had to control for confounding factors of relatedness and environment (Wilson et al. 2010; Aulchenko et al. 2007b). In addition, I expected that BMR would show average differences for sex, age and period of measurement, and there would be an approximate linear relationship between body mass and BMR. Hence, factors and covariates of sex, age and mass were selected prior to analyses because of their biological association with BMR, as seen in similar studies of the trait (Rønning et al., 2014; Versteegh et al., 2008). Measurement time itself can also affect BMR values, hence period of measurement also had to be included in the model (Page et al., 2011). Effects of covariates and factors was controlled for by performing GWAS analysis in two steps using the repeated measurements function (rGLS) implemented in the RepeatABEL R-package (Rönnegård et al., 2016). In short, a linear mixed model was fitted with the fixed effects (fixed factors and covariates) and random factors of relatedness and identity. Then each marker was fitted as a fixed factor (with three levels) on the independent residuals from the first model in a subsequent generalized linear model where the environmental- and polygenetic effects on BMR are accounted for. This method is similar to the GRAMMAR-method (Aulchenko et al. 2007b), but with the inclusion of repeated measurements.

Likelihood ratio tests (LRT) (Burnham and Anderson, 2002) were utilized to determine the significance of each covariate and factor of sex, age , mass and period of measurement. See Appendix Table A1 for LRT statistics. Inconveniently, the mixed model part of the rGLS function is hardcoded, where likelihood of model is not calculated by default. To gauge if covariates explained a significant part of variation in BMR, a HGLM model was fitted only

with animal identity as a random factor and covariates and factors to calculate LRT p-values. Likelihood ratio of full model (including all covariates and factors) was compared against reduced models (one covariate is dropped in each reduced model) to get LRT p-values for covariates. Parameter estimates from HGLM model with no GRM was not significantly different from rGLS estimates, therefore, parameter estimates for covariates are shown for rGLS fit in results, and p-values for fixed effects are from LRT tests where no GRM is included.

Genetic analyses – Repeated measurements model (rGLS)

The linear mixed model in the first step of the rGLS function uses the estimation algorithm for hierarchical generalized linear models (HGLM) (Lee and Nelder, 1996) implemented in the R package hglm (Rönnegård et al., 2010). HGLM allows for high flexibility where fixed and random effects can come from diverse distributions, but rGLS assumes multivariate Gaussian errors. The first step of rGLS is therefore an animal model (Lynch and Walsh, 1998) fitted via HGLM, where phenotypic variance in BMR is partitioned into genetic- and non-genetic sources of variance. Variance components estimated for BMR include permanent environment variance (V_{pe}), due to differences within individuals caused by fixed effects of environment or non-additive effects, and additive genetic variance (V_a) due to genetic differences between individuals. The GRM was estimated by the GenABEL method (Aulchenko et al. 2007a) implemented in RepeatABEL rGLS function. For a practical review of the animal model methodology, see Wilson et al. (2010). Note that confidence intervals for variance estimates were estimated by the asymptotic normal approximation method, which assumes the sample size is large enough to meet normal critera for the likelihood estimators (Williams et al., 2002). Additionally, the dataset is not balanced (Table 1), where some individuals only have one observation. This causes V_{pe} to be underestimated, although V_a is not affected (Rönnegård, pers. comm. 2016).

The second part of the rGLS model included each marker as a fixed effect on BMR adjusted for relatedness, repeated measurements, and covariates and factors from the first step in rGLS model. P-values of SNP-effects come from Wald-tests. Additive action of SNPs is assumed, where the variance of a SNP (V_{snp}) equals $2pq(a + d(q - p))^2$ (Husby et al., 2015). Dominance deviation d is zero, a is the additive effect, p is MAF and q is the major allele frequency. Visual examination of residuals, and correlations between fixed effects showed

no deviances from model assumptions. Genome-wide significance thresholds was determined by dividing chosen significance value of 0.05 by number of markers ($p = 0.05/183\ 876 = 2.72 \times 10^{-7}$, for pooled analysis). Suggestive thresholds for pooled GWAS was determined by dividing 1 by number of markers allowing for one false positive ($p= 5.44 \times 10^{-6}$) (Husby et al., 2015). The same protocol was used to construct within island GWAS significance thresholds: significance threshold within Leka = 2.75×10^{-7} and suggestive threshold within Leka = $p=5.42 \times 10^{-6}$, significance threshold within Lauvøya = 2.75×10^{-7}, suggestive threshold within Lauvøya = 5.50×10^{-6}, significance threshold within Vega = 2.76×10^{-7} , suggestive threshold within Vega = 5.52×10^{-6}.

Post-hoc mixed models were fitted to visualize the effects of SNPs of interest using the lme4 package in R (Bates et al., 2014), with confidence intervals constructed through Kenward-Roger approximation implemented in the R package pbkrtest (Halekoh and Højsgaard, 2014).

Genome partitioning analysis

Genome partitioning analysis was performed with the genome-wide complex trait analysis v.1.02 (GCTA) software tool (Yang et al., 2011a). BMR measurements from 697 individuals from all three island-populations, and the 183 876 SNPs that passed QC were used in analysis. Additive genetic variance in BMR were estimated for each chromosome using chromosomal GRMs. Similar to rGLS, GCTA assumes random contributions of SNPs to variation in BMR, which are correlated between individuals who share similar genotypes. By fitting chromosomal GRMs as random effects, the particular contribution of a chromosome to BMR variation was estimated. The relationship between chromosome size and variation explained was studied by plotting phenotypic variance in BMR explained by each chromosome against chromosome length. Developers of GCTA denote phenotypic variation explained by SNPs as V_g, wheras rGLS developers which use V_a. Basically both methods estimate and assume additive gene action, but V_g specifically points to the fact that genetic effects are estimated via SNPs, and V_g/V_p is the variance explained by all the genome-wide SNPs (Visscher et al., 2014). V_{g1}/V_p would for instance be the variance in BMR explained by all SNPs on chromosome 1.

The genetic architecture was visually examined to check for correspondence with either a polygenic-, medium-, or large effect QTL -architecture. See Fig. 2 in Robinson et al. (2013) for

simulations of different genetic architecture. A polygenic signature would be represented as an increase in variation explained with increasing size of the chromosome, but if large effect loci are present one would expect those to break up the linear relationship (Robinson et al., 2013). In the GCTA analyses, BMR was controlled for the same covariates and factors as in the rGLS analyses. However, there was no support for repeated observations in this version of GCTA, hence only unique observations were used to control for pseudo-replication. For individuals with repeated observations, the first observation was chosen. Several chromosomes were excluded in order to make the model converge. Non-convergence is due to estimates of variance explained by chromosomes move too far away from their parameter space (0 -1), and they become constrained to a negative value by default (Yang et al., 2011a). Because non-convergence is more likely when chromosomes contain few markers (Santure et al., 2013), the smallest chromosomes where excluded in ascending order until the model converged. This was a compromise to be able to perform the analyses.

BLAST search

The annotated collared flycatcher genome assembly FicAlb_1.4 (Ellegren et al., 2012) was used to see where my SNPs are in relation to known genes. BLAST search was performed through the Ensembl project (Kersey et al., 2015), with a sequence 71 bases long which included the SNP of interest. E-values were used to assert the significance of the match.

Results

Visual examination of the distance transformed IBS matrix indicated genetic sub-structuring of the populations (Fig. 1). Additional k-means clustering analysis (Appendix Table A2, Appendix Fig. A3) showed the same pattern in detail; where cluster 3 corresponds to a mix of all populations, cluster 1 corresponds mostly to Vega individuals and cluster 2 corresponds mostly to Leka individuals, where both cluster 1 and 2 contain translocated individuals in addition to recruits from Lauvøya.

The estimated variance components from the mixed model showed that BMR has significant additive genetic variance, which was larger than the variation within individuals (the permanent environmental effect) (Table 2). Factors and covariates explained a significant amount of variation in BMR (Appendix Table A1). BMR was on average 4.8 mL O_2 h^{-1} (±0.5, $p < 2e\text{-}16$) higher during night measurement than evening measurement. Females had on average 2.4 mL O_2 h^{-1} (±0.6, $p = 8.6 \times 10^{-6}$) higher BMR than males. An increase of one year in age on average reduced BMR by 0.9 mL O_2 h^{-1} (± 0.2, $p = 1.2 \times 10^{-4}$). Finally, an increase in mass by 1 gram increased on average BMR by 3.5 mL O_2 h^{-1} (± 0.2, $p < 2e\text{-}16$).

Pooled genome-wide association analyses

After the covariates and factors had been accounted for, a second model (the GWAS), including SNP effects were fitted. Genomic inflation was estimated as 1.0198 (± 2.09×10^{-5}) which was used to adjust p-values for residual population structure. No SNPs were significant after adjustment for genomic inflation, but SNPa91021 (p-value = 7.71×10^{-7}) and SNPa91020 (p-value = 7.85×10^{-7}) passed the suggestive threshold for pooled GWAS ($p = 5.44 \times 10^{-6}$) (Fig. 2). Table 3 shows the top 10 SNPs from the GWAS scan, where there is almost an order of magnitude difference in p-values between the two top SNPs and the third best SNP.

The suggestive SNPs are on the same scaffold (1931 bp distance between), and linkage analyses shows that they are in strong linkage disequilibrium (LD ; $r^2 = 0.99$). Hence, suggestive SNPs can therefore be considered as proxies for each other. Patterns of LD between SNPs near the two suggestive SNPs is shown in Fig. 3. Fig. 3 shows that SNPs closer to the two suggestive SNPs, in general increase their strength of association with BMR, and

in LD with the top SNP. SNPa90999 (r^2 =0.72) SNPa91032 (r^2=0.71), SNPa91033 (r^2=0.71, shows the highest degree of LD with SNPa91021.

Within islands association analysis

Within island association, analyses gave mixed results. There were no significant or suggestive results within Vega or Lauvøya (Appendix Table A3 and Table A4, Appendix Fig. A4 and Appendix Fig. A5). However, SNPi43117 (P-value = 2.32 x 10^{-7}) on chromosome 5 passed significance within Leka (significance threshold = 2.75 x 10^{-7}). Table 4 shows summary statistics for the top 10 markers from the Leka GWAS scan, with the corresponding Manhattan plot shown in figure 4. SNPi43117 show no strong LD with snips in near vicinity. Variance component estimates from Leka are more uncertain, with V_a = 14.3 (95 % CI = 10.2 - 20.1), V_{pe} = 7.9 (95% CI= 5.4 - 11.6), and V_e = 42.6 (95% CI = 40.5 - 44.7), but not significantly different from the pooled analysis (Table 2). The same is valid for Lauvøya: V_a = 18.4 (95 % CI = 13.0 - 26.1), V_{pe} = 3.4 (95 % CI =1.7 - 6.7), V_e = 52.9 (95 % CI = 44.3 - 63.0), and Vega : V_a = 6.2 (95 % CI = 3.5 - 11.0) , V_{pe}=6.1 (95 % CI = 3.6 - 10.6), V_e = 37.1. (95 % CI = 30.0 - 45.8). SNPa91021 and SNPa91020 where present in all within island analyses but did not reach suggestive or significant thresholds in within-populations analyses.

BLAST results

SNPa91021 and SNPa91020 were blasted against the collared flycatcher genome and was matched with a gene free region approximately 100 kbp away from the TOX3 (TOX High Mobility Group Box Family Member 3) gene (E-value = 4 x 10^{-15}, alignment = 96%). The next best match had an E-value of 0.19. The closest gene to SNPi43117 was SEL1L 300 kbp away (E-value= 4 x 10^{-17}, alignment = 95 %).

SNP effects

The estimated additive effect of SNPa91021 was 2.42 mL $0_2 h^{-1}$ (± 0.484), and this explains 3.5 % of phenotypic variance of BMR. Figure 5 corroborates this effect where the median BMR differs across genotype, with the highest BMR for the "TT" genotype class. Fig. 5 shows also that predicted effects of SNP91021 genotypes on BMR. The estimated additive effect of SNPi43117 was 9.11 mL 0_2 h^{-1} (±1.551), and this SNP explains 8.6% of the phenotypic variance BMR within the Leka population.

Genome partitioning of additive genetic variance

Variance in BMR explained by the different chromosomes is shown in Fig. 6. There was non-significant negative relationship between the variance explained by a chromosome and its length (- 3.1 x 10^{-10} (± 1.7 x 10^{-10}), p = 0.088). Chromosomes 7, 5, 4 and 11 were estimated to explain most variation in BMR, whereas the larger chromosomes 1, 2 and 3 explained little variance. Estimated variance components for each chromosome had high uncertainty and chromosome 1 and 8 had parameter estimates that were constrained to zero.

Discussion

In this study, I found significant additive genetic variance in BMR using genome-wide SNP-markers in three genetically linked island-populations of house sparrows (Table 2). Two suggestive SNPs, in almost complete LD with each other ($r^2 = 0.99$), explaining 3.5 % of variance in BMR within and across populations, were discovered. One SNP explaining 8.6 % of variance in BMR was significant within one island-population (Leka). The suggestive and significant SNPs were not within any known genes, but the two suggestive SNPs in the pooled analysis (SNPa91021 and SNPa91020, from now considered in singular form) were the closest to any known gene, being approximately 100 kbp away from the TOX3 gene. The genomic partitioning suggested oligogenic rather than polygenic inheritance of BMR (Figure 6). There was a trend in genetic architecture corresponding to loci of large or medium effect located on the small and medium sized chromosomes. However, estimates of variance explained by individual chromosomes were highly uncertain (Fig. 6).

Pooled GWAS

The suggestive SNPs (SNPa91021 and SNPa91020, from here considered in singular form) from the pooled GWAS show mixed importance as QTL for BMR. Regarding the fact that the SNP did not reach genome-wide significance, is a red flag when considering future fine-mapping of the region of interest. Corrections for multiple tests applied in the present study assumes independence of tests, which is clearly not accurate when SNPs are in LD (Fig. 3). Hence, the genome wide significance is overly stringent, because it overestimates the effective number of tests performed. The effect size of the suggestive SNP (3.5 %) corresponds to a SNP of large effect (Fig. 5), and might be overestimated (Slate, 2013; Xu, 2003). Nevertheless, the sample size of the present study (N=697 individuals, N=920 records) is larger than similar studies in outbred populations (Schielzeth and Husby, 2014), which should reduce upward bias of SNP effects.

Santure et al. (2015) studied the genetic architecture of life history traits in great tits, based on 5591 SNPs typed on between 416 and 1949 individuals, where their results indicated that the marker density might not have been sufficient for any SNPs to be in strong LD with causal variants. The present study however, has over 30 times the marker density compared to that of Santure et al. (2015). Thus, this study has higher probability of tagging causal variants. The present study contains the largest marker density in any GWAS in a wild outbred vertebrate population to date, with a relatively large sample size (see Table 3 in Schielzeth and Husby, 2014). Besides, average distance between SNPs in these analyses were approximately 6000 bp, which should increase the probability of strong LD between typed SNPs and causal genes. Furthermore, power-simulation studies show that power to detect large effect loci increases markedly with reduction in effective population size (N_e) (Kardos et al., 2015). In general, island-populations are presumed to have low N_e compared to mainland populations. In the present study system along the Norwegian coast, relatively low values of N_e have been documented, including in the populations at Leka and Vega (Baalsrud et al. 2014). Thus, the experimental "hybrid" population on Lauvøya should in theory have even lower N_e and larger linkage blocks, reducing the numbers of marker needed to tag causal loci (Mackay et al., 2009). Other studies in outbred populations, with smaller sample sizes and less marker dense QTL mapping, have discovered QTLs. For instance Husby et. al (2015) used 37 309 SNPs after QC, and discovered significant and suggestive SNPs for clutch size in analyses female collared flycatchers (N=313 records, N = 656 observations). Johnston et al. (2011) mapped QTL for both horn- type (N = 445 observations) and size (N = 160 observations) in soay sheep (*Ovis aries*) typed on 35 831 SNPs (after QC). Although the LD structure differs, comparing the sample size and marker density of the present study with the studies described indicate that large effect loci present for BMR should show up in GWA studies in these populations.

Within islands GWAS

The GWAS on data from the Leka population suggested that one significant SNP (SNPi43117) explained 8.6 % of BMR variation. However, one cannot exclude the possibility that this result may be a false positive. The MAF (0.028) is very low and the effect size estimate high (Table 4) compared to the suggestive SNP from the pooled GWAS (Table 3). The SNPi43117 effect can be upwardly biased due to reduced sample size (Table 1) (Slate, 2013).

Nonetheless, context dependent effects, such as gene *x* environment-interactions (G x E), are documented to be common in mice (*Mus musuculs*), where the content dependent effects can explain more of the variance in phenotype than fixed effects (Valdar et al., 2006). It is plausible that the pooled GWAS averages effects across environments or genetic backgrounds, which may mask the effect of SNPi43117 (Mackay et al., 2009). Hence, the effect of SNPi43117 is only detectable within Leka. Another plausible explanation is that the linkage structure within Leka could have changed due to selection, and that SNPi43117 is only in linkage with a causative variant on Leka.

Challenges of population structure

Interestingly, the population structure analysis reflects the experimental manipulation of the populations. Both cluster analysis (Appendix Table A2, Appendix Fig. A3) and visual inspection of genetic similarity between individuals, corresponds with present knowledge of the populations. The common garden population on Lauvøya seems to be genetically intermediate between the Leka and Vega populations (Fig. 1). As expected, only a subset of translocated individuals established and bred on Lauvøya, which is reflected in Fig. 1, where the individuals on Lauvøya are not uniformly distributed. For example, some individuals translocated from Leka are clustered near recruits from Lauvøya or near individuals from their native population. Parentage analyses determining which individuals that established and bred on Lauvøya might improve genetic analyses and shed light on the effects of the suggestive SNP from the pooled analysis. Considering that sample size is reduced within islands, and power to detect loci is also reduced (Mackay et al., 2009). A linkage mapping approach, similar to the study done by Tarka et al. (2010) on the great reed warblers (*Acrocephalus arundinaceu*), might be able detect genomics regions containing genes explaining variation in BMR with higher power, albeit at the loss of resolution, in the pedigreed mixed population on Lauvøya (Schielzeth and Husby, 2014). Then, correspondence between findings of linkage analysis and pooled GWAS could be investigated.

This study also highlights an important issue regarding population structure and the distribution of phenotype variation in BMR between and islands and between years. The present study assumes that phenotypic differences in BMR between islands and years (Table 1, Appendix Fig. 1) were due to genetic effects (i.e caused by artificial selection). SNPs were

therefore assumed to be correlated with QTLs coding for high or low BMR in Vega or Leka, respectively. However, subpopulations can also differ genetically by chance, i.e. genetic drift. If mean BMR between islands varies proportionally with allele frequencies present in the population by random chance, an association signal between SNPs and trait can be spurious (Marchini et al., 2004). Fitting Island as a fixed factor in the pooled GWAS can control for average differences in BMR between islands, but it will also control for differences due to genes having an additive effect on BMR, and this will cause a reduction in power in the GWAS. Johnston et al. (2014) had a similar population structure in atlantic salmon (*Salmo salar*) in a study on sea age variation, where they applied an F_{st} outlier test to identify differential selection between population clusters of salmon. Applying F_{st} outlier tests in my populations might be a future prospect for this study; in any case, a quantification of how population structure affects GWAS results is an aim in future studies.

SNPs as possible QTLs

Exact knowledge of all genes located on chromosome 11 near the suggestive SNP (SNPa91021/SNPa91020) in house sparrows is yet unknown. In this study, BLAST results showed that the gene with highest proximity to the suggestive SNP is TOX3. TOX3, have been associated with human breast cancer in a previous GWAS study (Easton et al., 2007). TOX3 contains a high-mobility group (HMG-box) (O'Flaherty and Kaye, 2003), which affects DNA modification, and can hence change chromatin structure (Bustin, 1999). In a study by Yuan et al. (2009), TOX3 was identified as a neuronal transcription factor, indicated to affect neuron survival (Dittmer et al., 2011). The function of TOX3 in birds is yet unknown, but the annotated House Sparrow genome (Elgvin et al. ms) will show whether TOX3 is a possible QTL for BMR in house sparrows.

Genome partitioning analysis

The genome partitioning analysis reflects the GWAS results, that large effect loci might be present for BMR (Fig. 6). Chromosome 4, 5, and 11 explain each a substantial part of phenotypic variance in BMR (Fig. 6). These chromosomes have top 10 SNPs in the pooled analysis (Table 3). The relationship between chromosome size and variance in BMR explained by each chromosome (Fig. 6) support BMR having large to moderate effect QTLs (see also Fig. 2 in Robinson et al., 2013). One outlier in the analysis is chromosome 7, which had no top 10 SNPs from the GWAS analyses, except for a non-significant SNP from GWAS

analysis within Lauvøya (Appendix Table A4) with relatively large effect (14.66 mL O_2 h^{-1}). This SNP might describe why chromosome 7 explains a large part of BMR variation. One must also consider that uneven LD across the genome, or clustered SNPs of small effect could give a false signature of large effect loci (Schielzeth and Husby, 2014). Nevertheless, it is interesting that results from GCTA analysis correspond with GWAS scans, and a trend in moderate/large effect QTL genetic architecture is visible. However, the uncertainty of the GCTA estimates were high (Fig. 6), and one cannot exclude the possibility of a polygenic basis for BMR in these house sparrows.

Variance components of BMR

From the estimated variance components of BMR, an approximate heritability of 0.2 (h^2= $V_a/V_a+V_p+V_e$) can be calculated. This estimate disregards covariance between- and uncertainty in variance component estimates, and the heritability estimate should therefore be taken with caution (pers. comm. Rönnegård, 2016). Nevertheless, the estimate for V_a (Table 2) should be unbiased and confirms there is a genetic component of BMR variation in the study populations. For the sake of comparison, estimates of evolvability (Hansen et al., 2011) could be better suited for comparison of variance components from this study with other studies, because it enables comparison of the evolutionary potential of traits in different populations and species (Hansen et al., 2011). An approximate estimate for evolvability can be calculated from Table 1 and Table 2: *(12.4/ 79.4^2) x 100 % = 0.19 %.* This measurement of evolvability predicts a 0.19 % change in mean BMR per generation under one unit strength of selection (Hansen et al., 2011). One unit of selection means that a 1 % change in the trait would yield a 1 % change in fitness. See Hansen et al (2011) for a biological interpretation of evolvability. Hence, the estimated h^2 and evolvability for BMR indicates that the "cost of living" in these house sparrows is an adaptable trait across evolutionary time. Robust calculations of heritability and evolvability with unbiased confidence estimates are elements considered in the future development of the rGLS model (pers. comm., Rönnegård, 2016). Setting this issue aside, the approximate estimate of h^2 for the house sparrows is similar to the estimate from captive zebra finches (*Taeniopygia guttata*) found in Rønning et al. (2007), but lower than the estimate from a natural population of pied flycatchers (*Ficedula hypoleuca*) (Bushuev et al., 2012). Estimates of h^2 for BMR show great range between vertebrate species (White and Kearney, 2013). It is

therefore plausible that the approximate h^2 for BMR in house sparrows is within the range found in other bird species. The same is valid for the approximate evolvability estimate (see Table 1 in Hansen et al., 2011). To my knowledge, this is the first study that estimates genetic variance components for BMR via SNPs, which should be of use in comparative studies of BMR evolvability in the future.

Conclusion

In this study, genetic variation in BMR in three genetically linked house sparrow populations has been explored using next-generation molecular methods. A new genome-wide association procedure utilizing repeated observations was successfully applied for the second time, as first seen used in Husby et. al (2015). Additive genetic variance is for the first time established for BMR in a free-living vertebrate population using SNPs. Another novel result is the discovery of a suggestive SNP of large effect associated with BMR in free-living birds, explaining 3.5 % of variation in BMR in three genetically linked populations. However, further research on phenotype differences between islands and population sub-structuring is needed, in order to clarify the relationship between the suggestive SNP and BMR variation. Additionally, future studies should further examine the genomic region in LD with the suggestive SNP from the present study. The TOX3 gene could possibly be a candidate for a QTL underlying BMR. However, we will have to wait until the positions of genes close to the suggestive SNP will be revealed by the annotated house sparrow genome. Genomic partitioning of additive genetic variance across chromosomes corroborated by GWAS results, lending support to an oligogenic genetic architecture for BMR, where loci of large to medium effect might be present. Such a genetic architecture may allow future studies on the eco-evolutionary dynamics of BMR at the genetic level.

References

Allendorf, F.W., England, P.R., Luikart, G., Ritchie, P. a., Ryman, N., 2008. Genetic effects of harvest on wild animal populations. Trends Ecol. Evol. 23, 327–337.

Aulchenko, Y.S., De Koning, D.J., Haley, C., 2007a. Genomewide rapid association using mixed model and regression: A fast and simple method for genomewide pedigree-based quantitative trait loci association analysis. Genetics 177, 577–585.

Aulchenko, Y.S., Ripke, S., Isaacs, A., van Duijn, C.M., 2007b. GenABEL: An R library for genome-wide association analysis. Bioinformatics 23, 1294–1296.

Bates, D., Mächler, M., Bolker, B., Walker, S., 2014. Fitting Linear Mixed-Effects Models using lme4. arXiv 67, arXiv:1406.5823. doi:10.18637/jss.v067.i01

Berteaux, D., Thomas, D., Bergeron, J., Lapierre, H., 1996. Repeatability of daily field metabolic rate in female Meadow Voles (Microtus pennsylvanicus). Funct. Ecol. 10, 751–759. doi:10.2307/2390510

Bligh, J., Johnson, K.G., 1973. Glossary of terms for thermal physiology. J. Appl. Physiol. 35, 941–961. doi:10.1016/S0306-4565(02)00055-4

Burnham, K.P., Anderson, D.R., 2002. Model selection and multimodel inference: a practical information-theoretic approach, Ecological Modelling. doi:10.1016/j.ecolmodel.2003.11.004

Bushuev, A. V, Husby, A., Sternberg, H., Grinkov, V.G., 2012. Quantitative genetics of basal metabolic rate and body mass in free-living pied flycatchers. J. Zool. 288, 245–251. doi:10.1111/j.1469-7998.2012.00947.x

Bustin, M., 1999. Regulation of DNA-Dependent Activities by the Functional Motifs of the High-Mobility-Group Chromosomal Proteins. Mol. Cell. Biol. 19, 5237–5246. doi:10.1128/MCB.19.8.5237

Colosimo, P.F., Peichel, C.L., Nereng, K., Blackman, B.K., Shapiro, M.D., Schluter, D., Kingsley, D.M., 2004. The genetic architecture of parallel armor plate reduction in threespine sticklebacks. PLoS Biol. 2, 635–641. doi:10.1371/journal.pbio.0020109

Dittmer, S., Kovacs, Z., Yuan, S.H., Siszler, G., Kögl, M., Summer, H., Geerts, A., Golz, S., Shioda, T., Methner, A., 2011. TOX3 is a neuronal survival factor that induces transcription depending on the presence of CITED1 or phosphorylated CREB in the transcriptionally active complex. J. Cell Sci. 124, 252–60. doi:10.1242/jcs.068759

Easton, D., Pooley, K., Dunning, A., Pharoah, P., Thompson, D., Ballinger, D., Struewing, J., Morrison, J., Field, H., Luben, R., Wareham, N., Ahmed, S., Healey, C., Bowman, R., Collaborators, S., Meyer, K., Haiman, C., Kolonel, L., Henderson, B., Le Marchand, L., Brennan, P., Sangrajrang, S., Gaborieau, V., Odefrey, F., Shen, C.-Y., Wu, P.-E., Wang, H.-

C., Eccles, D., Evans, D., Peto, J., Fletcher, O., Johnson, N., Seal, S., Stratton, M., Rahman, N., Chenevix-Trench, G., Bojesen, S., Nordestgaard, B., Axelsson, C., Garcia-Closas, M., Brinton, L., Chanock, S., Lissowska, J., Peplonska, B., Nevanlinna, H., Fagerholm, R., Eerola, H., Kang, D., Yoo, K.-Y., Noh, D.-Y., Ahn, S.-H., Hunter, D., Hankinson, S., Cox, D., Hall, P., Wedren, S., Liu, J., Low, Y.-L., Bogdanova, N., Schürmann, P., Dörk, T., Tollenaar, R., Jacobi, C., Devilee, P., Klijn, J., Sigurdson, A., Doody, M., Alexander, B., Zhang, J., Cox, A., Brock, I., MacPherson, G., Reed, M., Couch, F., Goode, E., Olson, J., Meijers-Heijboer, H., van den Ouweland, A., Uitterlinden, A., Rivadeneira, F., Milne, R., Ribas, G., Gonzalez-Neira, A., Benitez, J., Hopper, J., McCredie, M., Southey, M., Giles, G., Schroen, C., Justenhoven, C., Brauch, H., Hamann, U., Ko, Y.-D., Spurdle, A., Beesley, J., Chen, X., kConFab, Group, A.M., Mannermaa, A., Kosma, V.-M., 2007. Genome-wide association study identifies novel breast cancer susceptibility loci. Nature 447, 1087–1093. doi:10.1038/nature05887

Ekblom, R., Galindo, J., 2011. Applications of next generation sequencing in molecular ecology of non-model organisms. Heredity (Edinb). 107, 1–15. doi:10.1038/hdy.2010.152

Ellegren, H., Sheldon, B.C., 2008. Genetic basis of fitness differences in natural populations. Nature 452, 169–175. doi:10.1038/nature06737

Ellegren, H., Smeds, L., Burri, R., Olason, P.I., Backström, N., Kawakami, T., Künstner, A., Mäkinen, H., Nadachowska-Brzyska, K., Qvarnström, A., Uebbing, S., Wolf, J.B.W., 2012. The genomic landscape of species divergence in Ficedula flycatchers. Nature 491, 756–760. doi:10.1038/nature11584

Hagen, I.J., Billing, A.M., Ronning, B., Pedersen, S.A., Parn, H., Slate, J., Jensen, H., 2013. The easy road to genome-wide medium density SNP screening in a non-model species: development and application of a 10 K SNP-chip for the house sparrow (Passer domesticus). Mol Ecol Resour 13, 429–439. doi:10.1111/1755-0998.12088

Halekoh, U., Højsgaard, S., 2014. A Kenward-Roger Approximation and Parametric Bootstrap Methods for Tests in Linear Mixed Models - The R Package pbkrtest. J. Stat. Softw. 59, 1–32. doi:10.18637/jss.v059.i09

Hansen, T.F., Pélabon, C., Houle, D., 2011. Heritability is not Evolvability. Evol. Biol. 38, 258–277. doi:10.1007/s11692-011-9127-6

Hao, K., Di, X., Cawley, S., 2007. LdCompare: Rapid computation of single- and multiple-marker r2 and genetic coverage. Bioinformatics 23, 252–254. doi:10.1093/bioinformatics/btl574

Hartigan, J.A., Wong, M.A., 1979. Algorithm AS 136: A K-Means Clustering Algorithm. J. R. Stat. Soc. Ser. C (Applied Stat. 28, 100–108. doi:10.2307/2346830

Hill, W.G., 2012. Quantitative genetics in the genomics era. Curr. Genomics 13, 196–206. doi:10.2174/138920212800543110

Hirschhorn, J.N., Daly, M.J., 2005. Genome-wide association studies for common diseases and complex traits. Nat. Rev. Genet. 6, 95–108. doi:10.1038/nrg1521

Hudson, J.W., Kimzey, S.L., 1966. Temperature regulation and metabolic rhythms in populations of the house sparrow (Passer domesticus). Comp. Biochem. Phsiology 17, 203–217.

Hulbert, A.J., Else, P.L., 2000. Mechanisms underlying the cost of living in animals. Annu. Rev. Physiol. 62, 207–235. doi:10.1146/annurev.physiol.62.1.207

Husby, A., Kawakami, T., Rönnegård, L., Smeds, L., Ellegren, H., Qvarnström, A., 2015. Genome-wide association mapping in a wild avian population identifies a link between genetic and phenotypic variation in a life-history trait. Proc. Biol. Sci. 282, 20150156–. doi:10.1098/rspb.2015.0156

Husby, A., Saether, B.E., Jensen, H., Ringsby, T.H., 2006. Causes and consequences of adaptive seasonal sex ratio variation in house sparrows. J. Anim. Ecol. 75, 1128–1139. doi:10.1111/j.1365-2656.2006.01132.x

Jensen, H., Moe, R., Hagen, I.J., Holand, A.M., Kekkonen, J., Tufto, J., Saether, B.E., 2013. Genetic variation and structure of house sparrow populations: is there an island effect? Mol Ecol 22, 1792–1805. doi:10.1111/mec.12226

Jensen, H., Saether, B.E., Ringsby, T.H., Tufto, J., Griffith, S.C., Ellegren, H., 2003. Sexual variation in heritability and genetic correlations of morphological traits in house sparrow (Passer domesticus). J Evol Biol 16, 1296–1307. doi:10.1046/j.1420-9101.2003.00614.x

Jensen, H., Steinsland, I., Ringsby, T.H., Saether, B.E., 2008. Evolutionary dynamics of a sexual ornament in the house sparrow (Passer domesticus): the role of indirect selection within and between sexes. Evolution (N. Y). 62, 1275–1293. doi:10.1111/j.1558-5646.2008.00395.x

Johnston, S.E., McEwan, J.C., Pickering, N.K., Kijas, J.W., Beraldi, D., Pilkington, J.G., Pemberton, J.M., Slate, J., 2011. Genome-wide association mapping identifies the genetic basis of discrete and quantitative variation in sexual weaponry in a wild sheep population. Mol Ecol 20, 2555–2566. doi:10.1111/j.1365-294X.2011.05076.x

Johnston, S.E., Orell, P., Pritchard, V.L., Kent, M.P., Lien, S., Niemelä, E., Erkinaro, J., Primmer, C.R., 2014. Genome-wide SNP analysis reveals a genetic basis for sea-age variation in a wild population of Atlantic salmon (Salmo salar). Mol. Ecol. 23, 3452–3468. doi:10.1111/mec.12832

Kardos, M., Husby, A., Mcfarlane, S.E., Qvarnström, A., Ellegren, H., 2015. Whole-genome resequencing of extreme phenotypes in collared flycatchers highlights the difficulty of detecting quantitative trait loci in natural populations. Mol. Ecol. Resour. 727–741. doi:10.1111/1755-0998.12498

Kersey, P.J., Allen, J.E., Armean, I., Boddu, S., Bolt, B.J., Carvalho-Silva, D., Christensen, M., Davis, P., Falin, L.J., Grabmueller, C., Humphrey, J., Kerhornou, A., Khobova, J., Aranganathan, N.K., Langridge, N., Lowy, E., McDowall, M.D., Maheswari, U., Nuhn, M., Ong, C.K., Overduin, B., Paulini, M., Pedro, H., Perry, E., Spudich, G., Tapanari, E., Walts, B., Williams, G., Tello-Ruiz, M., Stein, J., Wei, S., Ware, D., Bolser, D.M., Howe, K.L., Kulesha, E., Lawson, D., Maslen, G., Staines, D.M., 2015. Ensembl Genomes 2016: more genomes, more complexity. Nucleic Acids Res. 44, 574–580. doi:10.1093/nar/gkv1209

Kierczak, M., Jablonska, J., Forsberg, S.K.G., Bianchi, M., Tengvall, K., Pettersson, M., Scholz, V., Meadows, J.R.S., Jern, P., Carlborg, Ø., Lindblad-Toh, K., 2015. Cgmisc: Enhanced genome-wide association analyses and visualization. Bioinformatics 31, 3830–3831. doi:10.1093/bioinformatics/btv426

Konarzewski, M., Ksiazek, A., 2013. Determinants of intra-specific variation in basal metabolic rate. J. Comp. Physiol. B Biochem. Syst. Environ. Physiol. 183, 1–26. doi:10.1007/s00360-012-0676-5

Kvalnes, T., Ringsby, T., Jensen, H., Sæther, B.-E., 2013. Correlates of egg size variation in a population of house sparrow Passer domesticus. Oecologia 171, 391–402. doi:10.1007/s00442-012-2437-2

Lee, Y., Nelder, J.A., 1996. Hierarchical Generalized Linear Models. J. R. Stat. Soc. Ser. B 58, 619–678.

Lynch, M., Walsh, B., 1998. Genetics and analysis of quantitative traits, Genetics and analysis of quantitative traits. Sinauer Associates, Inc. {a}, 108 North Main Street, Sunderland, Massachusetts 01375, USA.

Mackay, T.F.C., Stone, E.A., Ayroles, J.F., 2009. The genetics of quantitative traits: challenges and prospects. Nat Rev Genet 10, 565–577.

Marchini, J., Cardon, L.R., Phillips, M.S., Donnelly, P., 2004. The effects of human population structure on large genetic association studies. Nat. Genet. 36, 512–517. doi:10.1038/ng1337

Mardia, K. V, 1978. Some properties of clasical multi-dimesional scaling. Commun. Stat. Methods 7, 1233–1241.

Mardis, E.R., 2008. The impact of next-generation sequencing technology on genetics. Trends Genet. 24, 133–141. doi:10.1016/j.tig.2007.12.007

O'Flaherty, E., Kaye, J., 2003. TOX defines a conserved subfamily of HMG-box proteins. BMC Genomics 4, 13. doi:10.1186/1471-2164-4-13

Page, A.J., Cooper, C.E., Withers, P.C., 2011. Effects of experiment start time and duration on measurement of standard physiological variables. J. Comp. Physiol. B Biochem. Syst. Environ. Physiol. 181, 657–665. doi:10.1007/s00360-011-0551-9

Pärn, H., Jensen, H., Ringsby, T.H., Sæther, B.E., 2009. Sex-specific fitness correlates of dispersal in a house sparrow metapopulation. J. Anim. Ecol. 78, 1216–1225. doi:10.1111/j.1365-2656.2009.01597.x

Pärn, H., Ringsby, T.H., Jensen, H., Sæther, B.E., 2012. Spatial heterogeneity in the effects of climate and density-dependence on dispersal in a house sparrow metapopulation. Proc. R. Soc. B-Biological Sci. 279, 144–152. doi:10.1098/rspb.2011.0673

R Core Team, 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/. ISBN 3–900051–07–0, URL http://www.R–project.org/.

Ringsby, T.H., Saether, B.E., Tufto, J., Jensen, H., Solberg, E.J., 2002. Asynchronous spatiotemporal demography of a house sparrow metapopulation in a correlated environment. Ecology 83, 561–569. doi:10.2307/2680035

Robinson, M.R., Santure, A.W., Decauwer, I., Sheldon, B.C., Slate, J., 2013. Partitioning of genetic variation across the genome using multimarker methods in a wild bird population. Mol. Ecol. 22, 3963–3980. doi:10.1111/mec.12375

Roff, D.A., 2007. A centennial celebration for quantitative genetics. Evolution (N. Y). 61, 1017–1032. doi:10.1111/j.1558-5646.2007.00100.x

Rönnegård, L., McFarlane, E., Husby, A., Kawakami, T., Ellegren, H., Qvarnström, A., Sciences, C., Sci-, A., Ecology, A., Centre, E.B., Dynamics, B., Uni-, N., 2016. Increasing the power of genome wide association studies in natural populations using repeated measures evaluation and implementation. doi:10.1111/2041-210X.12535

Rönnegård, L., Shen, X., Alam, M., 2010. hglm: A Package for Fitting Hierarchical Generalized Linear Models. R J. 2, 20–28.

Rønning, B., Broggi, J., Bech, C., Moe, B., Ringsby, T.H., Pärn, H., Hagen, I.J., Sæther, B.E., Jensen, H., 2015. Is basal metabolic rate associated with recruit production and survival in free-living house sparrows? Funct. Ecol. 1–9. doi:10.1111/1365-2435.12597

Rønning, B., Jensen, H., Moe, B., Bech, C., 2007. Basal metabolic rate: heritability and genetic correlations with morphological traits in the zebra finch. J Evol Biol 20, 1815–1822. doi:10.1111/j.1420-9101.2007.01384.x

Rønning, B., Moe, B., Berntsen, H.H., Noreen, E., Bech, C., 2014. Is the Rate of Metabolic Ageing and Survival Determined by Basal Metabolic Rate in the Zebra Finch? PLoS One 9, e108675. doi:10.1371/journal.pone.0108675

Santure, A.W., De Cauwer, I., Robinson, M.R., Poissant, J., Sheldon, B.C., Slate, J., 2013. Genomic dissection of variation in clutch size and egg mass in a wild great tit (Parus major) population. Mol. Ecol. 22, 3949–3962. doi:10.1111/mec.12376

Santure, A.W., Poissant, J., De Cauwer, I., van Oers, K., Robinson, M.R., Quinn, J.L., Groenen, M. a M., Visser, M.E., Sheldon, B.C., Slate, J., 2015. Replicated analysis of the genetic architecture of quantitative traits in two wild great tit populations. Mol. Ecol. 6148–6162. doi:10.1111/mec.13452

Sax, K., 1923. The Association of Size Differences with Seed-Coat Pattern and Pigmentation in Phaseolus vulgaris. Genetics 8, 552–560.

Schielzeth, H., Husby, A., 2014. Challenges and prospects in genome-wide quantitative trait loci mapping of standing genetic variation in natural populations. Ann. N. Y. Acad. Sci. 1320, 35–57. doi:10.1111/nyas.12397

Skjelseth, S., Ringsby, T.H., Tufto, J., Jensen, H., Saether, B.E., 2007. Dispersal of introduced house sparrows Passer domesticus: an experiment. Proc. R. Soc. B-Biological Sci. 274, 1763–1771. doi:10.1098/rspb.2007.0338

Slate, J., 2013. From beavis to beak color: A simulation study to examine how much QTL mapping can reveal about the genetic architecture of quantitative traits. Evolution (N. Y). 67, 1251–1262. doi:10.1111/evo.12060

Slate, J., 2005. Quantitative trait locus mapping in natural populations: Progress, caveats and future directions. Mol. Ecol. 14, 363–379. doi:10.1111/j.1365-294X.2004.02378.x

Slate, J., Santure, A.W., Feulner, P.G.D., Brown, E.A., Ball, A.D., Johnston, S.E., Gratten, J., 2010. Genome mapping in intensively studied wild vertebrate populations. Trends Genet. 26, 275–284. doi:10.1016/j.tig.2010.03.005

Tarka, M., Akesson, M., Beraldi, D., Hernández-Sánchez, J., Hasselquist, D., Bensch, S., Hansson, B., 2010. A strong quantitative trait locus for wing length on chromosome 2 in a wild population of great reed warblers. Proc. Biol. Sci. 277, 2361–2369. doi:10.1098/rspb.2010.0033

Valdar, W., Solberg, L.C., Gauguier, D., Cookson, W.O., Rawlins, J.N.P., Mott, R., Flint, J., 2006. Genetic and environmental effects on complex traits in mice. Genetics 174, 959–984. doi:10.1534/genetics.106.060004

Versteegh, M. a., Helm, B., Dingemanse, N.J., Tieleman, B.I., 2008. Repeatability and individual correlates of basal metabolic rate and total evaporative water loss in birds: A case study in European stonechats. Comp. Biochem. Physiol. - A Mol. Integr. Physiol. 150, 452–457. doi:10.1016/j.cbpa.2008.05.006

Visscher, P.M., Hemani, G., Vinkhuyzen, A. a E., Chen, G.B., Lee, S.H., Wray, N.R., Goddard, M.E., Yang, J., 2014. Statistical Power to Detect Genetic (Co)Variance of Complex Traits Using SNP Data in Unrelated Samples. PLoS Genet. 10. doi:10.1371/journal.pgen.1004269

White, C.R., Kearney, M.R., 2013. Determinants of inter-specific variation in basal metabolic rate. J. Comp. Physiol. B Biochem. Syst. Environ. Physiol. 183, 1–26. doi:10.1007/s00360-012-0676-5

Wiersma, P., Munoz-Garcia, A., Walker, A., Williams, J.B., 2007. Tropical birds have a slow pace of life. Proc. Natl. Acad. Sci. U. S. A. 104, 9340–9345. doi:10.1073/pnas.0702212104

Williams, B.K., Nichols, J.D., Conroy, M.J., 2002. Analysis and management of animal populations, Academic Press New York.

Wilson, A.J., Réale, D., Clements, M.N., Morrissey, M.M., Postma, E., Walling, C. a., Kruuk, L.E.B., Nussey, D.H., 2010. An ecologist's guide to the animal model. J. Anim. Ecol. 79, 13–26. doi:10.1111/j.1365-2656.2009.01639.x

Xu, S.Z., 2003. Theoretical basis of the Beavis effect. Genetics 165, 2259–2268.

Yang, J., Lee, S.H., Goddard, M.E., Visscher, P.M., 2011a. GCTA: A tool for genome-wide complex trait analysis. Am. J. Hum. Genet. 88, 76–82. doi:10.1016/j.ajhg.2010.11.011

Yang, J., Manolio, T.A., Pasquale, L.R., Boerwinkle, E., Caporaso, N., Cunningham, J.M., de Andrade, M., Feenstra, B., Feingold, E., Hayes, M.G., Hill, W.G., Landi, M.T., Alonso, A., Lettre, G., Lin, P., Ling, H., Lowe, W., Mathias, R.A., Melbye, M., Pugh, E., Cornelis, M.C., Weir, B.S., Goddard, M.E., Visscher, P.M., 2011b. Genome partitioning of genetic variation for complex traits using common SNPs. Nat. Genet. 43, 519–U44. doi:10.1038/ng.823

Tables

Table 1 Sample sizes and BMR statistics for genetic analyses performed within populations and for pooled analysis. Founders at Lauvøya include 62 individuals from Leka and 64 individuals from Vega. Means and standard errors are from mixed models where only intercept and id (as a random factor) were fitted. N is total number of observations, id is the number of unique individuals, mean and variance obs. is the mean and variance of BMR for observations, respectively.

Sample	Leka	Vega	Lauvøya	Pooled analysis
N	362	225	333	920
id	267	172	258	697
Mean obs.	1.36	1.31	1.29	0.345
Variance obs.	0.359	0.308	0.355	1.32
Mean BMR	76.6 (± 0.5)	83.1 (± 0.6)	80.1 (± 0.6)	79.4 (± 0.3)

Table 2 Variance component estimates for BMR (mL O_2 h^{-1}) from repeated measurement model (rGLS). Estimates and their 95 % confidence intervals (CI) are presented. BMR was measured on 697 individuals with 920 observations from three genetically connected island populations of House Sparrows. BMR was controlled for for fixed effects of sex, age, period of measurement and mass. Variance in BMR was decomposed into variance due to genetic (V_a) and environmental (V_{pe}) differences between individuals. V_a and V_{pe} were estimated by by fitting the genomic relatedness matrix and animal identity as random effects, respectively. V_e is residual variance not explained by covariates or random effects. Confidence intervals were estimated by asymptotic normal approximation methods.

Variance components	Estimate	CI (95 %)
V_a (additive genetic variance)	12.4	9.74 - 15.7
V_{pe} (permanent environment variance)	6.64	5.02 - 8.78
V_e (residual variance)	42.2	38.0 - 47.0

Table 3 Summary statistics for the top 10 SNPs associated with BMR (mL $O_2 h^{-1}$) from repeated measurement GWA scan using 920 observations from 697 individuals from three genetically linked island-populations of house sparrows. For each SNP the table shows their chromosome, position within scaffolds, the reference allele A1, the effect allele A2, the minor allele frequency (MAF), estimated effect size of allele A2 with standard error, p-values adjusted for genomic inflation. SNP names in bold pass the pooled analysis suggestive threshold ($p<5.44 \times 10^{-6}$).

SNP-name	chromosome	position	A1	A2	MAF	effect (SE)	p-value	callrate
SNPa91021	**11**	**2987835**	**G**	**T**	**0.248**	**2.42 (± 0.484)**	**7.71×10^{-7}**	**1.000**
SNPa91020	**11**	**2985904**	**C**	**T**	**0.247**	**2.42 (± 0.486)**	**7.85×10^{-7}**	**1.000**
SNPa53531	5	10214309	G	A	0.096	3.18 (± 0.716)	1.10×10^{-5}	0.999
SNPa222365	4	7761289	G	A	0.460	1.74 (± 0.404)	2.06×10^{-5}	0.997
SNPa159077	5	4405934	A	G	0.138	2.61 (± 0.607)	2.11×10^{-5}	0.999
SNPa70021	8	10303911	C	T	0.081	-3.30 (± 0.783)	3.06×10^{-5}	1.000
SNPa91033	11	3001742	C	T	0.194	2.20 (± 0.528)	3.71×10^{-5}	0.996
SNPa91032	11	3000916	A	G	0.194	2.17 (± 0.527)	4.47×10^{-5}	1.000
SNPa121703	4	5392680	T	C	0.494	1.74 (± 0.428)	5.49×10^{-5}	0.996
SNPa520269	4	961417	G	A	0.326	1.89 (± 0.467)	5.91×10^{-5}	0.994

Table 4 Summary statistics for the top 10 SNPs associated with BMR (mL O_2h^{-1}) from repeated measurement GWAS using 362 observations from 267 individuals from Leka. For each SNP the table shows their chromosome, position within scaffolds, the reference allele A1, the effect allele A2, the minor allele frequency (MAF), estimated effect size of allele A2 with standard error, p-values adjusted for genomic inflation. SNP names in bold pass the bonferroni significance threshold within Leka (p=2.75 x 10-7).

SNP-name	chromosome	position	A1	A2	MAF	effect (SE)	p-value	callrate
SNPi43117	**5**	**418265**	**A**	**G**	**0.028**	**8.67 (± 1.638)**	**2.32×10^{-7}**	**1.000**
SNPa340652	29	3391002	C	T	0.075	4.96 (± 1.144)	2.29×10^{-5}	1.000
SNPa287941	9	5026343	C	A	0.165	3.47 (± 0.808)	2.67×10^{-5}	0.989
SNPi32953	1	2006319	A	G	0.077	4.24 (± 0.991)	2.89×10^{-5}	0.993
SNPa91720	11	4035219	C	T	0.163	-3.59 (± 0.840)	2.96×10^{-5}	1.000
SNPa67428	8	4546553	C	T	0.400	-2.67 (± 0.633)	3.67×10^{-5}	0.993
SNPa402163	3	768362	C	T	0.212	-3.24 (± 0.769)	3.84×10^{-5}	0.996
SNPa165502	18	5581805	G	A	0.153	3.55 (± 0.848)	4.40×10^{-5}	0.978
SNPa495505	30	1034	C	A	0.276	-3.04 (± 0.728)	4.46×10^{-5}	0.996
SNPa339958	29	2331321	C	T	0.483	-2.84 (± 0.683)	5.00×10^{-5}	0.993

Figures

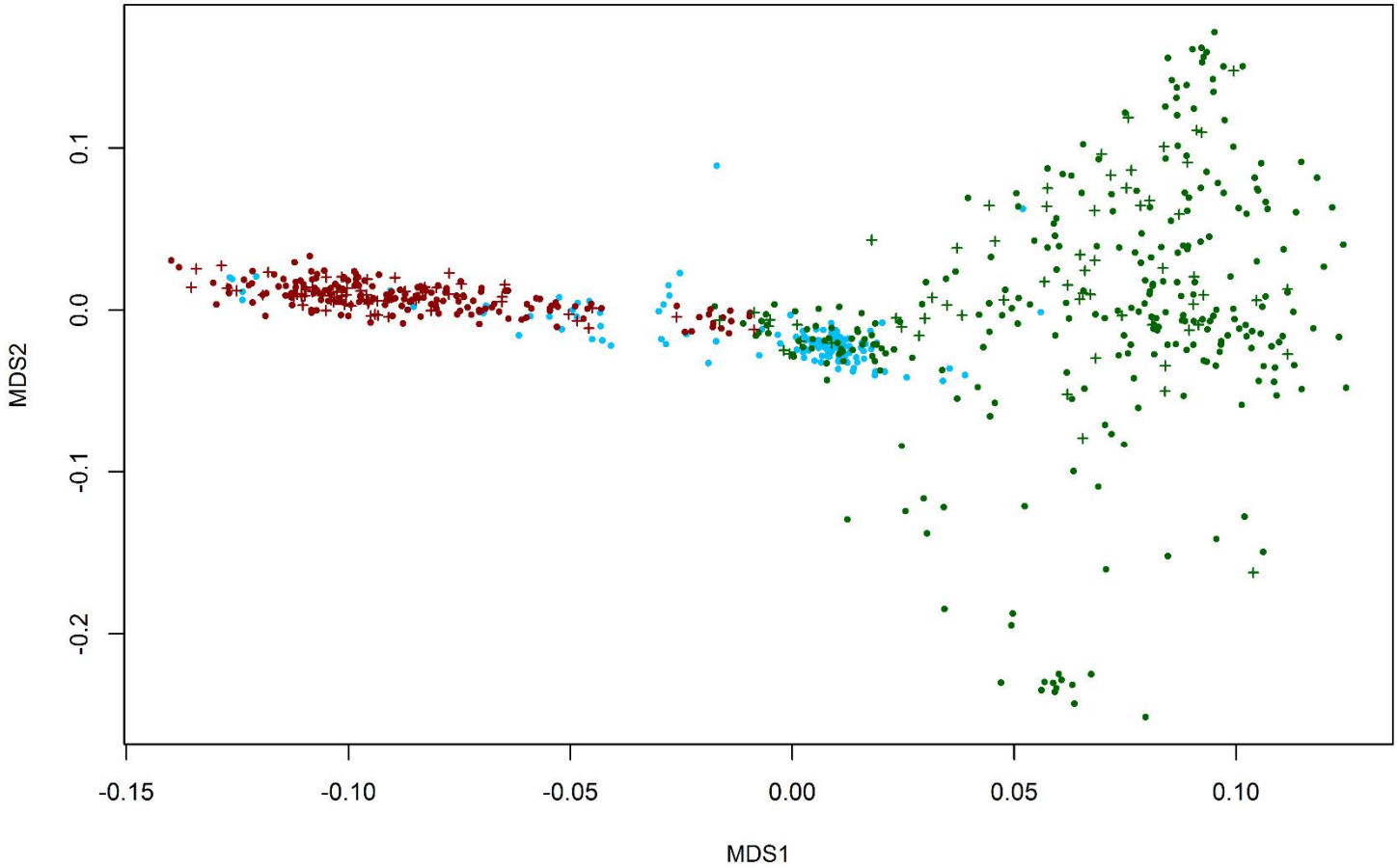

Fig. 1 Multidimensional scaling of the genomic relatedness matrix based on 697 individuals typed in 183 876 SNPs. Individuals originate from three genetically linked island populations off the coast of mid-Norway. Coloring is the population identifier: Leka =Green, Vega=Red, Lauvøya = Blue. Crosses are individuals that were translocated from either Leka (green cross) or Vega (red cross) to Lauvøya in a selection experiment (see main text). Dots are either individuals released back on the islands (Leka =red dots, Vega = green dots), or recruits from Lauvøya (blue dots).

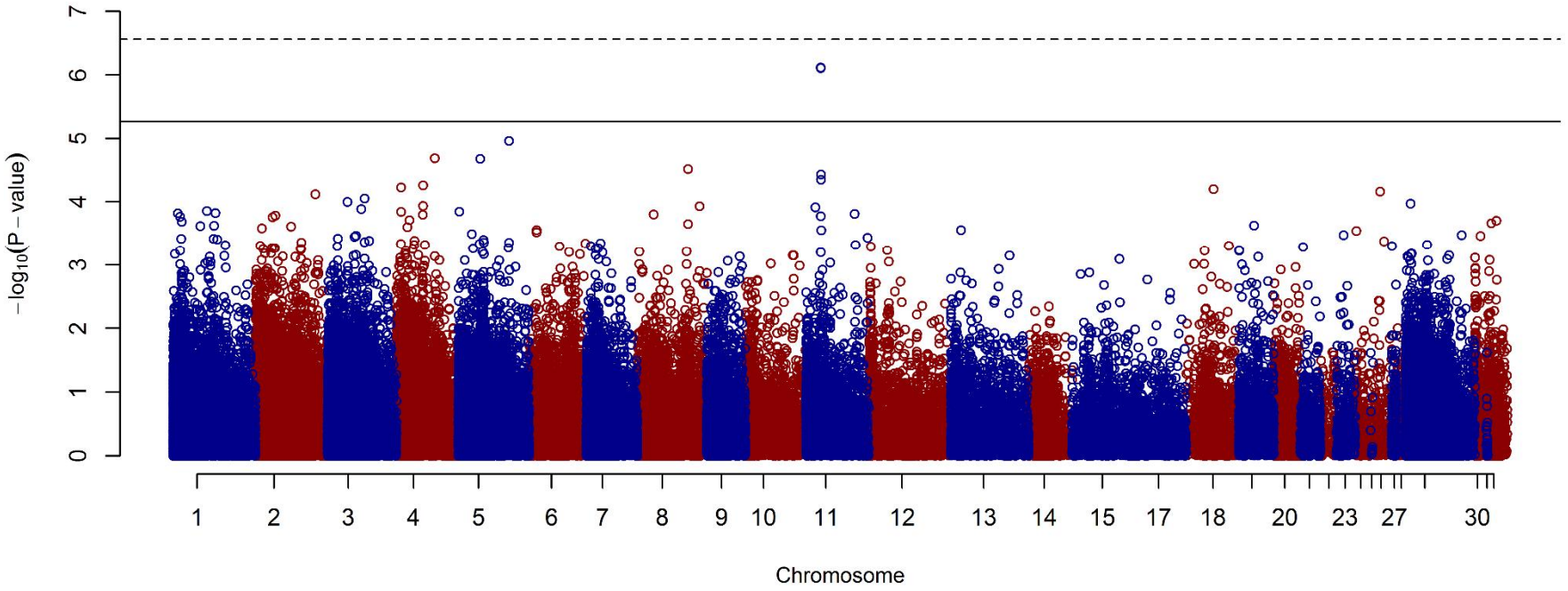

Fig. 2 Manhattan plot of pooled analysis repeated measures GWA scan for SNPs associated with BMR (mL O_2h^{-1}) in 697 individuals (N=920 observations) from three genetically connected island-populations of House Sparrows. 183 876 SNPs are show with their chromosome (SNPs positions are within scaffolds, see main text) labelled on the x-axis and the negative-$\log_{10}$ p-values on the y-axis. No SNPs were included from chromosome 16, and chromosomes; 19, 21, 22, 24, 25, 26, 28, 29, 31, and 32 (Z) are not labelled. The dashed line is the genome-wide significance threshold for the pooled analysis($p=2.72 \times 10^{-7}$), and the solid line shows the suggestive threshold for the pooled analysis ($p=5.44 \times 10^{-6}$).

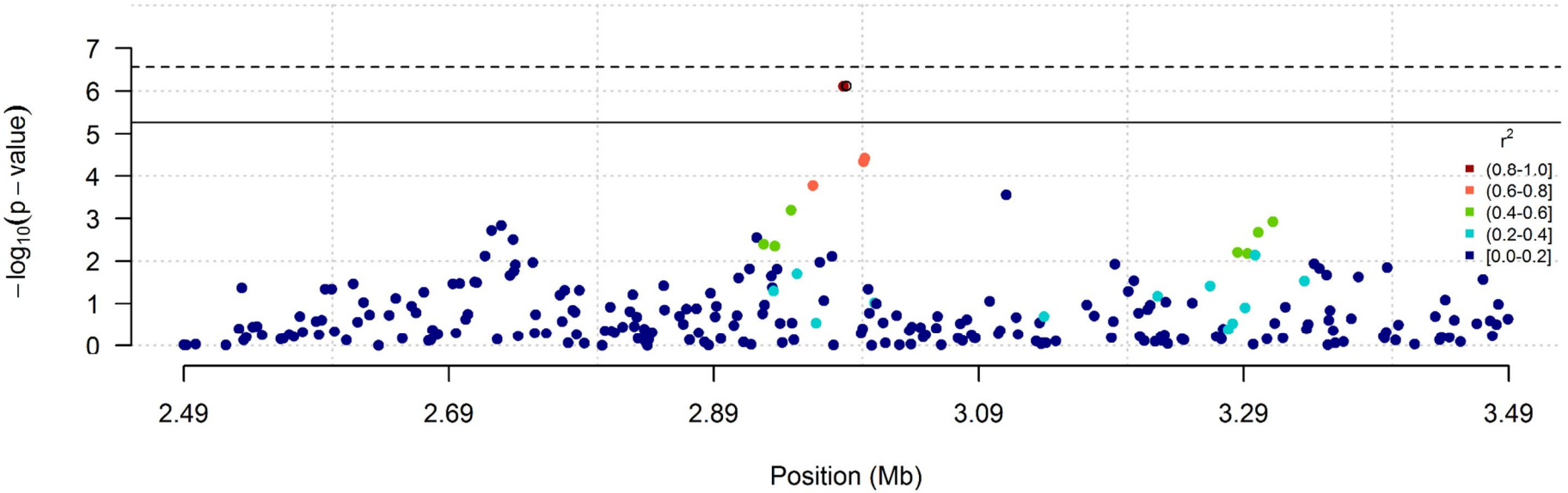

Fig. 3 Manhattan-linkage disequilbirum plot for a one million basepair (Mb) long region on chromosome 11 (within the same scaffold, see main text) containing the two top SNPs from the GWA scan for BMR. Results are from repeated measures GWAS scan for SNPs associated with BMR (mL 0_2h^{-1}) in 697 individuals (N=920 observations) from three genetically connected island-populations of house sparrows. Linkage disequilibrium (r^2) between SNPs and reference SNPa91021 is color coded after degree of LD, where r^2 color index is shown on right hand side of figure (SNPa91021 is the open black circle, overlapping with SNPa91020). Negative-$\log_{10}$ p-values are on the y-axis, where the dashed line is the genome-wide significance threshold for the pooled populations ($p=2.72 \times 10^{-7}$), and the solid line indicates the suggestive threshold for the pooled populations ($p=5.44 \times 10^{-6}$).

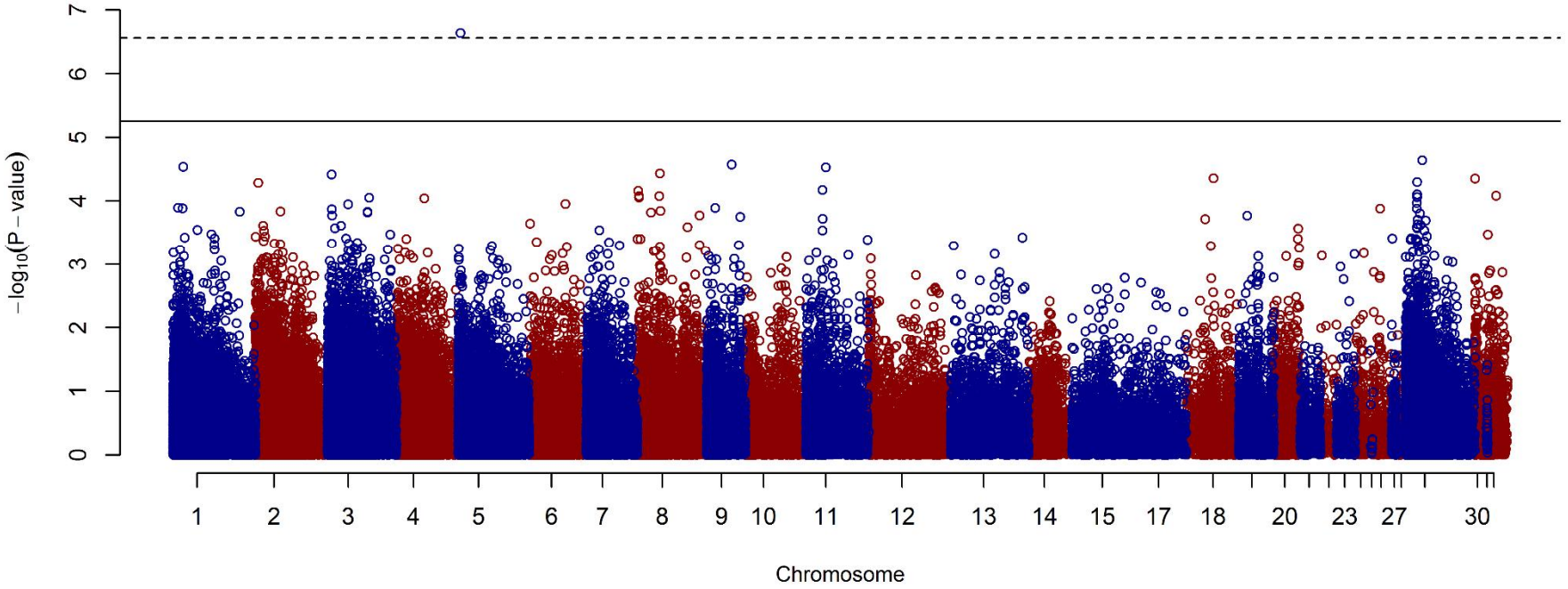

Fig. 4 Manhattan plot of repeated measures GWA scan for SNPs associated with BMR (mL O_2 h^{-1}) in 267 (N=362 observations) House Sparrows from Leka. 181 298 SNPs are shown with their chromosome labelled on the x-axis (SNPs positions are within scaffolds, see main text) and negative-log p-values on the y-axis. No SNPs were included from chromosome 16. Chromosomes; 19, 21, 22, 24, 25, 26, 28, 29, 31, and 32 (Z) are not labelled. Dashed line is the genome-wide significance threshold within Leka ($p=2.75 \times 10^{-7}$), and the solid line indicates the suggestive threshold within Leka ($p=5.42 \times 10^{-6}$).

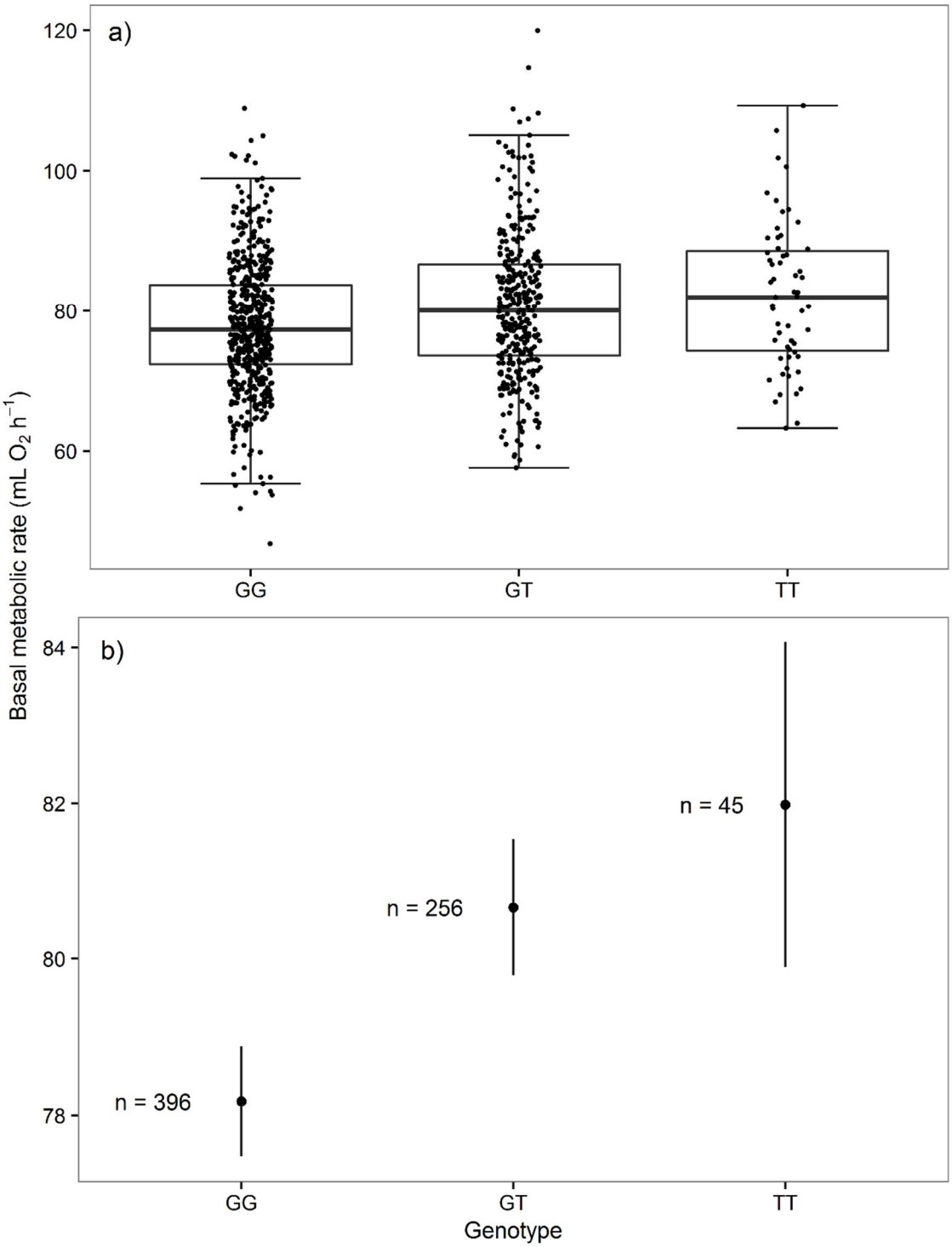

Fig. 5 a) Box-and-whisker plot of natural variation in BMR (mL O_2 h^{-1}) for all observations of individuals with different genotypes at the top SNP (SNPa91021) from the GWAS on the pooled population data. Median values (Q2) are shown as solid horizontal lines in the boxes, the bottom and top of boxes show the 25th (Q1) and 75th (Q3) percentile of data, respectively. Whiskers show max/min value or max/min points that are less than 1.5 times the interquantile range (Q3-Q1), which is approximately two standard deviations. b) Effect plot of the three genotype classes of SNPa91021 on BMR (mL O_2 h^{-1}) with 95 % confidence interval (CI). Predicted means and CI are computed through Kenward-Roger approximation a mixed model with fixed effect of SNP, fixed factors and covariates of age (years), sex, period of measurement (evening and night) and mass (grams) with individual identity as random effect. n is the number of individuals in genotype class.

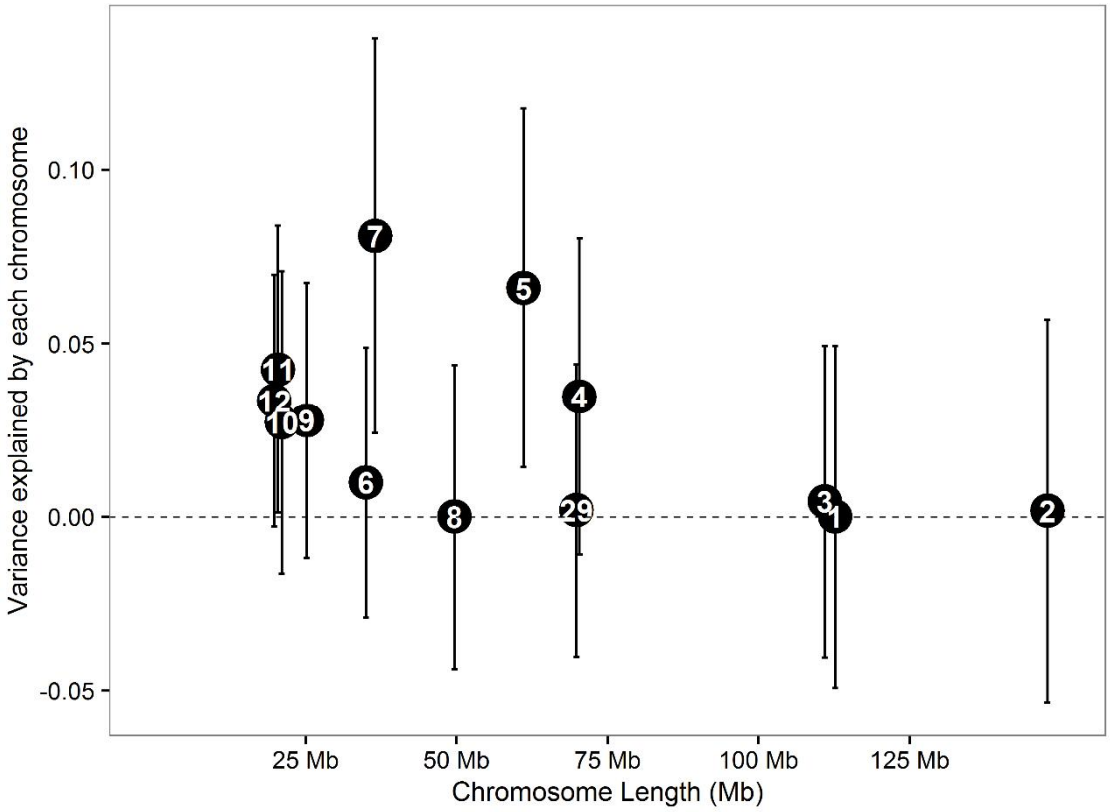

Fig. 6 Relationship between variance in BMR (mL O_2 h^{-1}) explained and chromosome size expressed in million basepairs (Mb). 697 individuals from three genetically linked island populations of house sparrows typed on 183 876 SNPs were used to partition additive genetic variance in BMR across chromosomes. Standard error is shown as whiskers for each labelled chromosome.

Appendix Tables

Table A1. Summary of likelihood ratio (LR) tests of factors and covariates explaining BMR (mL O_2 h^{-1}) Factors and covariates include; sex, a factor with two levels (males and females), period of measurement is a factor with two levels corresponding to evening and night measurements, age (years), and mass (grams). Mixed models were fitted by maximum likelihood with animal identity as a random factor. LRT-p-values are shown for each factor and covariate. DF = degrees of freedom. See main text for details on covariates and LRT.

Predictors	Log-likelihood ratio	DF	LRT-p-value
Age	14.866	1	p = 0.0001154
Sex	19.823	1	$P < 8.6 \times 10^{-6}$
Period	84. 991	1	p < 2.2 x-10^{-16}
Mass	384.74	1	P < 2.2 x-10^{-16}

Table A2. Summary table of k-means clustering analysis (clusters = 3) of the distance transformed genomic relatedness matrix for 697 individuals from three populations typed at 183876 SNPs. The Leka and Vega populations include only individuals that were released back on their respective islands during the selection and translocation experiment (see main text). The Lauvøya population consists of individuals translocated to the island and their recruits. Numbers in parentheses are the numbers of translocated individuals assigned to each cluster, whereas the remaining individuals in the Lauvøya population are recruits from Lauvøya assigned to each cluster.

	Cluster 1	Cluster 2	Cluster 3
Leka	0	176	91
Vega	157	0	15
Lauvøya	83 (62)	46 (43)	129 (21)

Table A3. Summary statistics for the top 10 markers associated with BMR(mL O_2 h^{-1}) from repeated measurement GWAS analysis using 225 observations from 172 individuals from Vega. For each SNP the table shows their chromosome, position within scaffolds, the reference allele A1, the effect allele A2, the minor allele frequency (MAF), estimated effect size of allele A2 with standard error, p-values adjusted for genomic inflation. No SNPs pass the genome-wide significance threshold within Vega ($p=2.76 \times 10^{-7}$).

SNP-name	chromosome	position	A1	A2	MAF	effect (SE)	p-value	callrate
SNPa296177	5	3671235	G	T	0.265	3.69 (± 0.818)	9.44×10^{-6}	1.000
SNPa222365	4	7761289	G	A	0.491	2.92 (± 0.659)	1.30×10^{-5}	0.992
SNPa221808	4	6915431	C	A	0.220	-3.62 (± 0.825)	1.60×10^{-5}	1.000
SNPa400839	9	1805070	A	C	0.407	2.99 (± 0.690)	2.09×10^{-5}	1.000
SNPa253598	29	6890225	G	T	0.119	-4.56 (± 1.074)	2.94×10^{-5}	1.000
SNPa56615	29	913883	T	C	0.083	5.16 (± 1.215)	2.95×10^{-5}	0.996
SNPa520255	4	938848	C	T	0.237	3.44 (± 0.817)	3.55×10^{-5}	1.000
SNPa254479	29	8222619	A	G	0.095	-5.16 (± 1.241)	4.35×10^{-5}	1.000
SNPa408985	3	1223497	A	G	0.271	-3.00 (± 0.721)	4.40×10^{-5}	1.000
SNPa259353	4	7482732	G	A	0.197	3.63 (± 0.875)	4.52×10^{-5}	1.000

Table A4. Summary statistics for the top 10 markers associated with BMR(mL O_2 h^{-1}) from repeated measurement GWAS analysis using 333 observations on 258 individuals from Lauvøya. For each SNP the table shows their chromosome, position within scaffolds, the reference allele A1, the effect allele A2, the minor allele frequency (MAF), estimated effect size of allele A2 with standard error, p-values adjusted for genomic inflation. SNP names in bold pass the genome-wide significance threshold within Lauvøya ($p=2.75 \times 10^{-7}$).

SNP-name	chromsome	position	A1	A2	MAF	effect (SE)	p-value	callrate
SNPa373199	7	3705623	C	A	0.232	14.66 (± 3.245)	1.07×10^{-5}	0.994
SNPa481642	20	340019	C	T	0.268	4.94 (± 1.114)	1.53×10^{-5}	0.994
SNPa95001	11	10039130	G	A	0.444	-4.43 (± 1.016)	2.17×10^{-5}	0.994
SNPa147493	6	9294894	A	G	0.366	6.15 (± 1.444)	3.35×10^{-5}	0.994
SNPa12575	12	3770892	A	G	0.376	4.26 (± 1.006)	3.59×10^{-5}	0.994
SNPa338847	29	538758	G	T	0.055	-4.94 (± 1.193)	5.49×10^{-5}	1.000
SNPa449664	2	215512	C	A	0.380	-5.04 (± 1.219)	5.55×10^{-5}	1.000
SNPa338852	29	549135	G	T	0.312	-5.17 (± 1.266)	7.03×10^{-5}	1.000
SNPa147421	6	9199435	G	A	0.201	7.55 (± 1.865)	7.94×10^{-5}	1.000
SNPa132401	10	9611318	G	A	0.386	3.56 (± 0.879)	8.09×10^{-5}	1.000

Appendix Figures

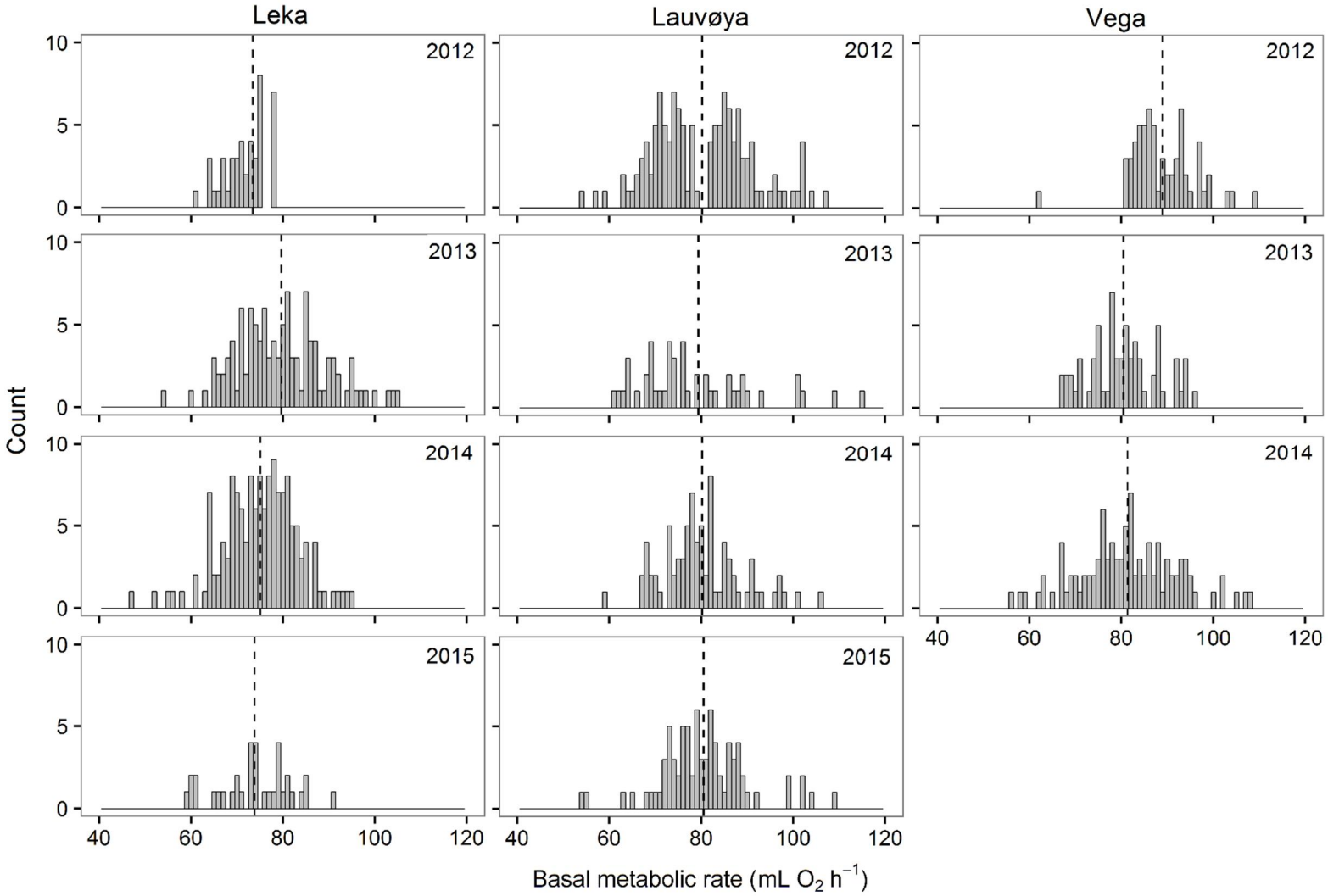

Fig. A1 Histogram over BMR measures across years within Leka, Vega, and Lauvøya. Column width is 1 mL O_2 h^{-1}. Dashed lines are means within islands and years.

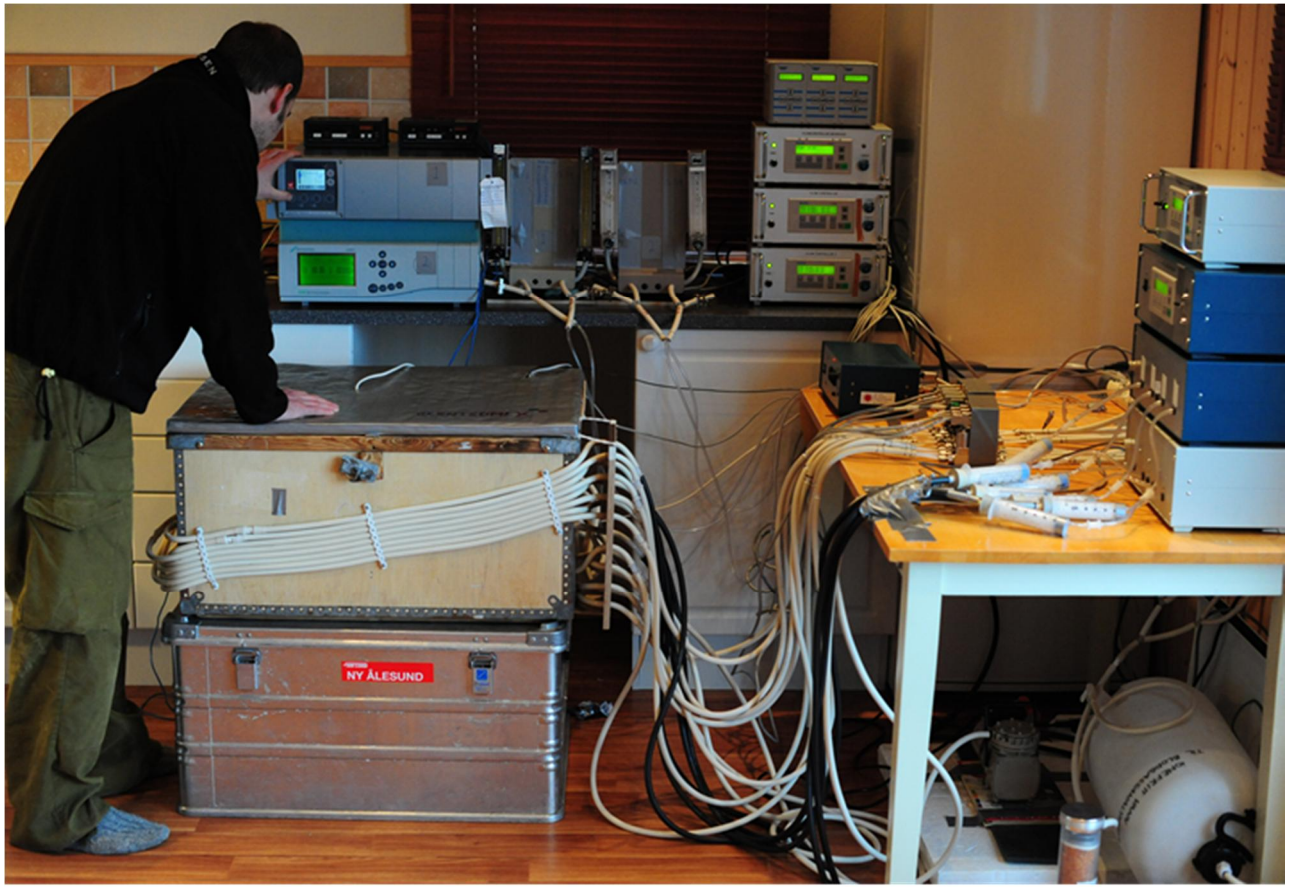

Fig. A2 Shows the custom made climate chamber used for collecting BMR data of eight individual house sparrows, housed within eight smaller respirometer chambers, for estimating O_2 consumption (mL O_2 h^{-1}). See main text for details. Photo: Henrik Jensen.

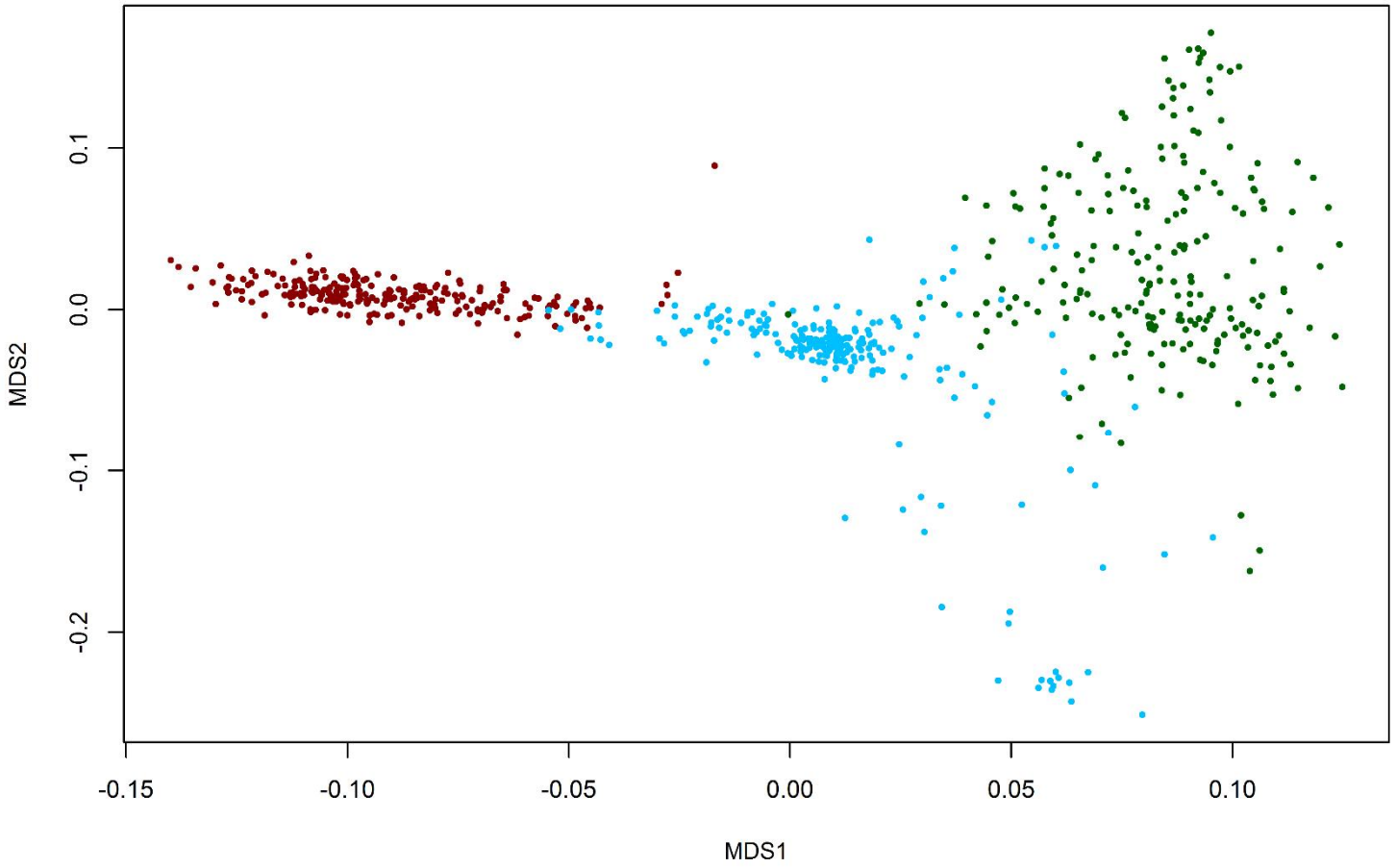

Fig. A3 Multidimensional scaling of the genomic relatedness matrix based on 697 individuals and 183 876 SNPs. Individuals originate from three island populations off the coast of mid-Norway. Coloring shows clusters identified by the k-means method with numbers of centers set to three. See Appendix Table A2 for

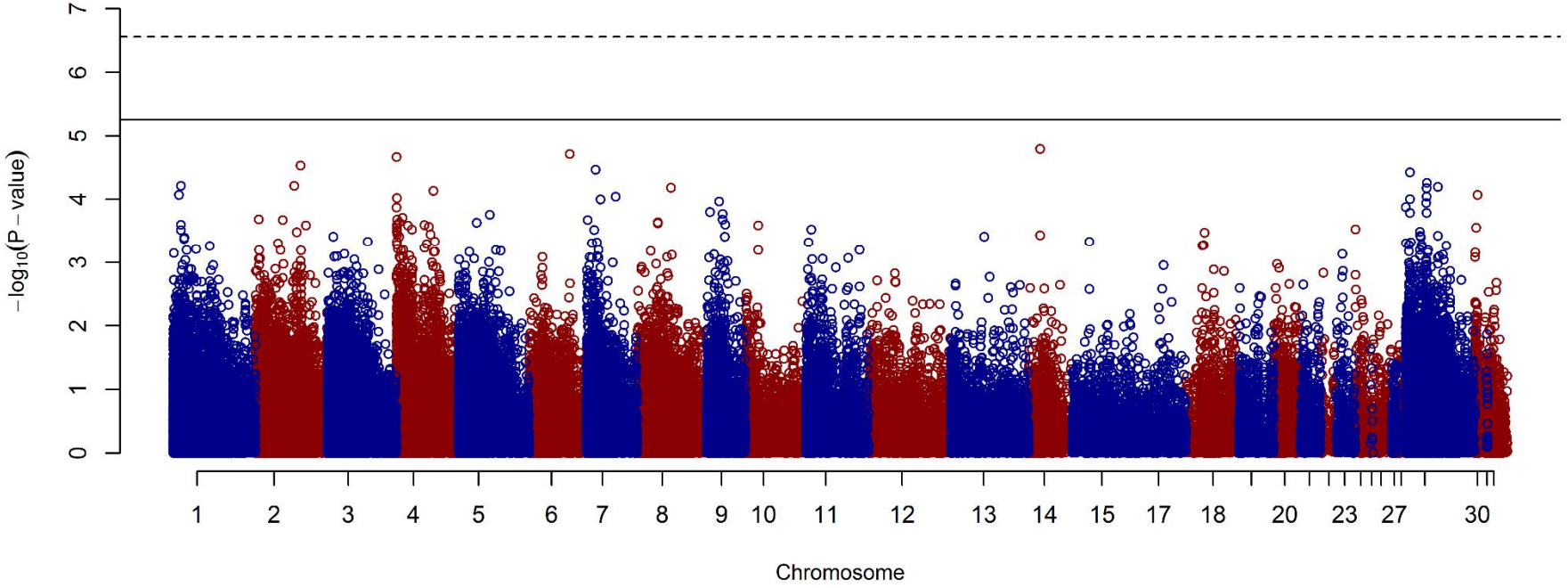

Fig. A4 Manhattan plot of repeated measures GWAS scan for SNPs associated with BMR (mL 0_2 h^{-1}) in 172 (N=225 observations) House Sparrows from Vega. 181 342 SNPs are shown with their chromosome labelled on the x-axis (SNPs positions are within scaffolds, see main text) and negative-log_{10} p-values on the y-axis. No SNPs were included from chromosome 16, and chromosomes; 19, 21, 22, 24, 25, 26, 28, 29, 31, and 32 (Z) are not labelled. The dashed line is the genome-wide significance threshold within Vega ($p=2.76 \times 10^{-7}$), and the solid line indicates the suggestive threshold within Vega ($p=5.52 \times 10^{-6}$).

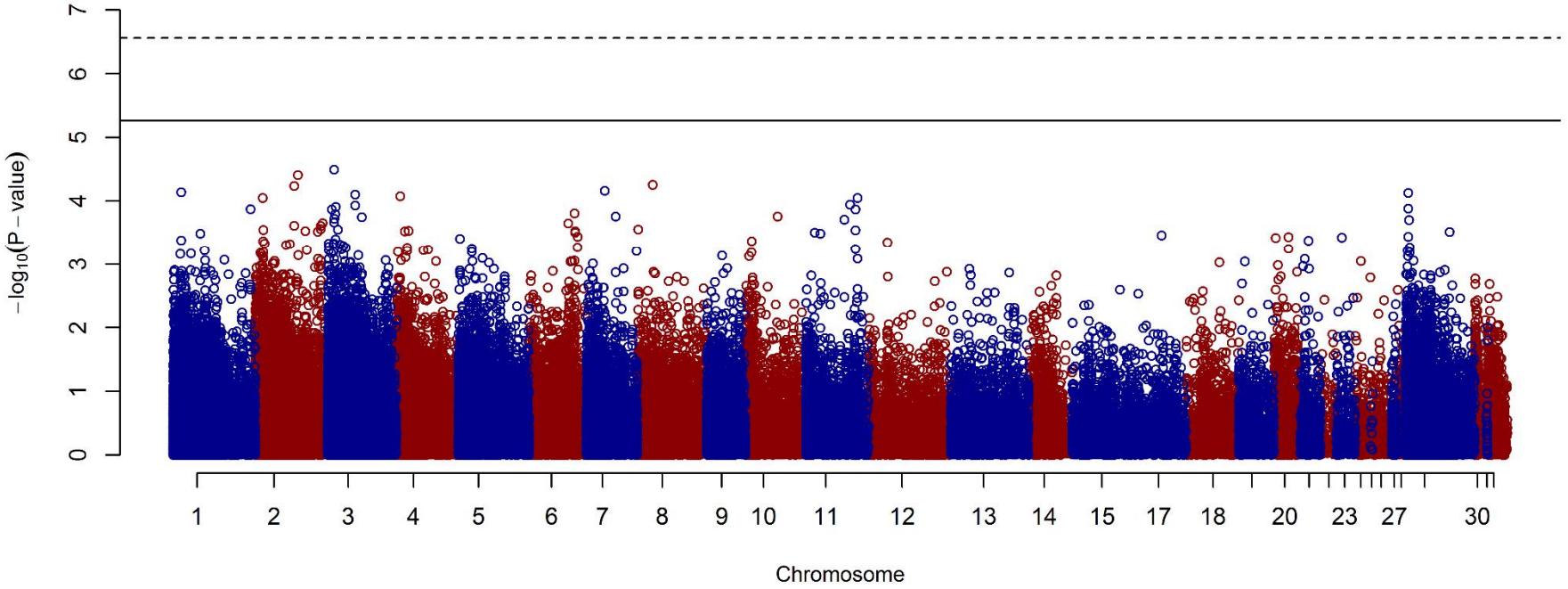

Fig. A5 Manhattan plot of repeated measures GWAS scan for SNPs associated with BMR (mL O_2 h^{-1}) in 258 (N=333 observations) House Sparrows from Lauvøya. 181 844 SNPs are shown with their chromosome labelled on the x-axis (SNPs positions are within scaffolds, see main text) and negative-$\log_{10}$ p-values on the y-axis. No SNPs were included from chromosome 16, chromosomes; 19, 21, 22, 24, 25, 26, 28, 29, 31, and 32 (Z) are not labelled. The dashed line is the genome-wide significance threshold within Lauvøya ($p=2.75 \times 10^{-7}$), and the solid line indicates the suggestive threshold within Lauvøya ($p=5.50 \times 10^{-6}$).