Polygenic basis for adaptive morphological variation in a threatened Aotearoa | New Zealand bird, the hihi (Notiomystis cincta)

Duntsch, Laura; Tomotani, B.M.; de Villemereuil, Pierre; Brekke, Patricia; Lee, Kate D.; Ewen, John G.; Santure, Anna W.

published in
Proceedings of the Royal Society B-Biological Sciences
2020

DOI (link to publisher)
10.1098/rspb.2020.0948

document version
Peer reviewed version

Polygenic basis for adaptive morphological variation in a threatened Aotearoa | New Zealand bird, the hihi (*Notiomystis cincta*)

Laura Duntsch[1] - Barbara M. Tomotani[2] - Pierre de Villemereuil[3] - Patricia Brekke[4] - Kate D. Lee[5] - John G. Ewen[4] - Anna W. Santure[1]

[1] School of Biological Sciences, University of Auckland, Auckland, NZ
[2] Netherlands Institute of Ecology, Wageningen, The Netherlands
[3] Institut de Systématique, Évolution, Biodiversité (ISYEB), École Pratique des Hautes Études PSL, MNHN, CNRS, Sorbonne Université, Université des Antilles, Paris, France
[4] Institute of Zoology, Zoological Society of London, Regents Park, London, UK
[5] School of Fundamental Sciences, Massey University, Palmerston North, NZ

Abstract

To predict if a threatened species can adapt to changing selective pressures, it is crucial to understand the genetic basis of adaptive traits, especially in species historically affected by severe bottlenecks. We estimated the heritability of three hihi (*Notiomystis cincta*) morphological traits known to be under selection: nestling tarsus length, body mass and head-bill length, using 523 individuals and 39,699 single nucleotide polymorphisms (SNPs) from a 50K Affymetrix SNP chip. We then examined the genetic architecture of the traits via chromosome partitioning analyses and genome-wide association scans (GWAS). Heritabilities estimated using pedigree relatedness or genomic relatedness were low. For tarsus length, the proportion of genetic variance explained by each chromosome was positively correlated with its size, and more than one chromosome explained significant variation for body mass and head-bill length. Finally, GWAS analyses suggested many loci of small effect contributing to trait variation for all three traits, although one locus (a SNP within an intron of the transcription factor *HEY2*) was tentatively associated with tarsus length. Our findings suggest a polygenic nature for the morphological traits, with many small effect size loci contributing to the majority of the variation, similar to results from many other wild populations. However, the small effective population size, polygenic architecture and already low heritabilities suggest that both the total response and rate of response to selection are likely to be limited in hihi.

Keywords: *genomic relatedness matrix; chromosome partitioning; genome-wide association; polygenic traits; pedigree study; Notiomystis cincta*

1 Introduction

In small populations, genetic diversity is eroded over time through genetic drift, including events such as population bottlenecks and founding events (Bouzat, 2010). The continual loss of genetic diversity is of particular interest in species of conservation concern with already low levels of variation. The erosion of genetic diversity may result in the loss of putatively adaptive alleles, leaving the species with low adaptive potential when facing new selection pressures, such as a change in climate or habitat fragmentation (Hoffmann et al., 2017, Chen et al., 2016, de Villemereuil et al., 2019b). However, the complex genomic architecture of traits in non-model organisms is often unknown and, consequently, the underlying evolutionary processes and possible ecological consequences for a population or a species subject to selection cannot be directly assessed (Savolainen et al., 2013). Recently, with the advent of modern and inexpensive sequencing and genotyping technologies, the field of conservation genomics has entered a new era where it is possible to genotype a large number of markers for threatened species with small population sizes (Narum et al., 2013). Now, the challenge at hand is to adequately analyse and interpret these genomic snapshots in a way that we can predict the adaptive potential for threatened species and begin to inform conservation management (Harrisson et al., 2014; Shafer et al., 2015).

Genomic approaches have already proven to be a powerful tool to better understand the structure and evolution of adaptive traits (Pardo-Diaz et al., 2015; Luikart et al., 2018). For example, a closer look at the genetic basis of horn polymorphism in Soay sheep (*Ovis aries*) revealed that variation is maintained by heterozygote advantage at a single gene (Johnston et al., 2013), which would be impossible to infer from pedigree-based methods alone. Hence, today, many population genomic studies aim to investigate the genetic basis of variation in life-history, fitness and morphological traits in their organism of interest (Luikart et al., 2018). To make assumptions about adaptive potential, it is crucial to understand the proportion of the phenotypic variation observed in a population (V_P) that is explained by the additive genetic component (V_A), a factor referred to as narrow-sense heritability (h^2) (Falconer and Mackay 1996). Traditionally, h^2 has been estimated using the expected relatedness between individuals captured by pedigree data. However, those pedigrees may contain errors, not capture all relationships between sampled individuals due to shallow and inconsistent pedigree depth, and only capture expected, rather than realised, genome sharing between two related individuals (Visscher et al., 2006). Therefore, it has recently become common practice to instead use genomic data when estimating trait heritabilities (Gienapp et al., 2017b), for example via inferring the realised genome sharing between individuals described by a genomic

relatedness matrix (GRM). Assuming that marker density is sufficient to accurately capture genomic relatedness, GRM-based heritabilities can be more precise, providing valuable insight into the evolutionary potential of the species (Berenos et al., 2014, Gienapp et al., 2017b, Jensen et al., 2014).

In addition to characterising the overall GRM-based heritability, the heritability can be partitioned across regions of the genome to give further insight into the genomic basis of trait variation (Yang et al., 2011b). If large chromosomes, containing more genes, explain a more substantial proportion of the genetic variation compared to smaller chromosomes, the trait is likely polygenic – i.e. there are many loci, scattered throughout the genome, contributing to trait differences (Yang et al., 2011b, Robinson et al., 2013). More than one chromosome explaining significant variation also suggests that many loci contribute to trait variation (Kemppainen & Husby, 2018a).

Another way of investigating the genetic basis of adaptive traits using genomic data involves a scan for genome-wide associations (GWAS). This detects specific loci with small to large effect on the phenotype of the study species. The power to detect association is dependent, among other factors, on the effect size of the locus. If a trait with significant heritability has no significant GWAS peaks, it suggests that the trait is polygenic, with many loci of small effect contributing to variation (Manilo et al., 2009). A further hurdle to finding significant genome-wide associations is presented when individuals are closely related, which is commonly the case when sampling from natural populations, particularly those with small population size. The similarity in genetic background must first be adjusted for before testing for an association, which may further reduce the power to detect an association (Yang et al., 2014).

It appears that a polygenic architecture underlies many evolutionarily significant traits under investigation (Jensen et al., 2014; Kardos et al., 2016). Variation in these traits is maintained by the balance between genetic drift, mutation and selection, with smaller heritabilities predicted in populations with small effective population size (Hoffman et al., 2017). For these populations, theoretical and simulation results suggest that polygenic adaptation in response to selection is likely to proceed via allele frequency changes at a subset of loci with large effect and/or high allele frequency, rather than frequency shifts across all loci (Höllinger et al., 2019). Variation at other loci is susceptible to be lost via genetic drift and will, therefore, constrain both the rate of response and the total response to selection (John & Stephan, 2020). In addition, the evolution of polygenic traits can be further constrained by linkage disequilibrium between deleterious and beneficial alleles, and trade-offs with other (potentially unmeasured) traits (Hill & Robertson 1966; Hospital & Chevalet 1996; Lande 1982). Therefore,

understanding the genetic architecture of adaptive traits is likely to enable better predictions of the adaptive potential of small populations.

The hihi is an endemic passerine of Aotearoa | New Zealand and is becoming a model species for studying the genetics of endangered birds, and species with small population sizes more generally (Brekke et al., 2011; de Villemereuil et al., 2019a; de Villemereuil et al., 2019b). The hihi dataset on the island of Tiritiri Matangi (36°36'S, 174°53'E) contains a multi-generational pedigree and morphological and life history data for every individual born since 1995 when the population was established through reintroduction. Recently, de Villemereuil et al. (2019a, 2019b) predicted little adaptive potential of this species based on signatures of low genetic diversity, low additive genetic variance of fitness, and small pedigree-based heritabilities for traits under selection. All but one life-history trait had very small heritabilities (<0.0001), however, there was good support for low but non-zero heritabilities for beak length, body mass and tarsus length (de Villemereuil et al., 2019b). While previous hihi studies already examined the age-dependent effect of body size on fitness (Low 2006), this large dataset (>3000 birds) for the first time revealed that all three morphological traits had positive linear selection and negative non-linear selection gradients (de Villemereuil et al., 2019b). These gradients suggest that there are optimal values for nestling tarsus length, head-bill length and mass that maximise fitness, but that all trait means are significantly smaller than these optima, and hence are under directional selection to increase body size. Morphological traits are commonly recorded in avian studies, with body size directly associated with fitness in many species (reviewed in Merilä and Sheldon, 2001; Gardner et al., 2011), and are the focus of a growing number of genomic studies (reviewed in Husby et al., 2019).

Here, we used genomic data to reveal previously undiscovered genetic patterns underlying these three adaptive morphological phenotypes in the Tiritiri Matangi population. We employed state-of-the-art software and data for 523 individuals genotyped on a recently generated custom SNP array to generate heritability estimates from pedigree relatedness and genomic relatedness, testing whether a genomic snapshot provides comparable results to a pedigree over many generations for the additive genetic variance component of the traits. To explore the genomic architecture of each trait, we further partitioned the variance explained by the genotyped SNPs across chromosomes to test for a possible polygenic architecture. Finally, we performed a genome-wide association study in order to find loci of small, medium and large effect on any of the traits under investigation. For traits with low heritability, we wanted to discern whether a GWAS could detect loci of large effect, or whether there was evidence that

traits are controlled by many loci of small effect, in agreement with studies on many other wild populations. From a conservation management perspective, a low heritability estimate and polygenic makeup of these morphological traits will complement our previous work demonstrating that hihi has little adaptive potential when facing future changes in selection pressures.

2 Methods

Study population and phenotypic data

The hihi was once abundant on the North Island of Aotearoa, until habitat loss and predation by introduced mammals confined its range to only one naturally occurring population on Te Hauturu-o-Toi (Little Barrier Island), in the Hauraki Gulf (36°12'S, 175°05'E) (Brekke et al., 2011). In 1995, the hihi was successfully introduced to the nearby island of Tiritiri Matangi, with a population size of 210 birds prior to the 2019/2020 breeding season (Parlato et al., *in review*). As a result of extensive monitoring since its translocation, we now have a wealth of environmental, life history, fitness, morphological and pedigree data for the population (de Villemereuil et al., 2019b). Hihi do not migrate between the populations and predominantly nest in nest boxes on Tiritiri Matangi Island. During every breeding season (ranging from October through to February), nestlings are banded and measured for body mass (g), but also for tarsus length (mm) and head-to-bill length (mm) as those two measures stabilize after 21 days. Blood has been routinely taken (around 60µl per individual via brachial vein puncture) from nestlings since 2004.

Genotyping and pedigree data

Due to a very high rate of extra-pair copulations in this species, the paternity is determined via routine genotyping of a panel of 19 microsatellite markers, allowing the construction of a multi-generational pedigree for the population (for details see Brekke et al., 2012; de Villemereuil et al., 2019a). For nestlings without blood samples available (i.e. those born before 2004), fathers are assumed to be unknown in the pedigree. In addition, in 2016, 1,536 birds (from five different hihi populations) were genotyped using a custom 50k Affymetrix single nucleotide polymorphism (SNP) array (manuscript *in prep*). This array was developed based on the identification of SNPs from *de-novo* assembly of restriction site-associated DNA sequencing (RAD-seq) of 31 individuals from the Te Hauturu-o-Toi and Tiritiri Matangi populations, and low-coverage whole-genome sequencing of ten of these birds (three of which were from Tiritiri Matangi). The flanking sequences of identified variants were aligned to the Ensemble 86 zebra finch (*Taeniopygia guttata*) genome, which, given the very high synteny of avian genomes

(Zhang et al., 2014), is used in the following analyses as a proxy for the distribution of markers in the hihi genome (electronic supplementary material, table S1).

Of the 58,466 SNPs included on the array, 45,553 markers were designated as polymorphic high resolution according to the default quality control metrics in the Axiom Analysis Suite software. *PLINK* (Purcell et al., 2007) filtering (--maf 0.01, --hwe 1e-20) resulted in the removal of 325 and 69 loci among the autosomes, respectively. 2,734 SNPs appeared as duplicates (i.e. mapped to the same chromosome at the same position) and hence only one SNP from each position was retained. To avoid population structure as a confounding factor, we only included birds from Tiritiri Matangi Island in this study, which shows no structure (electronic supplementary material, figure S1). No individual was removed due to low genotyping rates (--mind), but some were excluded for having incomplete phenotypic information. The final dataset contained 39,699 autosomal genomic loci with a mean call rate of 0.997398 and complete sex and phenotype information on all three traits for 523 birds.

Trait heritabilities

We employed several different methods of retrieving heritability estimates for our three traits under investigation, using generalised linear mixed models (an 'animal model', Wilson et al., 2009) and either the pedigree relatedness or the relatedness estimated from the genotyped markers (i.e. the genomic relatedness matrix; GRM). First, pedigree- and GRM-based heritabilities (h^2_{ped}, h^2_{GRM1}) were estimated in a Bayesian framework using the R (R Core Team, 2013c) package *MCMCglmm* (Hadfield 2010). To estimate h^2_{ped}, we fit the relatedness matrix calculated from the full pedigree. To estimate h^2_{GRM1}, the GRM was constructed using an approach that scales by the actual variance in relatedness (approach 3; Robinson et al., 2013). Following de Villemereuil (et al., 2019b), the additive genetic variance (V_A) for each trait was estimated by fitting sex and clutch size as fixed effects and accounting for the random effects of mother, social father, year and month when the nestling was born. Models were run for 503,000 iterations with a burn-in period of 3,000, sampling every 10th output from the chain, and convergence checked graphically and using the Heidelberger and Welch convergence test. The heritability was then calculated as the ratio of V_A to V_P (total phenotypic variance), where V_P and thus h^2 accounts for the variance explained by fixed effects (de Villemereuil et al., 2018).

Furthermore, to be able to compare our results with the heritability of morphological traits reported in other avian studies (reviewed in Husby et al., 2019), we also used the "prefitModel" function in the R package *RepeatABEL* (Rönnegård et al., 2016). While *RepeatABEL* is mainly

known for allowing the inclusion of repeated measures, it was chosen as it allows the inclusion of multiple random effects into the model. The GRM-based heritability (h^2_{GRM2}) was estimated by fitting the same fixed and random effects as above.

Finally, the package *BayesR* (Moser et al., 2015) was used to infer the phenotypic variation explained by SNPs in different effect size distributions and to estimate the SNP-based heritability (h^2_{BayesR}). The SNP-based heritability differs from the GRM-based heritability as it estimates the proportion of variance accounted for by linear regression on a set of genotyped SNPs (de los Campos et al., 2015). As *BayesR* does not allow for the inclusion of any fixed or random effects, the raw phenotypes were corrected for those effects in *MCMCglmm* (same model as for the heritability analysis, but excluding V_A) and the residuals then used as phenotypes in *BayesR*. We set the mean effect sizes of 0.01, 0.001, 0.0001 or 0 for the four mixture distributions and ran the programme for 50,000 iterations, with a burn-in of 20,000, recording every 10th sample in the chain after that (as suggested in the default settings).

Chromosome partitioning of genetic variance

For each chromosome, we estimated the contribution to overall trait variance by fitting a mixed model in *MCMCglmm* with a GRM estimated from markers on the focal chromosome, a GRM calculated from all other markers in the genome, and accounting for sex, clutch size, mother, social father, year and month of birth. As above, the GRMs were calculated using an approach that scales by the actual variance in relatedness. We ran 203,000 iterations with a burn-in period of 3,000, sampling every 10th output from the chain. Due to convergence issues for chromosomes with fewer than 850 markers, we merged multiple chromosomes based on their size according to the zebra finch karyotypes (electronic supplementary material, table S1). The variance was estimated individually for chromosomes 1-14, 1A and 20 and each merged chromosome group.

We also used *GCTA* (V1.24, Yang et al., 2011a) to estimate the variance explained by each chromosome. GRMs were generated with the --make-grm option for each individual chromosome and all but the focal chromosome. Because the model failed to converge when fitting all autosomes simultaneously, we fit the GRM pairs (e.g. GRM for Chr 1 and GRM for all but Chr 1, using the --mgrm option) in the default REML model one by one to estimate the additive genetic variance explained by each chromosome. Both pathways, however, should theoretically result in the same heritability estimate per autosomal chromosome (Yang et al., 2011a). Fixed effects included sex and clutch size, and we adjusted for population stratification using the –qcovar function on the top ten principle components.

For both methods, we used the cor.test function in R to test the significance of a positive Pearson's correlation between chromosome size (based on the zebra finch genome) and variance explained. To account for p-value inflation as a result of heteroscedasticity and censoring, we used 100,000 permutations to calculate corrected p-values following the approach of Kemppainen and Husby (2018b).

GWAS

To test for association between trait and genotype, we used the *RepeatABEL* package. The "prefitModel" function first fits a model without the SNP-effects and computes the variance components for the trait, using the *hglm* package. SNPs are then tested for association by accounting for these random effects in the "rGLS" function in *RepeatABEL*, using the same fixed and random effects as in the heritability analysis. P-values for each SNP are computed using Wald statistics, together with the SNP effect for each marker. We used a Bonferroni correction to control for family-wise error rate from multiple testing (significance value of 0.05 divided by the number of markers, giving a critical p-value of 1.26×10^{-6}). All genomic SNP were tested for association, including those that mapped to random or the Unknown zebra finch chromosomes, or were not mapped to a zebra finch chromosome (which we combine as 'chromosome 77', electronic supplementary material, table S1). SNPs of interest were further investigated based on the zebra finch genome annotation in the *Ensembl* database (Version 100.12: bTaeGut1_v1.p; Cunningham et al., 2019). In addition, we calculated the genomic inflation factor (λ, computed using lower 90 per cent of the distribution), to determine whether population structure was likely to have led to inflation of the test statistics.

3 Results

i) Trait heritability

The pedigree-based estimates for the heritability of nestling morphological traits tended to be higher than the GRM-based heritabilities, and *MCMCglmm* GRM-based estimates (h^2_{GRM1}) were lower than corresponding *RepeatABEL* heritabilities (table 1; electronic supplementary material, table S2, table S3, figure S2). The SNP-based heritabilities from *BayesR* were 0.117, 0.079 and 0.076, respectively (electronic supplementary material, table S5). All methods agreed on tarsus length being the 'most heritable' of the three morphological traits.

Table 1: Summary table of the outputs (electronic supplementary material, table S2, table S3) for the trait mean, total phenotypic variance (V_P), pedigree-based and GRM-based heritability (h^2) estimates and additive genetic variance (V_A) for tarsus length (mm), body mass (g) and head-bill length (mm). We report the mean h^2 and V_A for

the *MCMCglmm* (ped, GRM1) and *RepeatABEL* (GRM2) estimates for 523 hihi (268 males and 255 females). Credible / confidence intervals (*MCMCglmm*) or standard error (*RepeatABEL*) are reported in square brackets.

Trait	Mean	V_P	h^2_{ped}	$V_{A,\ ped}$	h^2_{GRM1}	$V_{A,\ GRM1}$	h^2_{GRM2}	$V_{A,\ GRM2}$
Tarsus length	27.576	1.440	0.133 [0.028, 0.253]	0.190 [0.040, 0.362]	0.077 [1.02e^{-9}, 0.154]	0.111 [1.39e^{-9}, 0.219]	0.081 [-0.020, 0.182]	0.082 [0.054, 0.126]
Body mass	37.011	43.662	0.066 [1.02e^{-10}, 0.151]	2.880 [4.18e^{-9}, 6.646]	0.042 [2.44e^{-10}, 0.105]	1.812 [1.14e^{-8}, 4.548]	0.065 [-0.031, 0.161]	2.050 [1.283, 3.276]
Head-bill length	38.793	2.845	0.120 [5.89e^{-8}, 0.252]	0.343 [1.45e^{-7}, 0.727]	0.032 [5.15e^{-16}, 0.094]	0.091 [1.42e^{-15}, 0.270]	0.070 [-0.028, 0.168]	0.159 [0.101, 0.251]

The correlation between off-diagonal elements of the pedigree-based and the GRM-based relatedness matrices was 0.749 (electronic supplementary material, figure S3). A correlation of less than one is expected due to factors such as Mendelian sampling and agrees with other studies (e.g. Pryce et al., 2012; Bérénos et al., 2014; Galla et al., 2020).

ii) Chromosome partitioning of genetic variance

There was evidence that multiple loci contributed to the variance in all three morphological traits (figure 1; electronic supplementary material, figure S4). For tarsus length, we found a positive but non-significant correlation between the proportion of additive genetic variance explained by each chromosome and the chromosome size (HC corrected p-value = 0.113). Chromosome 1, 1A and 6 explain the most variation, although the largest chromosome (2) explained only moderate variation.

For body mass, we found a weak and non-significant positive correlation between chromosome size and phenotypic variance explained (HC corrected p-value = 0.422). Here, chromosomes 14 and 9 explain the most variation.

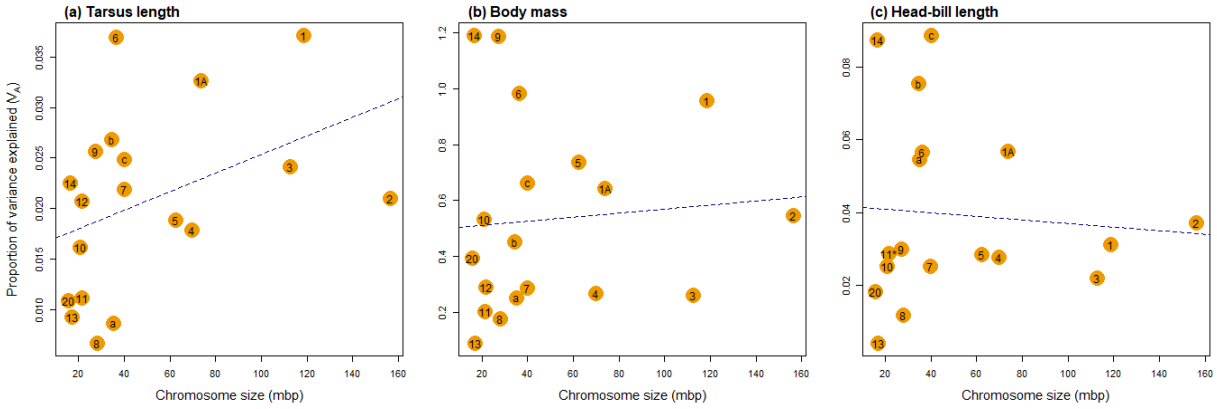

Figure 1: The variation explained by each chromosome or chromosome composite (in yellow circles, excluding 'Chr 77') for all three traits (tarsus length, body mass and head-bill length) as calculated in *MCMCglmm*. Plotted is additive genetic variance explained by each chromosome (V_A) against the length of the chromosome (in Mbp) based on the zebra finch reference genome. After HC correction, all p-values were non-significant (0.113, 0.422 and 0.574). a = chromosomes 15 and 4A, b = chromosomes 17-19, c = chromosomes 21-28 and 1B (electronic supplementary material, table S1).

Finally, for head-bill length, we found a weak negative correlation between the variation and the size of the genomic region. The three largest chromosomes explain relatively little of the overall variation, while chromosome 14, and micro-chromosome composites 'b' and 'c' each explain the most variation. Chromosome partitioning in GCTA resulted in a comparable picture for all three traits (electronic supplementary material, figure S5).

iii) GWAS

No SNP was significantly associated with the morphological traits after testing for association in *RepeatABEL* (figure 2). According to the well-curated zebra finch genome (which the hihi chromosome assignment is derived from), the SNP with the highest support for association with tarsus length is located in one of the intronic regions of the *HEY2* transcription factor (electronic supplementary material, figure S6), which is involved in cardiac morphogenesis, neurogenesis and somitogenesis (Ihara et al., 2020).

For head-bill length, there are two SNPs that, while not significant, show a stronger association with the trait than all others. One of them is placed in an intron of an uncharacterized protein-coding gene (ENSTGUG00000006844) on chromosome 1, the other in between two genes associated with cat eye syndrome (CECR1, CECR2) on chromosome 1A. The third most associated SNP for head-bill length was found on chromosome 20, a chromosome previously associated with bill length (Knief et al., 2012; Silva et al., 2017).

The top ten SNPs with the lowest p-values across all traits can be found in electronic supplementary material, tables S4. We refrained from calculating the additive genetic variance explained by each of the most associated SNPs, as p-values were generally high. The genomic inflation factor values ranged from 0.96 to 1.05 for all three phenotypic traits under investigation (electronic supplementary material, figure S7), suggesting that population structure has been adequately accounted for.

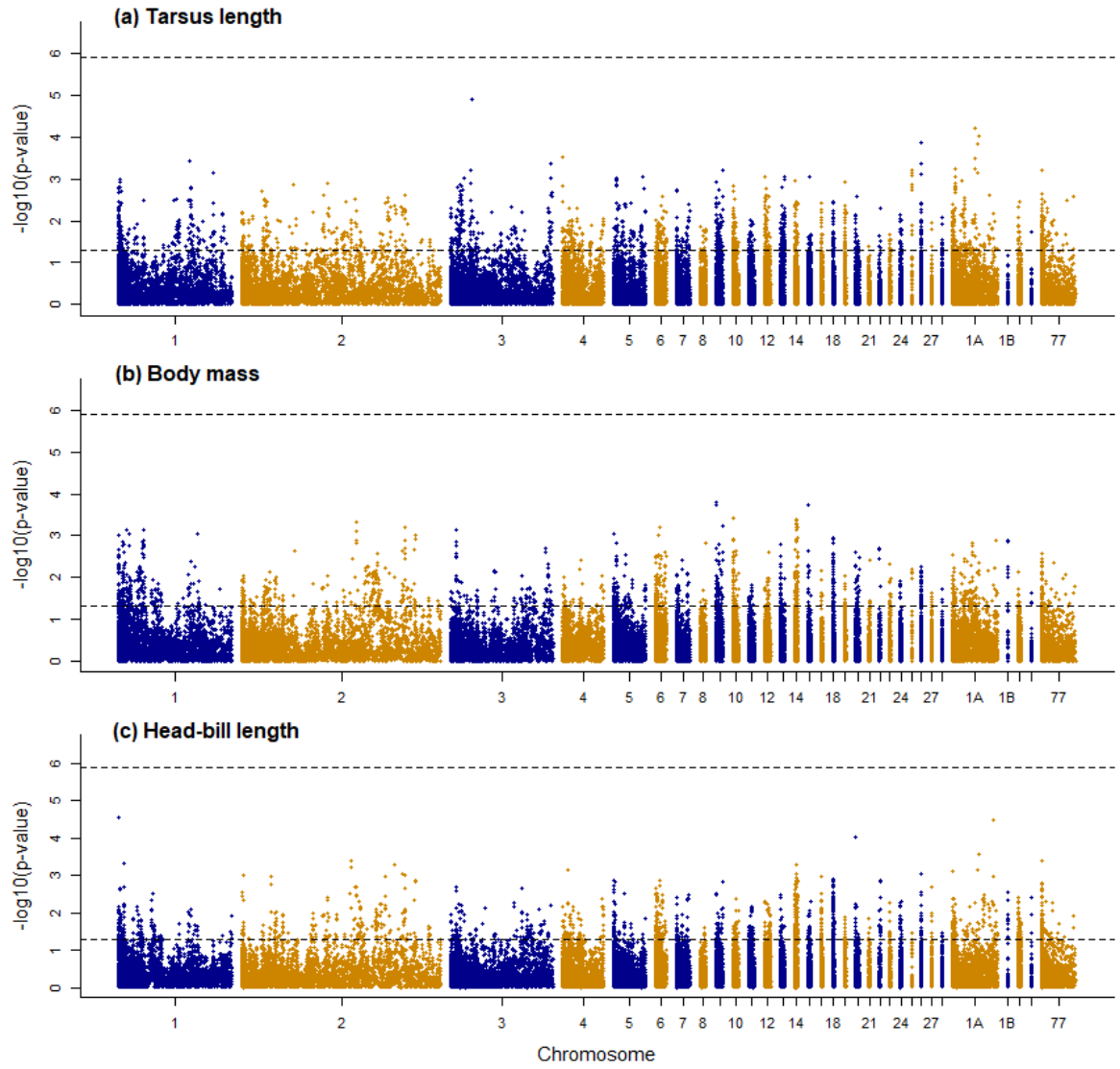

Figure 2: Manhattan plots for the association between single nucleotide polymorphisms and (a) tarsus length, (b) body mass and (c) head-bill length in 523 hihi nestlings of Tiritiri Matangi. Top black dashed line: Bonferroni significance threshold, bottom dashed line: nominal significance (p-value = 0.05). A description of the chromosome order and naming can be found in the electronic supplementary material, table S1.

The *BayesR* analyses confirmed that all traits under investigation appear polygenic, with less than 9% of the SNPs affecting trait variation, the majority of which contributed to the smallest effect size distribution (electronic supplementary material, table S5). A comparison of the SNPs with the lowest p-values from the GWAS showed reasonable agreement with those inferred to have a high probability of inclusion in the largest effect size distribution from *BayesR* (electronic supplementary material, tables S4).

4 Discussion

Adaptive response requires that there is genetic variation in traits linked to fitness, and further, that response to selection is unconstrained by trade-offs with other traits and loci. Here, we examined the genetic basis of variation in nestling tarsus length, body mass and head-bill length in the hihi, a threatened yet well-monitored bird species of Aotearoa | New Zealand, in order to understand whether these traits can respond to current and future selection pressures. Our analyses confirmed low heritability estimates, and support a polygenic basis for all three traits, with no loci of large effect contributing to trait variation. Given the low effective population size and the polygenic architecture, such that adaptive variation lost via drift is unlikely to be replenished by mutation in the short-term, we conclude that evolutionary potential in hihi is likely constrained even further than the already low heritabilities suggest.

We have three reasons to believe that a polygenic architecture underlies all three morphological traits under investigation: (i) for tarsus length, the variance explained by each chromosome scaled with chromosome size, and for all traits more than one chromosome contributed to variation, (ii) there were no significant SNP associations in the GWAS and (iii) the majority of loci explaining variation are of small effect size. This study is in good agreement with studies in other passerines, (e.g. great tits (Santure et al., 2013; 2015; Gienapp et al., 2017a; Bosse et al. 2017; Kim et al., 2018), collared flycatchers (Husby et al., 2015; Kardos et al., 2016; Silva et al., 2017), house sparrows (Silva et al., 2017; Lundregan et al., 2018), great reed warblers (Hansson et al. 2018), reviewed in Husby (et al., 2019)) as well as studies from other wild animal populations (e.g. Bérénos et al., 2015), where many loci with small effect are responsible for variation in adaptive phenotypes, rather than a few large-effect loci.

Our pedigree- and GRM-based heritability estimates support a small but non-zero heritability for all traits and agree with previous pedigree heritabilities estimated from a much larger number (>2,000) of individuals (de Villemereuil et al., 2019b). Interestingly, the GRM-based heritabilities were lower than the pedigree-based estimates, although the credible intervals are overlapping in all cases. One reason for this may be that our SNP array is not dense enough

to have captured true genome sharing more accurately than the pedigree has done. Given the number of SNPs genotyped, this seems unlikely (Berenos et al., 2014), although we note that we have not successfully genotyped SNPs from a number of microchromosomes, including chromosome 16. We acknowledge that, while 523 individuals is a large dataset for a threatened species, our sample size is likely to be somewhat limiting and has certainly contributed to the large credible intervals and standard errors around both the pedigree- and GRM-based heritability estimates, which may also explain some of the differences between estimates. Interestingly, a study of Corsican blue tits (*Cyanistes caeruleus*; Perrier et al., 2018) suggests that a sample size of around 500 birds and data from a 50K SNP chip can outperform heritability estimates from a seven-generation pedigree. The ability to estimate heritabilities without long-term and near-complete monitoring of a population offers exciting opportunity to infer adaptive potential in other wild populations and will be particularly valuable in assessing the vulnerability of endangered species (Gienapp et al., 2017b). The sample size, number of markers and low heritabilities is also likely to have limited power to detect association in the GWAS analyses. Linkage disequilibrium between neighbouring markers is moderate (mean of 0.518, calculated with all 523 individuals using plink --r2 and reporting all pairwise values with --ld-window-r2 0.0) and suggests that power to detect a locus explaining *all* of the trait heritability ranges from 0.306 to 0.516 (Wang and Xu 2019; electronic supplementary material, table S6). As expected, nestling mass, with the lowest heritability due to substantial maternal and month effects compared to the other traits (de Villemereuil et al. 2019b), has the lowest power to detect association.

Morphological traits are often the subject of wild bird studies, as they can reveal evolutionary patterns over time and in turn shed light on the adaptive response of traits known to correlate with fitness (Merilä and Sheldon, 2001; Gardner et al., 2011). The three morphological traits studied here are all under selection as previous analyses suggest that the population mean of all three traits is substantially smaller than the fitness optimum (de Villemereuil et al. 2019b) and that there is a trade-off between body size and agility mediated during the highly energetic breeding season of the hihi (Low 2006). Although there was no significant association detected in the GWAS, it is interesting to note the regions that were most highly associated with tarsus length and head-bill length. The tentative association of a locus on chromosome 3 with tarsus length and of variants on chromosomes 1 and 1A with head-bill length does not appear to be replicated in some studies of passerine body size (Silva et al., 2017; Bosse et al., 2017), while Lawson and Petren (2017) highlight a region on chromosome 1A in increased linkage with regard to beak shape in Darwin finches (*Geospiza*). Additionally, the third-highest association for head-bill, a SNP on chromosome 20, stands within an intron of NFS1 cysteine desulfurase

and is 7-12Mb distant from QTLs previously associated with beak morphology (Knief et al., 2012; Silva et al., 2017). The challenges in identifying significant loci in a genome-wide association analysis include low SNP density and small sample sizes (Santure & Garant 2018), both of which are likely to apply to the hihi study system to some degree. All the above-mentioned loci are interesting candidates for further study if sample size and SNP density in these regions can be increased. Creating a larger sample size through combining all monitored populations could improve the precision of estimates, and shed more light on the effects of translocations on trait variation and responses to selection.

Despite the low heritabilities of the traits, our genomic analyses have provided important further insights for hihi, and threatened species more generally. A more nuanced understanding of the genomic architecture of a trait under selection helps assess the likely speed of the adaptive response in a species and the longer-term total response to selection (John & Stephan, 2020; Höllinger et al., 2019). We have demonstrated that many loci of small effect are likely to contribute the majority of variation to all three morphological traits. Given the small effective population size of all hihi populations, many adaptive variants are likely to be lost to drift at a rate that is faster than the input of new variation via mutation (reviewed in Johnson & Barton, 2005; Willi & Buskirk, 2006). This is likely to slow the speed of adaptation and lead to further constraint on the adaptive potential of the species. Most importantly, other wild populations may be facing similar threats as the hihi, highlighting the need for continued collection and use of genomic resources to aid conservation measures. We therefore encourage further characterisation of the heritability and genomic architecture for species that face anthropogenic threats including habitat decimation and climate change, in order to better predict their evolutionary potential to adapt to such challenges (Gienapp & Brommer, 2014). Finally, we acknowledge that future studies will need to go beyond describing quantitative parameters in order to successfully combine genomic insights with active species conservation, including informing management actions that can buffer species with small population sizes from the impacts of low adaptive potential.

Acknowledgements

We extend many thanks to the volunteers, past students, and Department of Conservation staff who have contributed to monitoring the Tiritiri Matangi hihi population and to the Hihi Recovery Group, Department of Conservation, and Supporters of Tiritiri Matangi for maintaining such a long-term vision in monitoring and management of this population, with a special thank you to the current hihi conservation officer Mhairi McCready. We acknowledge Ngati Manuhiri as Mana Whenua and Kaitiaki of Te Hauturu-o-Toi and its taonga, including

hihi. Many thanks to Rachel Tucker and Jon Slate at the NERC Biomolecular Analysis Facility, University of Sheffield, for completing DNA extractions for the RADseq data and to Johanna Nielsen and Kang-Wook Kim for help with microsatellite genotyping. We acknowledge the use of New Zealand eScience Infrastructure (NeSI) high-performance computing facilities and greatly appreciated help from Catarina Silva and Lars Rönnegård with RepeatABEL. Finally, we are very grateful to Associate Editor Nancy Chen, Arild Husby and an anonymous reviewer for their insightful and helpful comments on an earlier version of the manuscript.

Funding

P.d.V., A.W.S., K.D.L., P.B., and J.G.E. were supported by a Marsden Grant (UOA1408) awarded to A.W.S. from the New Zealand Royal Society Te Aparangi. A.W.S. was also supported by a New Zealand National Science Challenge Biological Heritage Project Grant, Project 1.4. P.B. and the pedigree genotyping were also supported by a ZSL Fellowship. L.D. is supported by a University of Auckland Doctoral Scholarship.

References

Bérénos C, Ellis PA, Pilkington JG, Pemberton JM. Estimating quantitative genetic parameters in wild populations: a comparison of pedigree and genomic approaches. Molecular Ecology. 2014 Jul;23(14):3434-51.

Bérénos C, Ellis PA, Pilkington JG, Lee SH, Gratten J, Pemberton JM. Heterogeneity of genetic architecture of body size traits in a free-living population. Molecular Ecology. 2015 Apr;24(8):1810-30.

Bouzat JL. Conservation genetics of population bottlenecks: the role of chance, selection, and history. Conservation Genetics. 2010 Apr 1;11(2):463-78.

Bosse M, Spurgin LG, Laine VN, Cole EF, Firth JA, Gienapp P, Gosler AG, McMahon K, Poissant J, Verhagen I, Groenen MA. Recent natural selection causes adaptive evolution of an avian polygenic trait. Science. 2017 Oct 20;358(6361):365-8.

Brekke P, Bennett PM, Santure AW, Ewen JG. High genetic diversity in the remnant island population of hihi and the genetic consequences of re-introduction. Molecular Ecology. 2011 Jan;20(1):29-45.

Brekke P, Wang J, Bennett PM, Cassey P, Dawson DA, Horsburgh GJ, Ewen JG. Postcopulatory mechanisms of inbreeding avoidance in the island endemic hihi (*Notiomystis cincta*). Behavioral Ecology. 2012 Mar 1;23(2):278-84.

Chen N, Cosgrove EJ, Bowman R, Fitzpatrick JW, Clark AG. Genomic consequences of population decline in the endangered Florida scrub-jay. Current Biology. 2016 Nov 7;26(21):2974-9.

Cunningham F, Achuthan P, Akanni W, Allen J, Amode MR, Armean IM, Bennett R, Bhai J, Billis K, Boddu S, Cummins C. Ensembl 2019. Nucleic Acids Research. 2019 Jan 8;47(D1):D745-51.

de los Campos G, Sorensen D, Gianola D. Genomic heritability: what is it? PLoS Genetics. 2015 May;11(5):e1005048.

de Villemereuil P, Morrissey MB, Nakagawa S, Schielzeth H. Fixed-effect variance and the estimation of repeatabilities and heritabilities: Issues and solutions. Journal of Evolutionary Biology. 2018 Apr;31(4):621-32.

de Villemereuil P, Rutschmann A, Ewen JG, Santure AW, Brekke P. Can threatened species adapt in a restored habitat? No expected evolutionary response in lay date for the New Zealand hihi. Evolutionary Applications. 2019a Mar;12(3):482-97.

de Villemereuil P, Rutschmann A, Lee KD, Ewen JG, Brekke P, Santure AW. Little adaptive potential in a threatened passerine bird. Current Biology. 2019b Mar 4;29(5):889-94.

Falconer DS, Mackay TF. Introduction to quantitative genetics. 1996. Harlow, Essex, UK: Longmans Green. 1996;3.

Galla SJ, Moraga R, Brown L, Cleland S, Hoeppner MP, Maloney RF, Richardson A, Slater L, Santure AW, Steeves TE. A comparison of pedigree, genetic, and genomic estimates of relatedness for informing pairing decisions in two critically endangered birds: Implications for conservation breeding programmes worldwide. Evolutionary Applications. 2019 Jan 1.

Gardner JL, Peters A, Kearney MR, Joseph L, Heinsohn R. Declining body size: a third universal response to warming?. Trends in ecology & evolution. 2011 Jun 1;26(6):285-91.

Gienapp P, Brommer JE. Evolutionary dynamics in response to climate change. In Quantitative Genetics in the Wild. 2014 Apr 3:254-73.

Gienapp P, Laine VN, Mateman AC, van Oers K, Visser ME. Environment-dependent genotype-phenotype associations in avian breeding time. Frontiers in genetics. 2017a Aug 4;8:102.

Gienapp P, Fior S, Guillaume F, Lasky JR, Sork VL, Csilléry K. Genomic quantitative genetics to study evolution in the wild. Trends in Ecology & Evolution. 2017b Dec 1;32(12):897-908.

Hadfield JD. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. Journal of Statistical Software. 2010 Feb 2;33(2):1-22.

Hansson B, Sigeman H, Stervander M, Tarka M, Ponnikas S, Strandh M, Westerdahl H, Hasselquist D. Contrasting results from GWAS and QTL mapping on wing length in great reed warblers. Molecular Ecology Resources. 2018 Jul;18(4):867-76.

Harrisson KA, Pavlova A, Telonis-Scott M, Sunnucks P. Using genomics to characterize evolutionary potential for conservation of wild populations. Evolutionary Applications. 2014 Nov;7(9):1008-25.

Hill WG, Robertson A. The effect of linkage on limits to artificial selection. Genetics Research. 1966 Dec;8(3):269-94.

Hoffmann AA, Sgrò CM, Kristensen TN. Revisiting adaptive potential, population size, and conservation. Trends in Ecology & Evolution. 2017 Jul 1;32(7):506-17.

Höllinger I, Pennings PS, Hermisson J. Polygenic adaptation: From sweeps to subtle frequency shifts. PLoS Genetics. 2019 Mar 20;15(3):e1008035.

Hospital F, Chevalet C. Interactions of selection, linkage and drift in the dynamics of polygenic characters. Genetics Research. 1996 Feb;67(1):77-87.

Husby A, Kawakami T, Rönnegård L, Smeds L, Ellegren H, Qvarnström A. Genome-wide association mapping in a wild avian population identifies a link between genetic and phenotypic variation in a life-history trait. Proceedings of the Royal Society B: Biological Sciences. 2015 May 7;282(1806):20150156.

Husby A, McFarlane SE, Qvarnström A. Avian Population Studies in the Genomic Era. In Avian Genomics in Ecology and Evolution. 2019 (pp. 267-293). Springer, Cham.

Ihara D, Watanabe Y, Seya D, Arai Y, Isomoto Y, Nakano A, Kubo A, Ogura T, Kawamura T, Nakagawa O. Expression of Hey2 transcription factor in the early embryonic ventricles is controlled through a distal enhancer by Tbx20 and Gata transcription factors. Developmental Biology. 2020 Feb 5.

Jensen H, Szulkin M, Slate J. Molecular quantitative genetics. In Quantitative Genetics in the Wild. 2014 Apr 3:209-27.

John S, Stephan W. Important role of genetic drift in rapid polygenic adaptation. Ecology and Evolution. 2020 Jan 10:1278–1287.

Johnson T, Barton N. Theoretical models of selection and mutation on quantitative traits. Philosophical Transactions of the Royal Society B: Biological Sciences. 2005 Jul 29;360(1459):1411-25.

Johnston SE, Gratten J, Berenos C, Pilkington JG, Clutton-Brock TH, Pemberton JM, Slate J. Life history trade-offs at a single locus maintain sexually selected genetic variation. Nature. 2013 Oct;502(7469):93-5.

538 Kardos M, Husby A, McFarlane SE, Qvarnström A, Ellegren H. Whole-genome resequencing
539 of extreme phenotypes in collared flycatchers highlights the difficulty of detecting quantitative
540 trait loci in natural populations. Molecular Ecology Resources. 2016 May;16(3):727-41.

541 Kemppainen P, Husby A. Inference of genetic architecture from chromosome partitioning
542 analyses is sensitive to genome variation, sample size, heritability and effect size distribution.
543 Molecular Ecology Resources. 2018a Jul;18(4):767-77.

544 Kemppainen P, Husby A. Accounting for heteroscedasticity and censoring in chromosome
545 partitioning analyses. Evolution Letters. 2018b Dec;2(6):599-609.

546 Kim JM, Santure AW, Barton HJ, Quinn JL, Cole EF, Great Tit HapMap Consortium, Visser
547 ME, Sheldon BC, Groenen MA, van Oers K, Slate J. A high-density SNP chip for genotyping
548 great tit (Parus major) populations and its application to studying the genetic architecture of
549 exploration behaviour. Molecular Ecology Resources. 2018 Jul;18(4):877-91.

550 Knief U, Schielzeth H, Kempenaers B, Ellegren H, Forstmeier W. QTL and quantitative
551 genetic analysis of beak morphology reveals patterns of standing genetic variation in an
552 Estrildid finch. Molecular Ecology. 2012 Aug;21(15):3704-17.

553 Lande R. A quantitative genetic theory of life history evolution. Ecology. 1982 Jun;63(3):607-
554 15.

555 Lawson LP, Petren K. The adaptive genomic landscape of beak morphology in Darwin's
556 finches. Molecular Ecology. 2017 Oct;26(19):4978-89.

557 Low M. Sex, age and season influence morphometries in the New Zealand Stitchbird (or
558 Hihi; *Notiomystis cincta*). 2006;106(4):297-304.

559 Luikart G, Kardos M, Hand BK, Rajora OP, Aitken SN, Hohenlohe PA. Population genomics:
560 advancing understanding of nature. In Population Genomics. 2018 (pp. 3-79). Springer,
561 Cham.

562 Lundregan SL, Hagen IJ, Gohli J, Niskanen AK, Kemppainen P, Ringsby TH, Kvalnes T,
563 Pärn H, Rønning B, Holand H, Ranke PS. Inferences of genetic architecture of bill
564 morphology in house sparrow using a high-density SNP array point to a polygenic basis.
565 Molecular Ecology. 2018 Sep;27(17):3498-514.

566 Merilä J, Sheldon BC. Avian quantitative genetics. In Current ornithology 2001 (pp. 179-255).
567 Springer, Boston, MA.

568 Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy MI, Ramos
569 EM, Cardon LR, Chakravarti A, Cho JH. Finding the missing heritability of complex diseases.
570 Nature. 2009 Oct;461(7265):747-53.

571 Moser G, Lee SH, Hayes BJ, Goddard ME, Wray NR, Visscher PM. Simultaneous discovery,
572 estimation and prediction analysis of complex traits using a Bayesian mixture model. PLoS
573 Genetics. 2015 Apr;11(4):e1004969.

574 Narum SR, Buerkle CA, Davey JW, Miller MR, Hohenlohe PA. Genotyping-by-sequencing in
575 ecological and conservation genomics. Molecular Ecology. 2013 Jun;22(11):2841-7.

576 Pardo-Diaz C, Salazar C, Jiggins CD. Towards the identification of the loci of adaptive
577 evolution. Methods in Ecology and Evolution. 2015 Apr;6(4):445-64.

578 Parlato EH, Ewen JG, McCready M, Parker KA. Armstrong DP. A modelling framework for
579 integrating reproduction, survival and count data when projecting the fates of threatened
580 populations. Oecologia. *In review*.

581 Perrier C, Delahaie B, Charmantier A. Heritability estimates from genomewide relatedness
582 matrices in wild populations: Application to a passerine, using a small sample size. Molecular
583 Ecology Resources. 2018 Jul;18(4):838-53.

584 Pryce JE, Hayes BJ, Goddard ME. Novel strategies to minimize progeny inbreeding while
585 maximizing genetic gain using genomic information. Journal of Dairy Science. 2012 Jan
586 1;95(1):377-88.

587 Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, De
588 Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and
589 population-based linkage analyses. The American Journal of Human Genetics. 2007 Sep
590 1;81(3):559-75.

591 R Core Team. R: A language and environment for statistical computing. R Foundation for
592 Statistical Computing,Vienna, Austria. 2013. https://www.R-project.org/.

593 Robinson MR, Santure AW, DeCauwer I, Sheldon BC, Slate J. Partitioning of genetic
594 variation across the genome using multimarker methods in a wild bird population. Molecular
595 Ecology. 2013 Aug;22(15):3963-80.

596 Rönnegård L, McFarlane SE, Husby A, Kawakami T, Ellegren H, Qvarnström A. Increasing
597 the power of genome wide association studies in natural populations using repeated
598 measures–evaluation and implementation. Methods in Ecology and Evolution. 2016
599 Jul;7(7):792-9.

600 Santure AW, De Cauwer I, Robinson MR, Poissant J, Sheldon BC, Slate J. Genomic
601 dissection of variation in clutch size and egg mass in a wild great tit (*Parus major*) population.
602 Molecular Ecology. 2013 Aug;22(15):3949-62.

603 Santure AW, Poissant J, De Cauwer I, Van Oers K, Robinson MR, Quinn JL, Groenen MA,
604 Visser ME, Sheldon BC, Slate J. Replicated analysis of the genetic architecture of

quantitative traits in two wild great tit populations. Molecular Ecology. 2015 Dec;24(24):6148-62.

Santure AW, Garant D. Wild GWAS—association mapping in natural populations. Molecular Ecology Resources. 2018 Jul;18(4):729-38.

Savolainen O, Lascoux M, Merilä J. Ecological genomics of local adaptation. Nature Reviews Genetics. 2013 Nov;14(11):807-20.

Shafer AB, Wolf JB, Alves PC, Bergström L, Bruford MW, Brännström I, Colling G, Dalén L, De Meester L, Ekblom R, Fawcett KD. Genomics and the challenging translation into conservation practice. Trends in Ecology & Evolution. 2015 Feb 1;30(2):78-87.

Silva CN, McFarlane SE, Hagen IJ, Rönnegård L, Billing AM, Kvalnes T, Kemppainen P, Rønning B, Ringsby TH, Sæther BE, Qvarnström A. Insights into the genetic architecture of morphological traits in two passerine bird species. Heredity. 2017 Sep;119(3):197-205.

Visscher PM, Medland SE, Ferreira MA, Morley KI, Zhu G, Cornes BK, Montgomery GW, Martin NG. Assumption-free estimation of heritability from genome-wide identity-by-descent sharing between full siblings. PLoS Genetics. 2006 Mar;2(3):e41.

Wang M, Xu S. Statistical power in genome-wide association studies and quantitative trait locus mapping. Heredity. 2019;123(3):287-306.

Willi Y, Van Buskirk J, Hoffmann AA. Limits to the adaptive potential of small populations. Annual Review of Ecology, Evolution, and Systematics. 2006 Dec 1;37:433-58.

Wilson AJ, Reale D, Clements MN, Morrissey MM, Postma E, Walling CA, Kruuk LE, Nussey DH. An ecologist's guide to the animal model. Journal of Animal Ecology. 2010 Jan;79(1):13-26.

Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. The American Journal of Human Genetics. 2011a Jan 7;88(1):76-82.

Yang J, Manolio TA, Pasquale LR, Boerwinkle E, Caporaso N, Cunningham JM, De Andrade M, Feenstra B, Feingold E, Hayes MG, Hill WG. Genome partitioning of genetic variation for complex traits using common SNPs. Nature Genetics. 2011b Jun;43(6):519.

Yang J, Zaitlen NA, Goddard ME, Visscher PM, Price AL. Advantages and pitfalls in the application of mixed-model association methods. Nature Genetics. 2014 Feb;46(2):100.

Yang J, Zeng J, Goddard ME, Wray NR, Visscher PM. Concepts, estimation and interpretation of SNP-based heritability. Nature Genetics. 2017 Sep;49(9):1304.

636 Zhang G, Li C, Li Q, Li B, Larkin DM, Lee C, Storz JF, Antunes A, Greenwold MJ, Meredith
637 RW, Ödeen A. Comparative genomics reveals insights into avian genome evolution and
638 adaptation. Science. 2014 Dec 12;346(6215):1311-20.