Supplementary Materials for

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

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

Correspondence to: ldun612@aucklanduni.ac.nz

Journal: *Proceedings of the Royal Society B: Biological Sciences*

DOI: 10.1098/RSPB-2020-0948

This PDF file includes:

Tables S1 to S6

Figures S1 to S7

Table S1*:* Overview of the SNP markers genotyped per chromosome in the hihi, the chromosome size based on the zebra finch reference genome and the original and adjusted/composite chromosome names for each analysis. Chromosome 77 represents SNPs that mapped to random or the 'Unknown' zebra finch chromosomes or were not mapped to a zebra finch chromosome.

Zebra finch chromosome number	Zebra finch chromosome length	Inferred hihi chromosome number	Hihi markers genotyped	*MCMCglmm* merge code
1	118,548,696	1	4,189	1
2	156,412,533	2	5,569	2
3	112,617,285	3	3,962	3
4	69,780,378	4	2,580	4
5	62,374,962	5	2,209	5
6	36,305,782	6	1,323	6
7	39,844,632	7	1,478	7
8	27,993,427	8	1,010	8
9	27,241,186	9	975	9
10	20,806,668	10	1,183	10
11	21,403,021	11	1,254	11
12	21,576,510	12	1,216	12
13	16,962,381	13	915	13
14	16,419,078	14	919	14
15	14,428,146	15	806	a
17	11,648,728	17	622	b
18	11,201,131	18	603	b
19	11,587,733	19	574	b
20	15,652,063	20	850	20
21	5,979,137	21	326	c
22	3,370,227	22	180	c
23	6,196,912	23	327	c
24	8,021,379	24	431	c
25	1,275,379	25	57	c
26	4,907,541	26	277	c
27	4,618,897	27	209	c
28	4,963,201	28	258	c
1A	7,3657,157	1A	2,683	1A
1B	1,083,483	1B	38	c
4A	20,704,505	4A	734	a
LGE422	883,365	LGE422	62	c
-	-	77	1,880	77

Table S2: Output from the pedigree-based *MCMCglmm* run of generalized mixed models for tarsus length (mm), body mass (g) and head-bill length (mm) including 523 hihi. f.effect = fixed effects fitted in the model, mean = trait means, vp = total phenotypic variance, cvp = total coefficient of variation, va = additive genetic variance, cva = additive genetic coefficient of variation, h2 = heritability, va */ (mean^2)* = evolvability.

Trait	f.effect	mean	mean.mode	mean.low	mean.up	vp	vp.mode
Tarsus	Sex + ClutchSize	27.57486	27.57593	27.23846	27.9039	1.440098	1.398983
Mass	Sex + ClutchSize	37.01056	37.02464	35.47931	38.49293	43.66249	43.02144
Head	Sex + ClutchSize	38.79296	38.74639	38.37928	39.20361	2.845592	2.728489

	vp.low	vp.up	cvp	cvp.mode	cvp.low	cvp.up	va
Tarsus	1.149875	1.778238	4.344749	4.283339	3.89775	4.847318	0.190174
Mass	36.38902	51.66775	17.84305	17.59028	16.23056	19.68247	2.880239
Head	2.359098	3.385684	4.343733	4.300857	3.968953	4.751838	0.343071

	va.mode	va.se	va.low	va.up	cva	cva.mode	cva.low
Tarsus	0.151665	0.084186	0.04011	0.362312	1.538985	1.67087	0.821993
Mass	0.035848	2.035138	4.18E-09	6.645591	4.237356	4.672282	0.418941
Head	0.303741	0.21171	1.45E-07	0.727702	1.423394	1.475582	0.359751

	cva.up	h2.mean	h2.mode	h2.median	h2.se	h2.low	h2.up
Tarsus	2.245018	0.133024	0.112997	0.127697	0.058695	0.028387	0.253071
Mass	7.286765	0.065941	0.000651	0.059346	0.045915	1.02E-10	0.151345
Head	2.386515	0.120303	0.099629	0.112592	0.072297	5.89E-08	0.251649

	va / (mean^2)
Tarsus	0.00025010622
Mass	0.00210269951
Head	0.00022797024

Table S3: Output from the GRM-based *MCMCglmm* run of generalized mixed models for tarsus length (mm), body mass (g) and head-bill length (mm) including 523 hihi. f.effect = fixed effects fitted in the model, mean = trait means, vp = total phenotypic variance, cvp = total coefficient of variation, va = additive genetic variance, cva = additive genetic coefficient of variation, h2 = heritability, va / (mean^2) = evolvability.

Trait	f.effect	mean	mean.mode	mean.low	mean.up	vp	vp.mode
Tarsus	Sex + ClutchSize	27.53591	27.53206	27.26254	27.81346	1.453254	1.408385
Mass	Sex + ClutchSize	36.87699	36.88974	35.59676	38.12623	43.46233	42.25234
Head	Sex + ClutchSize	38.72925	38.6998	38.40582	39.05599	2.850764	2.797517

	vp.low	vp.up	cvp	cvp.mode	cvp.low	cvp.up	va
Tarsus	1.143742	1.810275	4.369806	4.292538	3.901697	4.899467	0.111009
Mass	36.37417	51.64145	17.86391	17.84542	16.25706	19.62097	1.811847
Head	2.346352	3.410876	4.354397	4.32261	3.973366	4.78087	0.090762

	va.mode	va.se	va.low	va.up	cva	cva.mode	cva.low
Tarsus	0.097637	0.060452	1.39E-09	0.218855	1.156055	1.252824	0.429957
Mass	0.021285	1.450521	1.14E-08	4.548253	3.298343	3.726922	0.009635
Head	0.000628	0.090347	1.42E-15	0.270281	0.672703	0.710392	9.74E-08

	cva.up	h2.mean	h2.mode	h2.median	h2.se	h2.low	h2.up
Tarsus	1.862343	0.077121	0.061612	0.073066	0.042292	1.02E-09	0.153476
Mass	5.794001	0.041772	0.000449	0.035736	0.033189	2.44E-10	0.104785
Head	1.34189	0.031791	0.000512	0.022964	0.031232	5.15E-16	0.094075

	va / (mean^2)
Tarsus	0.00014640615
Mass	0.00133232626
Head	0.00006050982

Table S4a*:* The ten most significant SNPs associated with tarsus length after *RepeatABEL* GWAS, their MAF, counts of individuals with each genotype and their SNP effect plus the PIP4 (posterior inclusion probability to be in mixture class 4) value from *BayesR* output.

SNP	Chromosome	Position	A1	A2	MAF	A1A1	A1A2	A2A2	effB	se_effB	P1df
AX-172486735	3	52404441	G	T	0.14053537	389	121	13	0.3672732	0.08399748	1.228707e-05
AX-171754927	1A	53396960	A	G	0.12523901	404	107	12	0.3688445	0.09187671	5.955715e-05
AX-171714808	1A	56256927	A	G	0.07101727	450	68	3	0.4549406	0.11634956	9.225069e-05
AX-171722226	26	378196	G	A	0.28202677	265	221	37	-0.2514108	0.06598627	1.389407e-04
AX-171712697	1A	56205950	T	C	0.06692161	454	68	1	0.4576042	0.12020177	1.406814e-04
AX-171764222	1A	56186866	C	T	0.06692161	454	68	1	0.4576042	0.12020177	1.406814e-04
AX-171757070	4	1553827	G	A	0.32088123	252	205	65	-0.2226038	0.06163531	3.042813e-04
AX-171730907	1A	52785060	G	T	0.32409178	236	235	52	0.2333929	0.06473413	3.116631e-04
AX-171755458	1	94512507	T	C	0.37284895	201	254	68	0.2253723	0.06346394	3.834990e-04
AX-171736922	26	94531	A	G	0.13601533	386	130	6	0.3088710	0.08778972	4.343135e-04

Table S4b*:* The ten most significant SNPs associated with body mass after *RepeatABEL* GWAS, their MAF, counts of individuals with each genotype and their SNP effect plus the PIP4 (posterior inclusion probability to be in mixture class 4) value from *BayesR* output.

SNP	Chromosome	Position	A1	A2	MAF	A1A1	A1A2	A2A2	effB	se_effB	P1df
AX-171747347	9	2151684	T	C	0.23900574	303	190	30	1.500777	0.3979115	0.0001621817
AX-171750282	15	5233802	G	A	0.18007663	354	148	20	1.671526	0.4477722	0.0001892249
AX-171734448	9	2151452	C	T	0.23800384	302	190	29	1.492961	0.4000869	0.0001902735
AX-171750191	10	9200724	C	T	0.24031008	291	202	23	1.462740	0.4122596	0.0003880159
AX-172485276	14	13725430	A	C	0.21839080	312	192	18	-1.468594	0.4156804	0.0004108891
AX-171716117	14	13678108	T	C	0.21743295	312	193	17	-1.467433	0.4173279	0.0004376733
AX-171728749	2	119853640	C	T	0.04022989	480	42	0	2.950912	0.8418522	0.0004561593
AX-171722003	14	13726090	A	G	0.21892925	312	193	18	-1.442127	0.4151397	0.0005130767
AX-171727176	14	13733955	A	G	0.21892925	312	193	18	-1.442127	0.4151397	0.0005130767
AX-171736667	14	13667736	G	A	0.21892925	312	193	18	-1.442127	0.4151397	0.0005130767

Table S4c*:* The ten most significant SNPs associated with head-bill length after *RepeatABEL* GWAS, their MAF, counts of individuals with each genotype and their SNP effect plus the PIP4 (posterior inclusion probability to be in mixture class 4) value from *BayesR* output.

SNP	Chromosome	Position	A1	A2	MAF	A1A1	A1A2	A2A2	effB	se_effB	P1df
AX-171740203	1	8013170	G	A	0.21032505	324	178	21	-0.4716194	0.1126187	2.817302e-05
AX-171759269	1A	70506060	G	A	0.18103448	345	165	12	0.5078508	0.1223601	3.318053e-05
AX-171742068	20	269524	G	A	0.27063340	277	206	38	-0.4113054	0.1054804	9.644949e-05
AX-171714808	1A	56256927	A	G	0.07101727	450	68	3	0.6633574	0.1826216	2.807810e-04
AX-171714154	2	117460692	T	C	0.09464627	429	89	5	0.5433429	0.1538614	4.133998e-04
AX-171730915	2	117426607	G	A	0.09464627	429	89	5	0.5433429	0.1538614	4.133998e-04
AX-171746228	77	46	A	G	0.09464627	429	89	5	0.5433429	0.1538614	4.133998e-04
AX-172485349	2	117409659	A	G	0.09464627	429	89	5	0.5433429	0.1538614	4.133998e-04
AX-171733157	1	26777224	C	A	0.13557692	386	127	7	0.4859009	0.1394753	4.943908e-04
AX-171721340	2	138161355	A	G	0.25239006	284	214	25	0.3844585	0.1107671	5.187710e-04

Table S5 Summary of the output from the *BayesR* analysis (50,000 iterations and burn-in of 20,000) on the corrected phenotypes of 523 hihi (*top half of the table*). In addition, we report the estimates for the SNP-based heritability (h^2_{BayesR}) for tarsus length, body mass and head-bill length, the total proportion of SNPs contributing to trait variation and the proportion of the number of SNPs assigned to the different effect size distributions of 0.0001, 0.001 and 0.01 (*bottom half of the table*).

	Number of SNPs in model	**Sum of squares of SNP effects in mixture component 2 (0.0001)**	**Sum of squares of SNP effects in mixture component 3 (0.001)**	**Sum of squares of SNP effects in mixture component 4 (0.01)**	**Genetic variance explained by SNPs (Va)**	**Residual variance (Ve)**
Tarsus length	3,431	0.0236	0.0237	0.0119	0.059	0.4460
Body mass	3,393	0.4121	0.4244	0.2729	1.110	12.960
Head-bill length	3,598	0.0405	0.03806	0.0279	0.100	1.2189

	h^2_{BayesR}	**Effect size distribution**			
		0.0001	**0.001**	**0.01**	**Total**
Tarsus length	0.117	0.077	0.009	0.001	0.086
Body mass	0.079	0.076	0.008	0.001	0.085
Head-bill length	0.076	0.082	0.007	0.001	0.091

Table S6 Power for association analyses with (*) as calculated from *RepeatABEL*. Following code provided by Wang and Xu (2019), we calculate the power for association from a mixed model analysis as a function of sample size (n = 523), the heritability of the locus (h^2_locus), the ratio of the additive genetic variance to the residual variance (lambda), the correlation between any pair of individuals in the kinship matrix (rho), the number of markers (m = 39,699) and the significance threshold (alpha = 0.05). This calculation gives the power to detect association when the causal variant is typed. Therefore, following Delongchamp et al. (2018), we adjust the sample size to the effective sample size (n') accounting for the average linkage disequilibrium between neighbouring markers (r^2 = 0.541) as n' = n * r^2 = 523 * 0.541 = 270.914. The average r^2 was calculated in plink as described in the main text while rho was calculated as the mean of relatedness values in the *RepeatABEL* GRM (rho = -3.090533e-06). To calculate the maximal power for association, for each trait we assume that the trait heritability is fully explained by one causal locus (i.e. h2_locus = total heritability).

Trait	h^2_locus *	Lambda *	Power
Tarsus length	0.081	0.152	0.516
Body mass	0.065	0.126	0.306
Head-bill length	0.071	0.113	0.371

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

[2]Delongchamp R, Faramawi MF, Feingold E, Chung D, Abouelenein S. The Association between SNPs and a Quantitative Trait: Power Calculation. Eur J Environ Public Health. 2018;2(2):10.

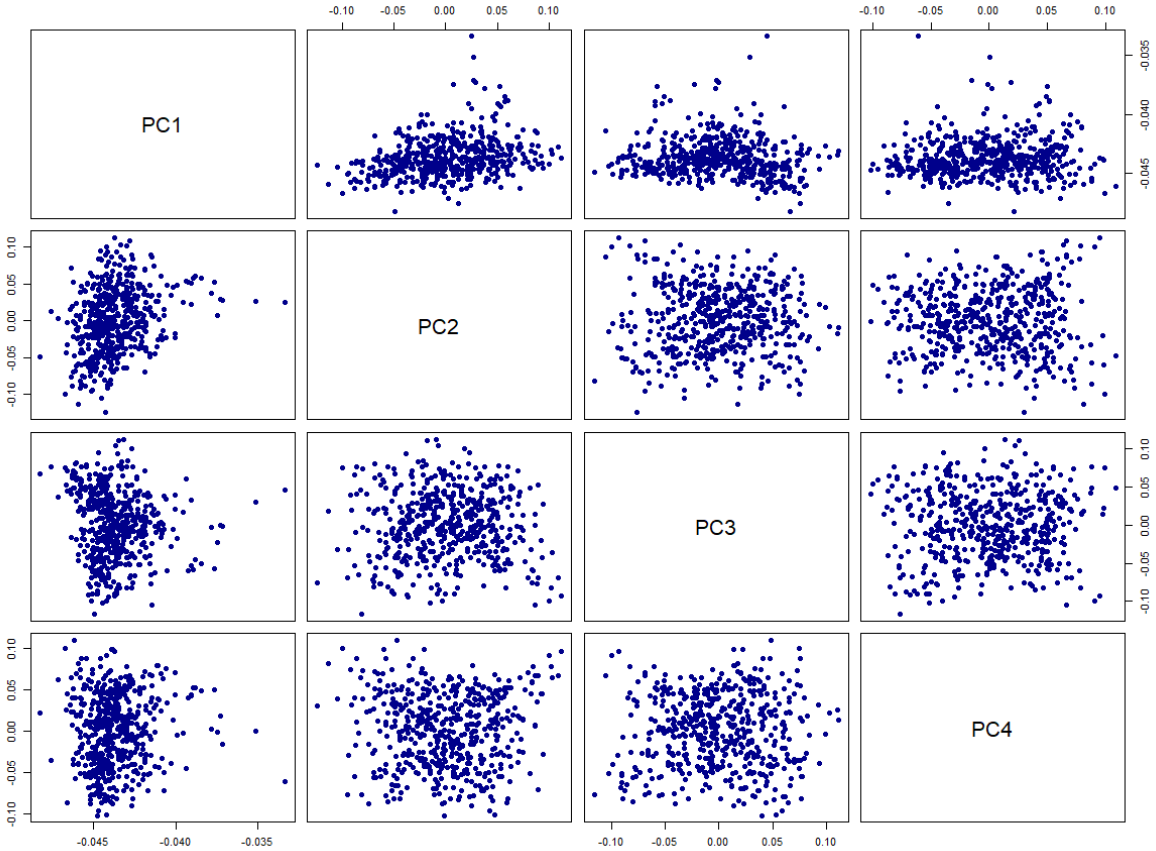


Figure S1: Plotted output from a principal component analysis using the 523 hihi genotypes from Tiritiri Matangi island (eigenvalues generated in *PLINK* using the –pca function). Shown are the first four principle components, explaining a proportion of 0.28 (PC1), 0.2 (PC2), 0.2 (PC3) and 0.2 (PC4) of the variance, respectively, revealing no with-population structure.

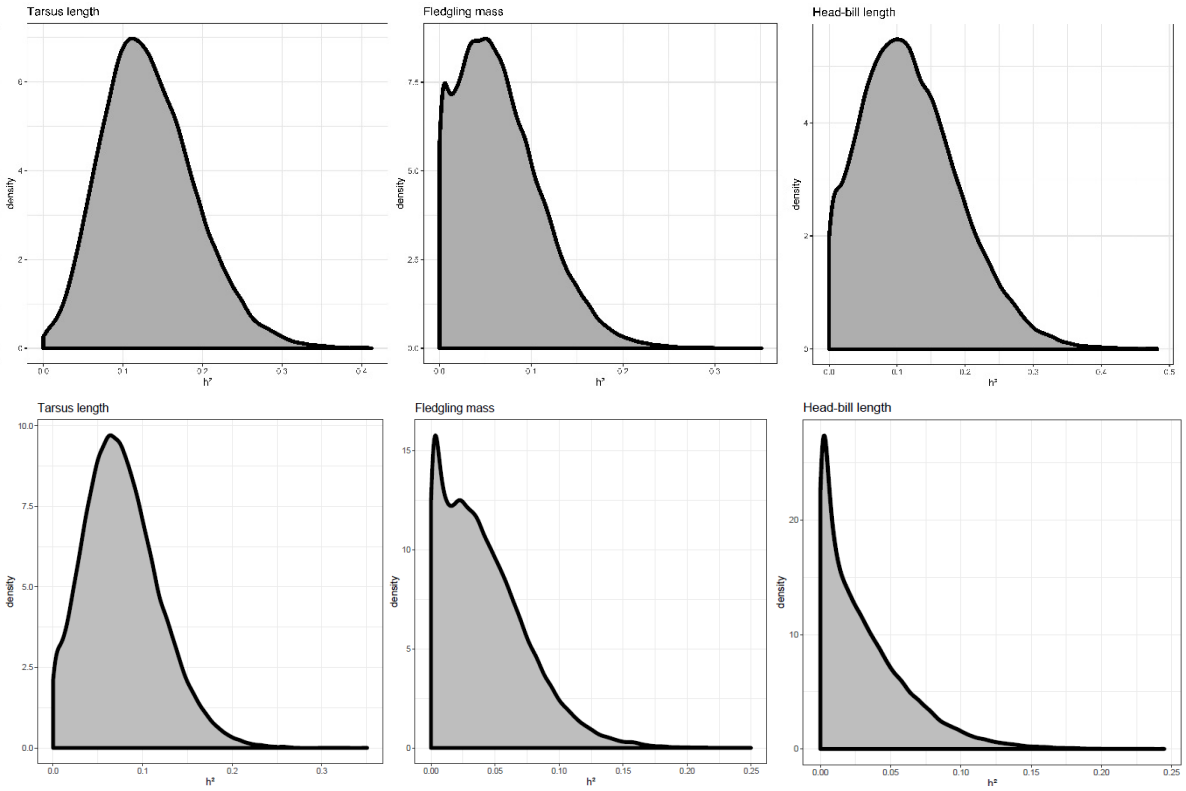


Figure S1: Density graph for the *MCMCglmm* heritability estimates using the pedigree (top row) and the genomic relatedness matrix (h^2_{GRM1}) (bottom row) for all three morphological traits using the model described in the main text. Plotted using the cairo_pdf/plot_grid function in R.

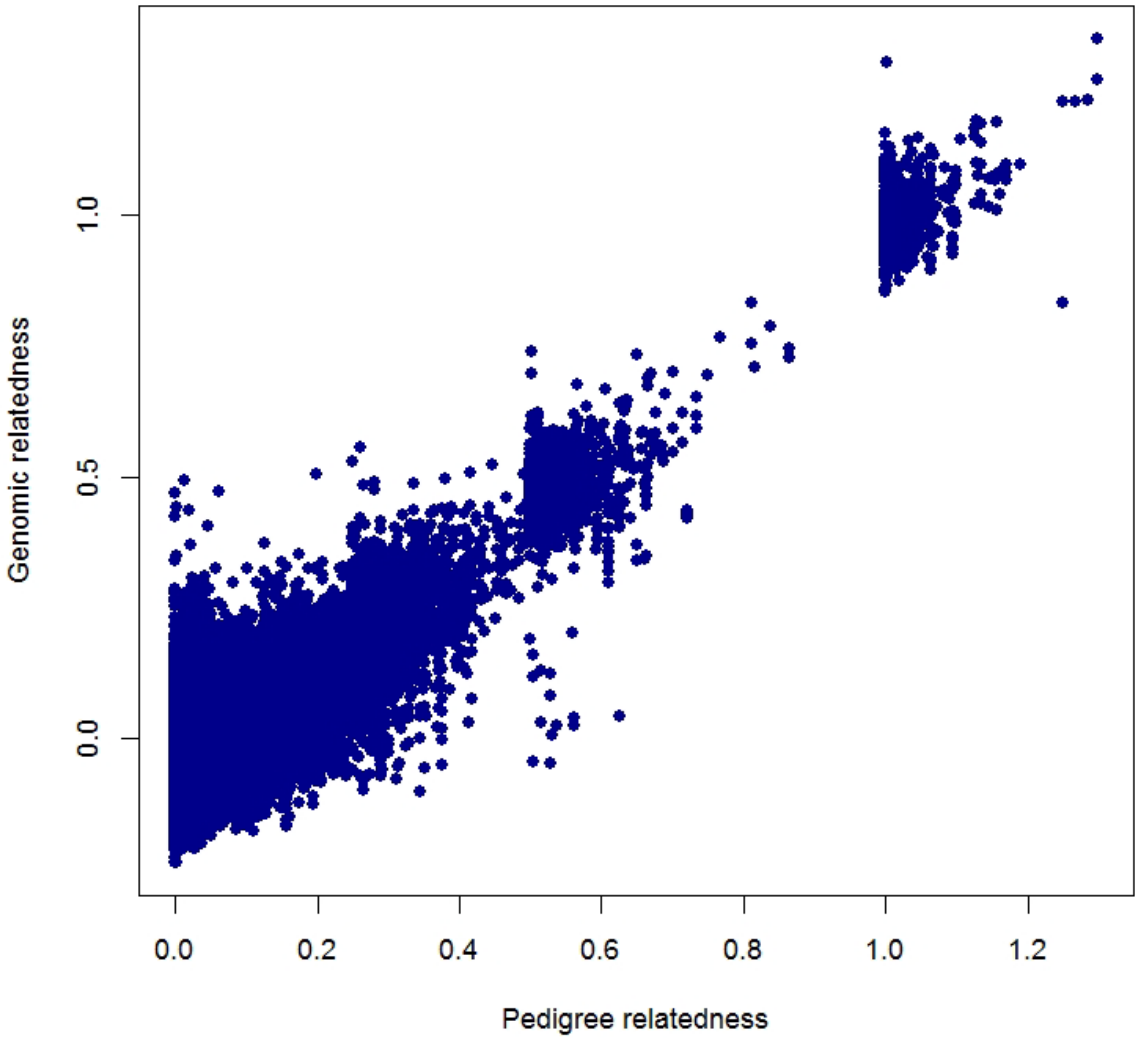
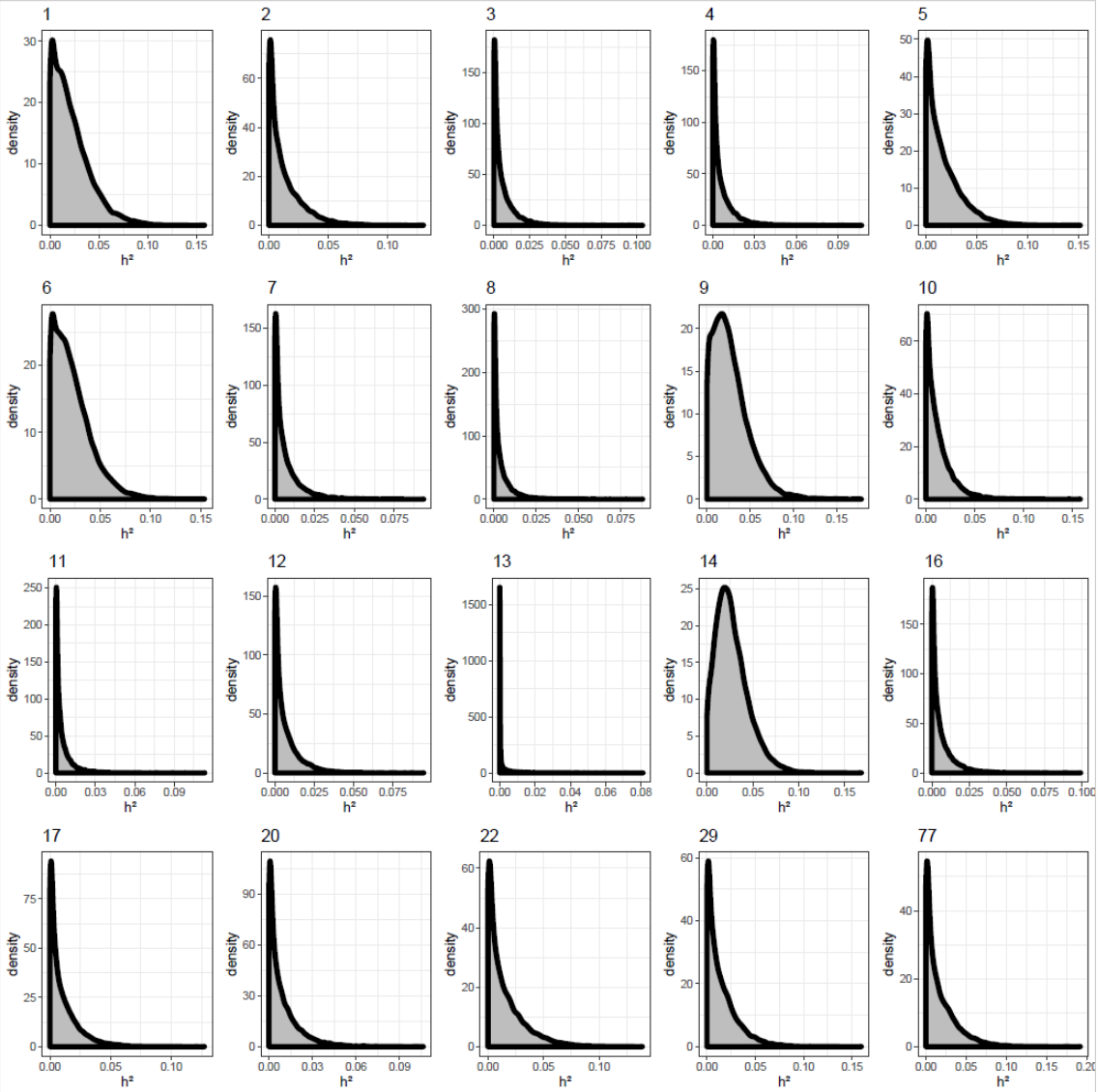
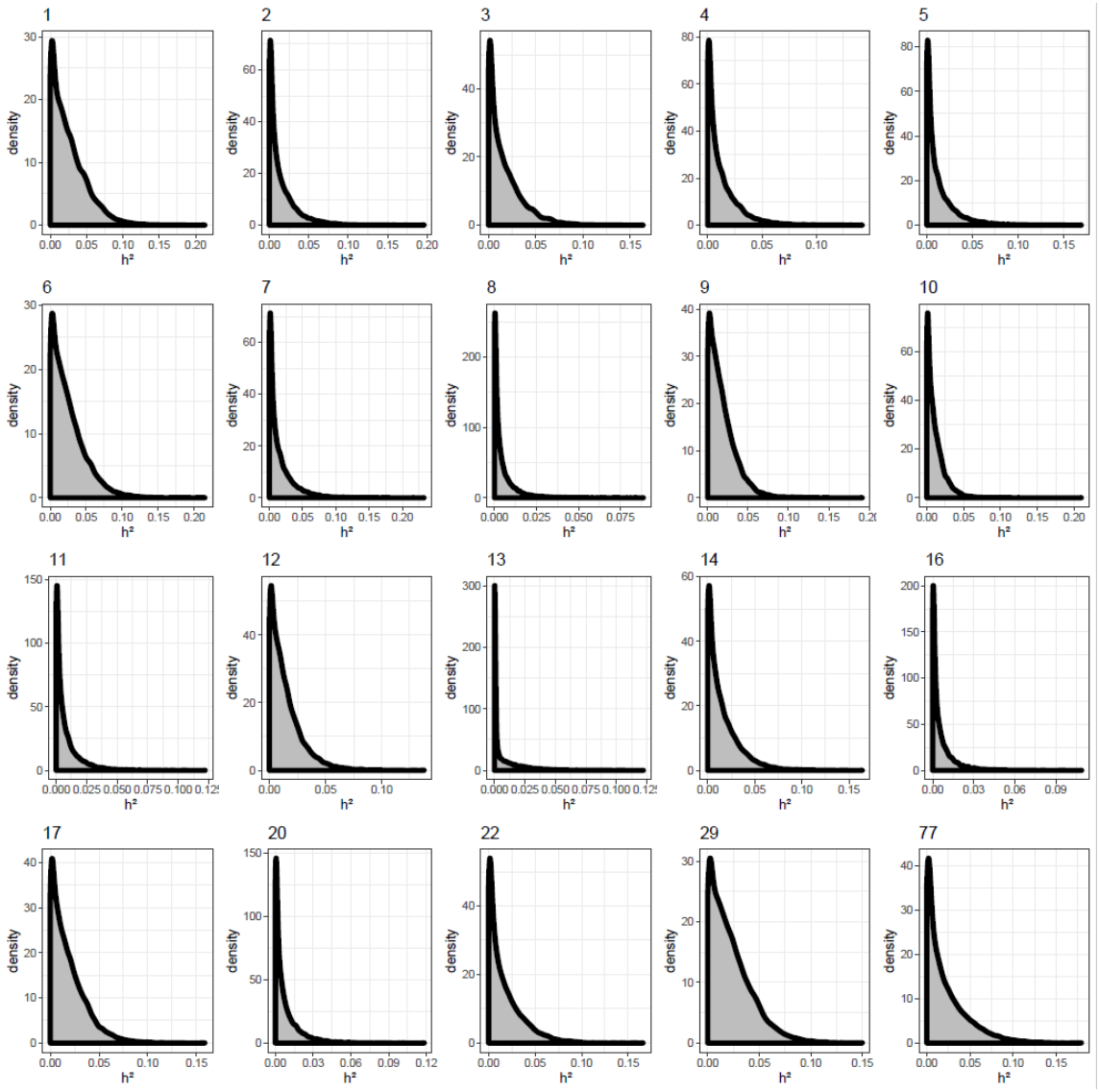


Figure S3: Graph showing the correlation between the pedigree-based relatedness matrix (A) and the genomic relatedness matrix (GRM) based on 39,699 SNPs for 523 hihi individuals. The overall correlation coefficient is 0.847 for all pairs of individuals (i.e. Pearson's correlation of the upper triangle of the relatedness matrices, including the diagonal), and 0.749 if self-self relatedness is excluded (i.e. Pearson's correlation of the upper triangle of the relatedness matrices, excluding the diagonal).

a) Tarsus length



b) Body mass



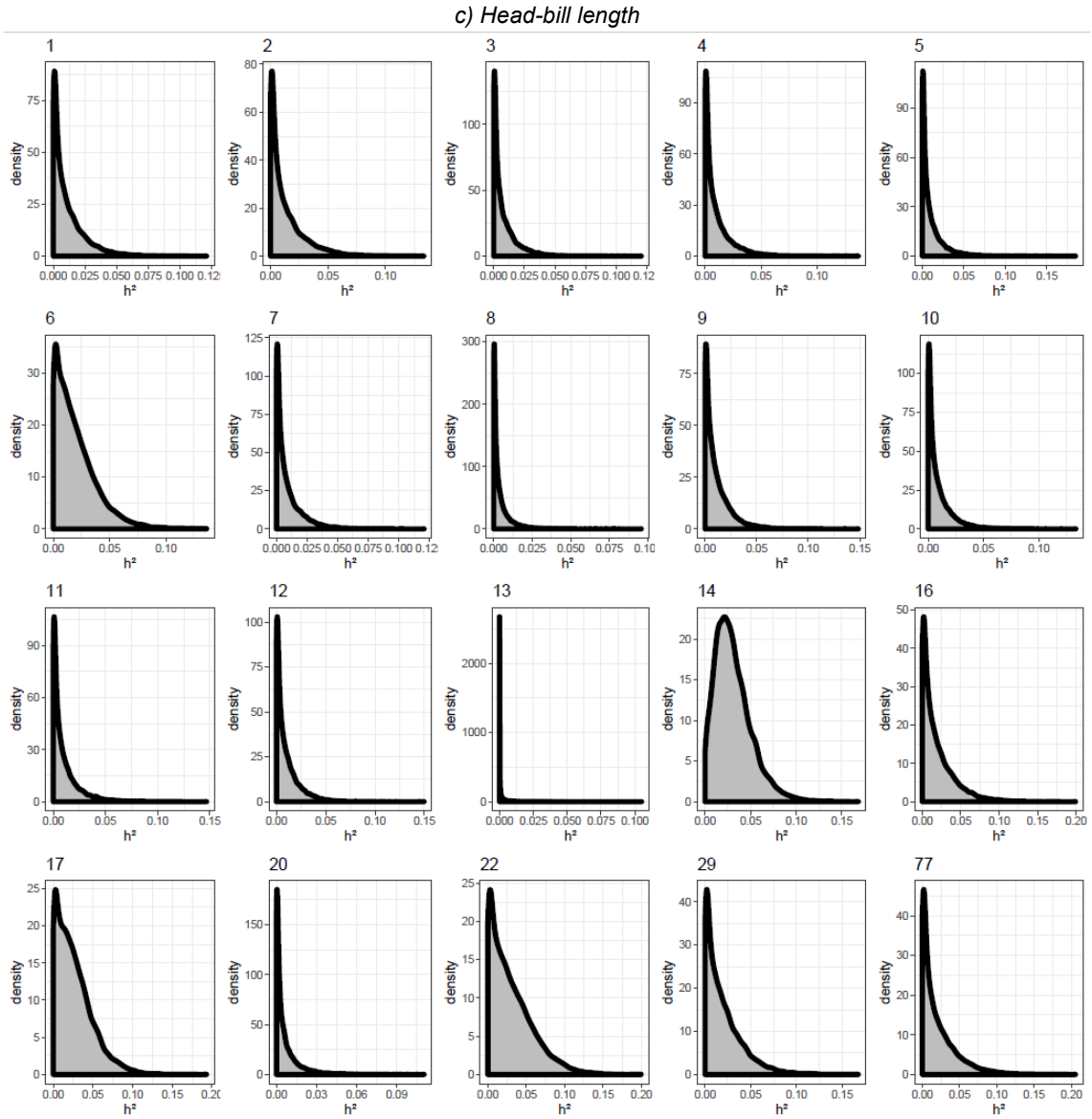


Figure S4: Density graphs of the *MCMCglmm* estimates for chromosome heritabilities for all three morphological traits after chromosome partitioning using the model described in the main text. Plotted using the cairo_pdf/plot_grid function in R. Chromosome numbers: 1-14 are estimated from markers mapped to zebra finch chromosomes 1-14, 16 = chromosome 4A and chromosome 15, 17 = chromosomes 17, 18 and 19, 20 = chromosome 20, 22 = chromosomes 21-28, 1B and LGE422, 29 = chromosome 1A, 77 = markers mapped to zebra finch Unknown or random chromosomes or with no alignment to the zebra finch genome.

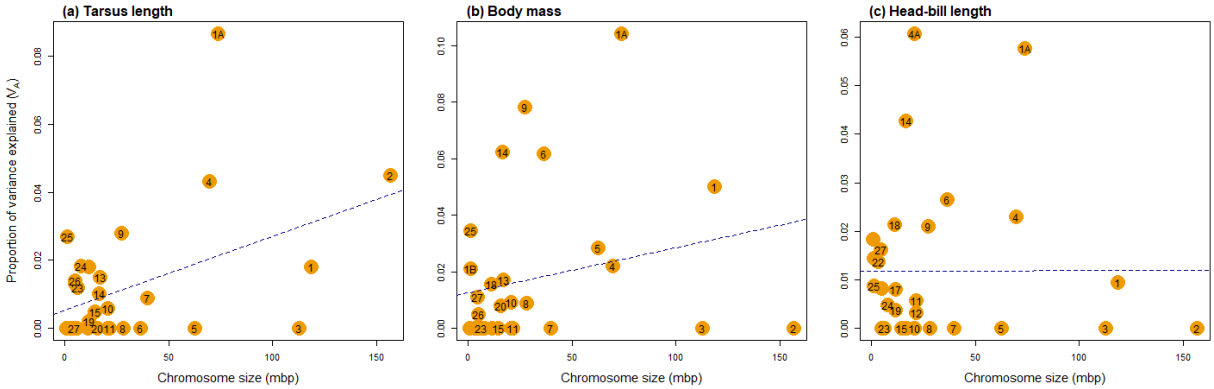


Figure S5: The proportion of variation explained by each chromosome (in yellow circles, except for 'Chr 77') for all three traits (tarsus length, body mass and head-bill length) as calculated in *GCTA* for the 523 hihi. Plotted is the mean additive genetic variance explained by each chromosome (V_A) against the length of the chromosome (in Mbp) based on the zebra finch reference genome (overlapping chromosome not shown). Raw phenotypes were used (only corrected for fixed effects) to avoid convergence errors and hence be able to include all chromosomes. After HC correction all p-values were non-significant (0.119402985, 0.3485535 and 0.6963788).

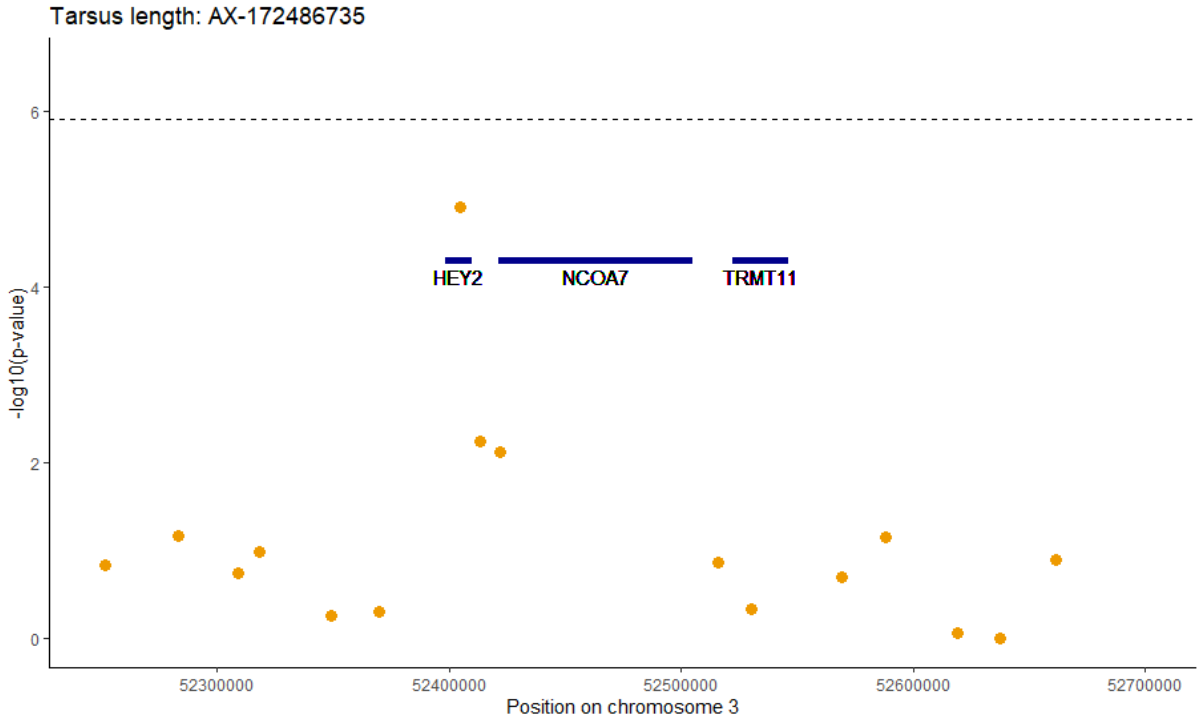


Figure S2: Zoom into the genomic area around the most significant SNP for tarsus length (p= 1.23e-05) according to the GWAS analysis in *RepeatABEL*. Gene annotations appear according to the annotated zebra finch genome.

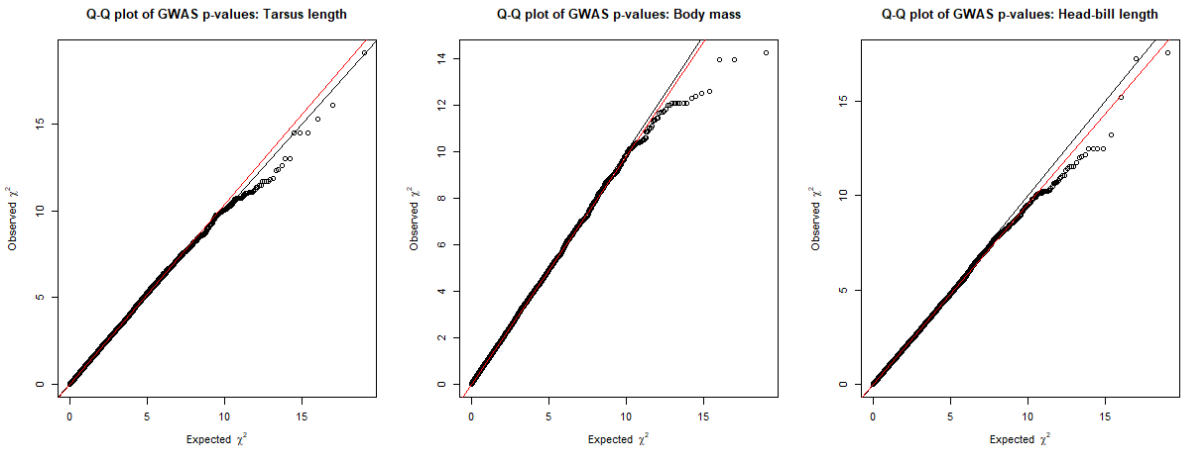


Figure S7: Q-Q plot of the p-value distributions for all the *RepeatABEL* GWAS runs using the "prefitModel" function. Lambda (genomic inflation) values for tarsus length: 1.036452; body mass: 0.9827692; head-bill length: 0.9572672.