Using the *Mus musculus* hybrid zone to assess covariation and genetic architecture of limb bone lengths

Neva Škrabar[1], Leslie M. Turner[1,3], Luisa F. Pallares[1,2], Bettina Harr[1], Diethard Tautz[1*]

[1] Max-Planck Institute for Evolutionary Biology, August-Thienemann-Str. 2, Plön 24306, Germany

[2] Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ 08544, USA

[3] Milner Centre for Evolution, Department of Biology and Biochemistry, University of Bath, Bath BA2 7AY, United Kingdom

* corresponding author: tautz@evolbio.mpg.de

running title: Mouse hybrid zone association mapping

ABSTRACT

Two subspecies of the house mouse, *Mus musculus domesticus* and *Mus musculus musculus*, meet in a narrow contact zone across Europe. Mice in the hybrid zone are highly admixed, representing the full range of mixed ancestry from the two subspecies. Given the distinct morphologies of these subspecies, these natural hybrids can be used for genome-wide association mapping at sufficiently high resolution to directly infer candidate genes. We focus here on limb bone length differences, which is of special interest for understanding the evolution of developmentally correlated traits. We used 172 first-generation descendants of wild-caught mice from the hybrid zone to measure the length of stylopod (humerus / femur), zeugopod (ulna / tibia) and autopod (metacarpal / metatarsal) elements in skeletal CT scans. We find phenotypic covariation between limb elements in the hybrids similar to patterns previously described in *M. m. domesticus* inbred strains, suggesting that the hybrid genotypes do not influence the covariation pattern in a major way. Mapping was performed using 143,592 SNPs and identified

several genomic regions associated with length differences in each bone. Bone length was found to be highly polygenic. None of the candidate regions include the canonical genes known to control embryonic limb development. Instead, we are able to identify candidate genes with known roles in osteoblast differentiation and bone structure determination, as well as recently evolved genes of, as yet, unknown function.

INTRODUCTION

Vertebrate limbs are an excellent model system to study morphological integration, developmental stability and phenotypic covariance patterns (Lande 1980; Hallgrimsson et al. 2002; Young and Hallgrimsson 2005; Schmidt and Fischer 2009; Rolian et al. 2010; Kolarov et al. 2011; Pavlicev et al. 2013; Young 2013). The forelimb and hind limb represent serially homologous structures, i.e. repeated parts that share a developmental architecture. During evolution, serially homologous structures are formed when the same underlying developmental program is expressed at two (or more) different locations along the body (Hall 1995). While each of the structures may diverge over time, they are still expected to show phenotypic and genetic covariation, due to the same general underlying developmental program. Hence, genetic variants are expected to have pleiotropic effects on more than one limb bone. Such pleiotropy is expected to constrain evolutionary divergence and adaptation. On the other hand, limbs have also often been subject to specialization during evolution (Young and Hallgrimsson 2005; Schmidt and Fischer 2009), suggesting that parts of the genetic architecture must be sufficiently free to allow such specializations. Accordingly, one should expect that there is genetic variation that determines the evolvability of the structures (Hansen et al. 2003; Hansen 2006). The pleiotropic effect of loci that affect two structures in parallel can be modified by loci that act on only one of the structures. The latter loci have been called relationship quantitative trait loci (*r*QTLs) (Cheverud et al. 2004; Pavlicev et al. 2008, 2013). The identification of *r*QTLs is of particular interest for serially homologous structures, such as the limbs, since they are thought to underlie their evolutionary specialization, such as the morphological differentiation that is evident between the forelimbs and hind limbs in mice (Figure 1).

QTLs influencing bone length have been mapped in various studies of laboratory mouse strains derived from the subspecies *M. m. domesticus* (Wolf et al, 2006; Norgard et al. 2008, 2009; Pavlicev et al. 2008; Pavlicev et al. 2013; Parmenter et al. 2016 - and references therein). These studies found a high degree of covariation between traits in forelimbs and hind limbs and pleiotropy of QTLs, including correlations with body weight (Parmenter et al. 2016). rQTLs can be investigated in this system through mapping the genomic regions that influence a phenotypic trait depending on the presence of another trait. Pavlicev et al. (2008 and 2013) used this approach to identify loci conveying variational independence between limb bones.

Employing a standard F2 QTL mapping approach for detecting loci connected to length variation of limb bones does not provide sufficient resolution to identify underlying causative genes (Flint et al. 2005). Although fine-mapping approaches using advanced intercrosses provide further resolution (e.g. Norgard et al. 2009), they require extensive breeding efforts. With the advent of high-throughput marker technology, it has become possible to make use of variation present in natural populations through association mapping (Flint et al. 2012). Natural populations differ from controlled crosses by longer recombination histories and thus lower levels of linkage disequilibrium, facilitating much higher mapping resolution. Typical levels of linkage disequilibrium in natural populations of mice enable direct identification of candidate genes (Laurie et al. 2007). Genome-wide association studies (GWAS) provide insights into the genetic basis of natural variation as well as in understanding complex traits. In GWAS, tests for association between each genetic variant, typically single-nucleotide polymorphisms (SNPs), and a phenotype of interest are performed (Bush et al. 2012), while controlling for population structure which can cause spurious associations (Sul et al. 2016). In addition, the proportion of phenotypic variance attributed to additive genetic effects can be estimated in this framework (Price et al. 2010).

Hybrid zones between subspecies are of particular interest for mapping, since they represent natural cases of advanced intercrosses (Rieseberg and Buerkle 2002). There are in fact multiple advantages for using hybrid zones for mapping. First, because they are formed between two evolutionary distinct lineages, they should harbor more phenotypic and genetic variation than single populations, which should increase the power for mapping. Second, it is expected that LD is lower than in family intercrosses, but still higher than in fully outbred situations, i.e. a

relatively lower marker density is required compared to standard genome wide association studies. Finally, hybrid zone mapping should be particularly effective for finding loci that are actually causative in the evolutionary distinction between the corresponding lineages.

In previous studies we used samples from the house mouse hybrid zone for mapping of hybrid sterility phenotypes (Turner and Harr 2014) and craniofacial traits (Pallares et al. 2014). The Western and Eastern house mouse subspecies (*M. m. domesticus* and *M. m. musculus*) form a narrow hybrid zone in the middle of Europe (Guenet and Bonhomme 2003; Phifer-Rixey and Nachman 2015). Mice in the hybrid zone are naturally admixed and linkage disequilibrium is sufficiently low to allow high resolution mapping, including the identification of individual candidate genes (Turner and Harr 2014; Pallares et al. 2014). Further, the continuous transition in genomic composition from one subspecies to the other allows inferences about the genetic architecture and evolution of morphological traits (Pallares et al. 2016).

Limbs in tetrapods are divided into three segments, from proximal to distal: the stylopod (humerus / femur) the zeugopod (ulna / tibia) and the autopod (metacarpal / metatarsal) (Figure 1). The goal of this study is to explore the genetic architecture controlling the length of individual limb bones, as well as their joint variation. We expect that homologous elements of forelimbs and hind limbs, and bones that belong to the same segment, i.e. stylopod, zeugopod or autopod share common developmental genetic networks. On the other hand, since forelimbs and hind limbs have different lengths and somewhat different morphological functions (Figure 1), we asked whether we can find genetic regions that contribute to length variation in single limb elements. We show that it is indeed possible to identify candidate genes that have previously not been implicated in limb length determination, but are known to be involved in bone structure determination. Candidate regions are identified in both, single bone measures, as well as with a second bone included as a covariate (rQTLs). Together, our results also point towards a general polygenic architecture for limb length determination in mice.

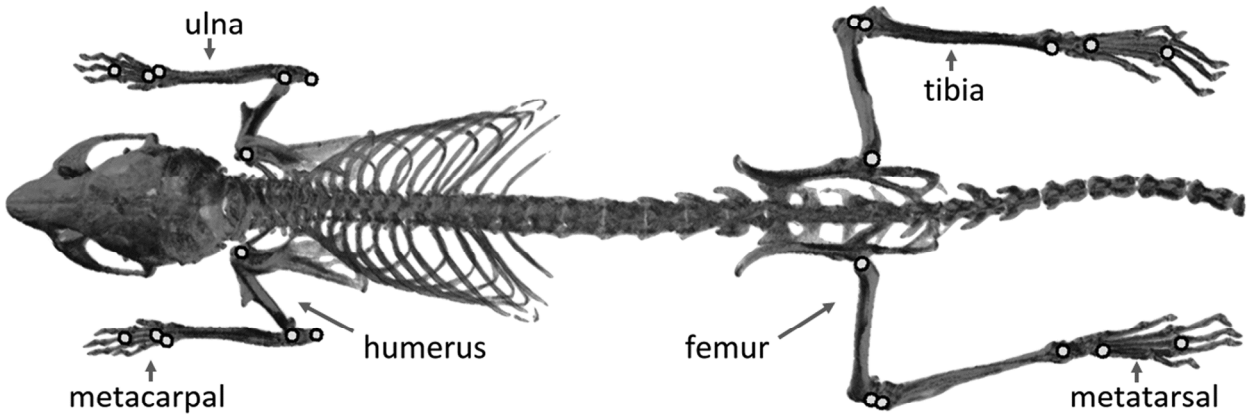


Figure 1. CT-scan of a mouse showing the limb bones measured in this study. The approximate positions of landmarks are indicated by dots. Note that the actual landmarks were set in 3D representations of the skeleton.

METHODS

Mapping population

Individuals included in this study are first-generation offspring of wild caught mice collected in the hybrid zone in Bavaria in 2008 (Turner et al. 2012). The sampling procedure and breeding was previously described in Turner et al. (2012) and Turner and Harr (2014). Mice were raised under standard laboratory conditions to reduce the environmental effect on the traits. They were sacrificed by CO_2 asphyxiation between 9 and 12 weeks of age. The mapping population consists of 172 males including full-siblings, half-siblings and unrelated individuals (Supplementary Table 1). Genome-wide association studies of sterility traits (Turner and Harr 2014) and craniofacial traits (Pallares et al. 2014) were previously reported for these mice.

Phenotype measurements

Mice were scanned with a computer tomograph (micro-CT-vivaCT 40, Scanco, Bruettisellen, Switzerland) with the following settings - energy: 70 kVp, intensity: 114 µA, voxelsize: 38 µm. We generated three-dimensional cross-sections with a resolution of one cross-section per 0.038 mm. The images were transformed into the DICOM (Digital Imaging and Communications in Medicine) format and landmarks were placed within the 3D representation at the endpoints of limb bones in the forelimb (humerus, ulna, metacarpal bone) and in the hind limb (femur, tibia and metatarsal bone) using the TINA landmarking tool (Schunke et al. 2012). Two landmarks were used per left and right limb bone and linear measurements were obtained for each pair of landmarks. Description of landmarks in proximo-distal direction: humerus: from the humeral head to the medial point of the trochlea; ulna: from the most proximal point of the olecranon to the styloid process; 3rd metacarpal bone: from the capitate-metacarpal articular surface of the base to the head; femur: from the greater trochanter to the articular surface for the patella; tibia from the intercondyloid eminence of medial condyle to the articular surface with talus, 3rd metatarsal bone from the articulate surface of the base to the head (Bab et al. 2007). The approximate positions of the landmarks are shown in Figure 1.

Measurement error was estimated based on double measurements of the same image in fifty individuals. The percentage of measurement error was calculated according to the ANOVA design described in Yezerinac et al. (1992) as the ratio of the within-measurement component of variance to the sum of the within- and among-measurement components (Claude 2008).

No significant differences between right and left sides of the corresponding limb bones were found (Supplementary Table 2). Because some individuals had partially damaged bones due to previous handling of the samples, all bones from the affected side for these individuals were excluded from analysis. Measurements from both sides of intact individuals (N=136) (i.e. specimens with preserved left and right side) were averaged and individuals with either a complete right (N=16) or a complete left (N=20) side were added, i.e. the animals with only one intact side are represented with only one measurement. A total of 172 individuals were included in the dataset for association mapping.

The individual bone lengths were normally distributed (Supplementary Tables 3 and 4, Supplementary Figures 1 and 2). Age showed small but significant correlations with metacarpal (r^2 = 0.026, p = 0.036, slope = - 0.003) and metatarsal (r^2 = 0.045, p = 0.005, slope = - 0.011) measurements. Hence, for consistency, we performed linear regressions of length on age for each bone (i.e. each phenotypic trait entered the model as dependent variable, while age was independent variable) and used residuals for further analysis. All statistical analyses for phenotypes were performed in R version 3.2.5 (R Core Team 2016).

Pearson (product-moment) correlation coefficients were calculated to estimate the strength of relationships between individual limb bones. In order to investigate the transition in limb length across the hybrid zone (for the results shown in Figure 2), the forelimb was represented as the sum of humerus, ulna and metacarpal bones, and hind limb as the sum of femur, tibia and metatarsal bones. Following the procedure described in Pallares et al. (2016) and based on a specific SNP dataset from 37 loci, we performed regressions of forelimb and hind limb lengths versus percentage of *M. m. domesticus* alleles ("PairMid" in Supplementary Table 1).

Association mapping

SNP genotype data generated using the Mouse Diversity Genotyping Array (Affymetrix, Santa Clara, CA, USA) (Yang et al. 2009) were previously reported in Turner and Harr (2014) and Pallares et al. (2014). SNP positions were converted to the coordinates of the GRCm38/mm10 assembly of the mouse genome using the LiftOver tool in the UCSC Genome Browser (Kent et al. 2002). SNPs were pruned in PLINK (Purcell et al. 2007) using a sliding window approach with the following settings, 30 SNPs window size, 5 SNPs step size and a VIF threshold of 1×10^{-6} (VIF = $1/(1-R^2)$) (see Turner and Harr 2014 for details). Essentially, this procedure removed nearby SNPs in strong LD from the dataset. The SNP genotype dataset contained genotype data for 185 individuals, which were filtered for SNPs $\geq$ 5% minor allele frequency (MAF). In order to use the same set of SNPs as in Turner and Harr (2014) and Pallares et al. (2014) for the present study, which involved a subset of only 172 individuals, and therefore could change the minor allele frequencies slightly, an additional MAF filter of 1% (default option in GEMMA) was imposed for mapping. In total, 143,592 SNPs (of 285,625 SNPs called from the array) were used for mapping in this study.

For association mapping, we implemented the univariate linear mixed model analysis in genome-wide efficient mixed-model association - GEMMA (version 0.94.1) (Zhou and Stephens 2012), using the Wald test (–lmm 1). With this approach, association between a marker and a single phenotype is tested using a variance component model that corrects for relatedness and population structure (Kang et al. 2010) by incorporating a relatedness matrix as a random effect. All LD-pruned SNPs from autosomes were included in the calculation of the centered kinship matrix. Because we used only males, the number of X-chromosomes sampled is half that of autosomes. We found no significant associations with SNPs on the X chromosome in preliminary mapping analyses, and therefore excluded it from further analysis.

We investigated the general genetic architecture of limb bone lengths using Bayesian sparse linear mixed models (BSLMM), a hybrid approach that simultaneously allows for a small number of individually large genetic effects and combined effects of many small genetic effects, with the relative contributions of both being inferred from the data itself (Zhou et al. 2013). Using BSLMM, we estimated PVE, the proportion of variance in phenotypes explained by all available SNPs and PGE, which is the proportion of total genetic variance explained by "large" effect size SNPs. Data were fitted with a standard linear BSLMM (–bslmm 1), using 500K burn-in steps followed by 5 million sampling steps. Posterior samples for the hyper-parameters (PVE, PGE) were recorded for every 10^{th} iteration (see GEMMA manual for details, and Zhou et al. 2013). The respective median values for PVE and PGE are reported from the second half of the sampling iterations (Wheeler et al. 2016).

To identify possible rQTLs specific for a given limb bone (Pavlicev et al. 2008), we fit univariate linear mixed models in GEMMA to test for the genetic effect of each SNP on one trait, including another phenotype in the model as a covariate. This analysis was conducted in both directions, i.e. traits were exchanged in the position of a response variable and the covariate (Pavlicev et al. 2013). Wald tests were used in mapping (–lmm 1), as described above for single bones. The specific combinations of bones used in the analysis with a covariate were selected based on higher phenotypic and genetic correlations (see Results). Therefore, the combinations included mostly bones within one limb (humerus - ulna, femur - tibia) and bones between limbs (humerus - femur, ulna - tibia, humerus - tibia, ulna - femur and metacarpal - metatarsal). We did not perform multivariate mapping (Zhou and Stephens

2014) of bone phenotypes because preliminary mapping analyses showed our sample size was too small to reliably reach convergence.

Statistical analysis

To define genome-wide significance thresholds for each individual limb bone, we randomly assigned (10,000 times) phenotypes to individuals (thus preserving genetic structure), and performed mapping in GEMMA, recording the lowest SNP association p-value for each permuted dataset. The significance thresholds for each bone were then defined as the 5^{th} percentile of values for 10,000 permutations. For mapping analyses with a second bone as a covariate, the focal pair of phenotypes from each individual were kept together (i. e. humerus and ulna as a covariate from one individual were randomized always in a pair). As for single bones, the lowest p-value was recorded for each permutation and significance thresholds defined as the 5^{th} percentile of values for 10,000 permutations for each analysis.

We also visually inspected the Manhattan plots for obvious "peaks", i.e. clusters of SNPs with low p-values and selected the lowest threshold that would contain all of the obvious peaks. In this way, $p < 10^{-5}$ was selected as a cutoff. A significance threshold of $p < 10^{-5}$ is expected to yield one false positive SNP association per trait in our dataset of 143,592 SNPs. In the supplementary material we provide extended tables with SNPs significant with a threshold of $p < 10^{-4}$; because this is expected to yield 14 false positives per trait in our dataset, we do not analyze these SNPs further.

For each significant SNP, we determined an "LD region" surrounding it by selecting the furthest SNP within 1 Mb upstream and downstream in strong LD ($r^2 \geq 0.8$) with the focal SNP. These LD analyses were performed using the full genotype dataset (i.e. 285,625 SNPs prior to LD pruning) in PLINK (version v1.90b3.32) (Purcell et al. 2007). For significant SNPs without any other SNP in LD within 1 Mb we used the median region size as an approximation. Candidate gene annotation in the LD regions containing significant SNPs was performed using the UCSC Genome Browser (GRCm38/mm10) (Kent et al. 2002) and MGI database (Blake et al. 2017).

215

Genetic correlation

We used bivariate restricted maximum likelihood analysis implemented in the genome-wide complex trait analysis - GCTA (version 1.26.0) (Yang et al. 2010; Yang et al. 2011b) to estimate the genetic correlation between all combinations of two traits. This analysis essentially compares the phenotypic similarity and the genetic similarity

220 between individuals within and across two traits. The genetic correlation (r_g) is defined as $r_g = \frac{cov_g(t1,t2)}{\sqrt{var_g(t1) * var_g(t2)}}$, where ($var_g(t_i)$) is the additive genetic variance of trait *i* and covariance ($cov_g(t_i, t_j)$) is the additive genetic covariance between the traits. The variances and covariances are estimated directly by REML in GCTA (Visscher et al. 2014).

The genetic correlation was examined between each pair of traits including ten principal component axes (i. e. the first ten eigenvectors of the principal component analysis) from the genetic relationship matrix to account for

225 population structure (Yang et al. 2011b). A likelihood-ratio test was applied to determine whether traits are genetically similar, by setting the value of the genetic correlation coefficient to zero (no genetic correlation). Additionally, we tested whether the correlation coefficient equals one, which would mean identical genetic background among two traits (Deary et al. 2012) (note that the value of exactly one would never be reached due to experimental variances). It is important to note that estimates of genetic correlations have high standard errors,

230 especially in small samples such as ours (Visscher et al. 2014).

Phenotypic variance explained by each chromosome

Chromosomal partitioning of variance was performed for each trait using the software GCTA. The variance in bone length explained by each chromosome was calculated with restricted maximum-likelihood analysis. Separate

235 analyses were performed for each bone and for each of the 19 autosomes, by including only data from one chromosome in the model. As above, the first ten principal components from the genetic relationship matrix were included as covariates to account for the effect of variance due to population structure. Individual per-chromosome variance estimates were inflated because of relatedness among individuals; therefore, we estimated the relative contribution of each chromosome to overall trait variance (Pallares et al. 2014).

240

RESULTS

Bone length measures and correlations

Mean bone length measurements with standard deviations and measuring error are provided for all bones in Table 1. Measuring error was larger for metacarpal and metatarsal bones due to the limited scanning resolution, but still low relative to variances and thus considered to be negligible in further analyses. We found no significant differences between right and left sides across all samples (Supplementary Table 2), hence we averaged these measures as well as their measuring errors.

Table 1. Bone length measurements.

Trait	Humerus	Ulna	Femur	Tibia	Metacarpal	Metatarsal
Mean length (mm)	11.11	13.57	14.51	16.57	3.22	7.35
Standard deviation (mm)	0.46	0.44	0.60	0.62	0.12	0.29
% measuring error*	0.25	0.20	0.18	0.12	5.69	0.53

* the percent of measurement error is independent of the units of the measured objects

Across the hybrids, lengths of forelimbs and hind limbs were significantly correlated with the proportion of genomic ancestry from each subspecies (Figure 2). A similar pattern was observed for the transition of skull shapes across the hybrid zone, where we have argued that this is compatible with a polygenic model (Pallares et al. 2016).

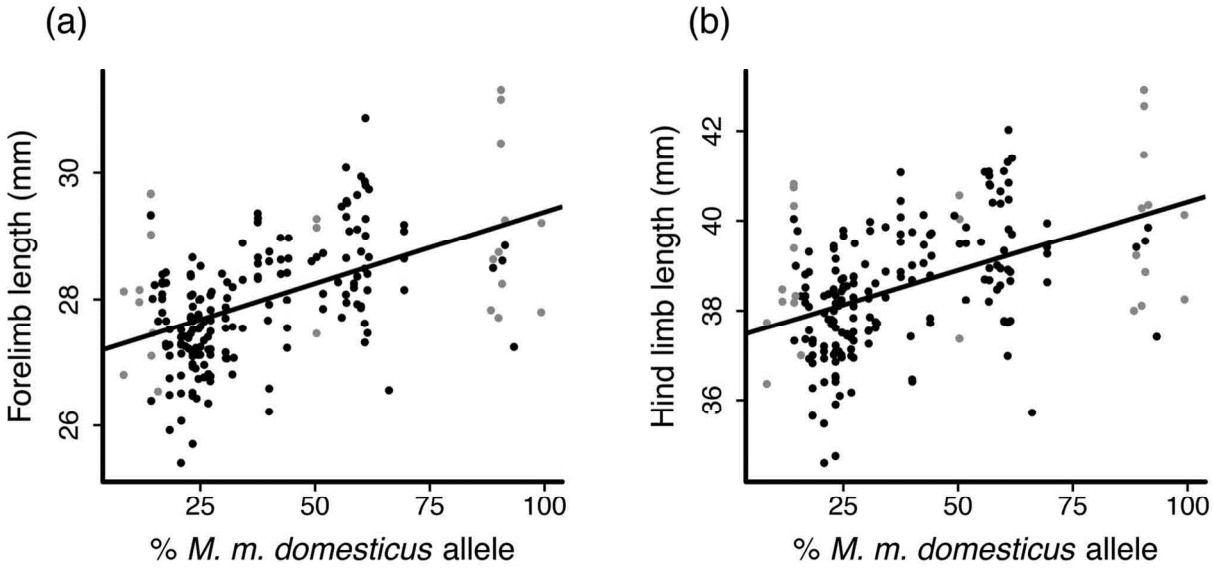


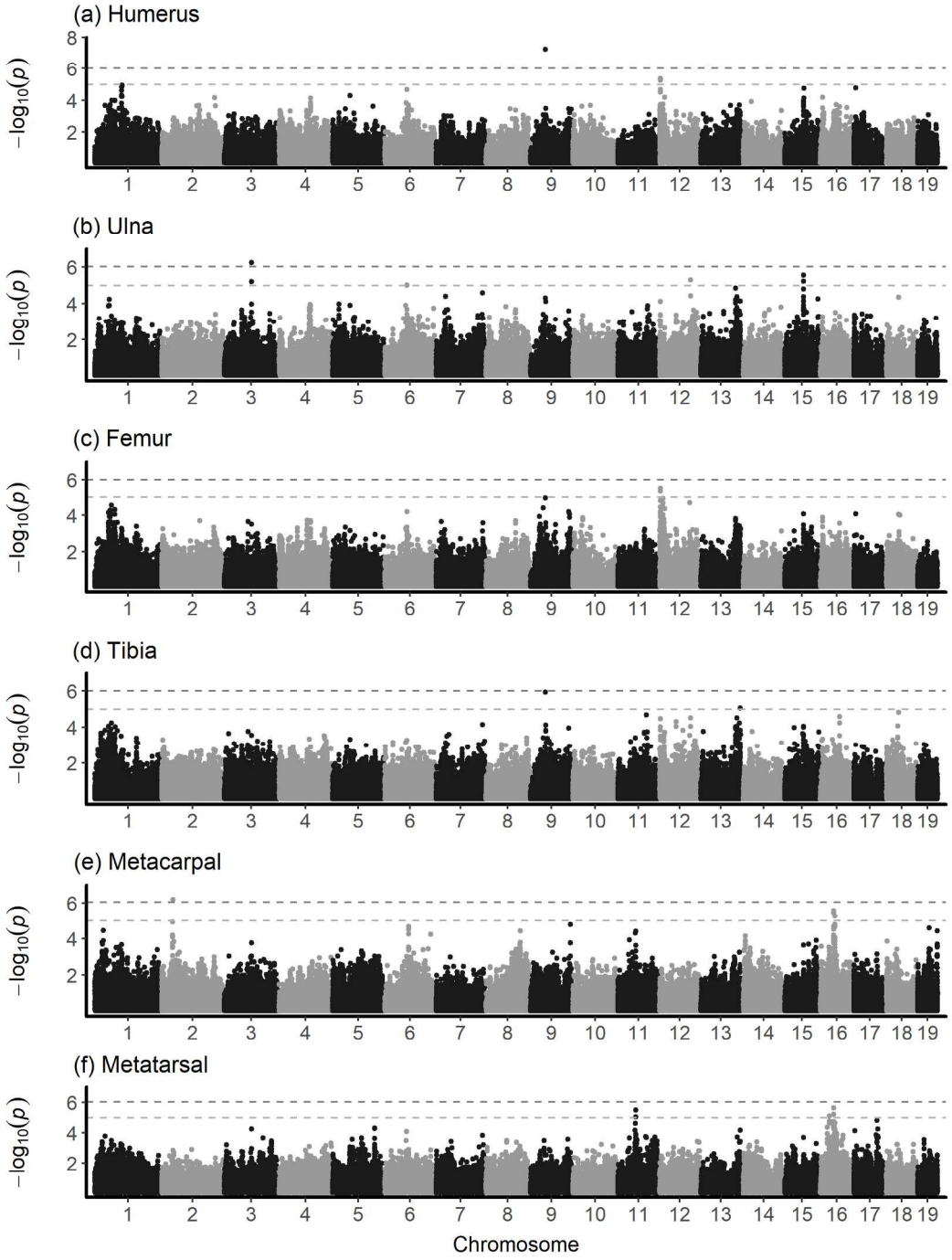
Figure 2. Correlation between limb length and genomic ancestry. Regression of forelimb (a) and hind limb (b) over the percentage of *M. m. domesticus* alleles. Forelimb values consist of summed values of humerus, ulna and metacarpal bone ($r^2 = 0.25$, $p = 8.8 \times 10^{-14}$), hind limb values are shown as a sum of femur, tibia and metatarsal bone ($r^2 = 0.22$, $p = 5.1 \times 10^{-12}$). Grey dots include values of 25 additional animals taken from Pallares et al. (2016) for which only partial genotype data were available, i.e. these were not included in the association analyses.

We observed phenotypic correlations between stylopod (humerus - femur), zeugopod (ulna - tibia) and autopod (metacarpal - metatarsal) elements (Table 2, upper diagonal). As expected (Schmidt and Fischer 2009; Pavlicev et al. 2013; Martin-Serra et al. 2015), stylopod and zeugopod elements revealed higher correlations in comparison to autopod elements. Similar relationships were observed for the genetic correlations (Table 2, lower diagonal), which are based on the genome-wide SNP-to-phenotype similarity, as estimated by the bivariate restricted maximum likelihood analysis implemented in GCTA (see Methods).

Genome wide association mapping and genetic architecture

Genome-wide association mapping included 143,592 SNPs and each bone was analyzed separately. We ran linear mixed models with and without covariates to identify SNPs significantly associated with phenotypes, and used Bayesian sparse linear mixed models to make inference about the overall genetic architecture of the traits.

The results from mapping individual bones are shown as Manhattan plots in Figure 3 (Q-Q-plots are provided in Supplementary Figure 3). When using genome-wide permutation-based thresholds, we find three significant SNPs associated with three different bones (chr2 in metacarpal bone, chr3 in ulna and chr9 in humerus). We also identified SNP associations using a more permissive significance threshold of $p < 10^{-5}$. At this level, we identified 23 significant SNPs within 13 separate "LD regions" (see Methods), distributed across 9 chromosomes and six bones, with an overlapping region on chr9 for humerus and tibia, and on chr16 for metacarpal and metatarsal bones (Supplementary Table 5). The size of LD regions around significant SNPs varied from 51 bp to 601 kb (median 114 kb).



285 **Figure 3. Manhattan plots showing SNP associations with lengths of six limb bones**. Dashed lines indicate significance thresholds with $p < 10^{-5}$ (lower line) and genome-wide threshold (GWT) based on 10,000 permutations (upper line) in (a) humerus (GWT = 9.87×10^{-7}), (b) ulna (GWT = 9.45×10^{-7}), (c) femur (GWT = 9.87×10^{-7}), (d) tibia (GWT = 9.65×10^{-7}), (e) metacarpal (GWT = 8.96×10^{-7}), (f) metatarsal bone (GWT = 9.47×10^{-7}).

Bayesian sparse linear mixed model analyses were used to calculate parameters relevant to the genetic architecture of the bone phenotypes. The proportion of phenotypic variance (PVE) explained by all SNPs in the dataset (i.e. "SNP heritability" - Wray et al. 2013) are reported in Table 2.

Table 2. Phenotypic and genetic correlations

Trait	Humerus	Ulna	Femur	Tibia	Metacarpal	Metatarsal
Humerus	**0.77**[3]	0.86[1]	0.88[1]	0.88[1]	0.40[1]	0.55[1]
Ulna	0.91[2]	**0.82**[3]	0.84[1]	0.91[1]	0.47[1]	0.64[1]
Femur	0.96[2]	0.92[2]	**0.85**[3]	0.89[1]	0.34[1]	0.51[1]
Tibia	0.88[2]	0.91[2]	0.96[2]	**0.85**[3]	0.43[1]	0.62[1]
Metacarpal	0.51[2]	0.48[2]	0.50[2*]	0.54[2]	**0.68**[3]	0.74[1]
Metatarsal	0.54[2]	0.62[2]	0.54[2]	0.60[2]	0.77[2]	**0.67**[3]

[1]Phenotypic correlation coefficients (off-diagonal, upper triangle, dark gray) in 172 genotyped individuals (Pearson product-moment correlation, all t-tests were significant at $p < 0.001$).
[2]Genetic correlations from bivariate analysis, based on genome-wide SNP similarity with phenotypic similarity (off-diagonal, lower triangle, white). In all but one case the estimated genetic correlation coefficients are significantly different from zero and from one (likelihood ratio test, $p < 0.05$, one-sided test). *For femur-metacarpal, the correlation was significantly different from one, but not from zero.
[3]Proportion of phenotypic variation explained (PVE) by all the SNPs used in the mapping from Bayesian sparse linear mixed model (diagonal – bold, light gray).

We observed high values for all bones, indicating limb bone lengths are highly heritable, at least under the laboratory breeding conditions that served to reduce environmental variation. In contrast, lower values (0.28 – 0.47) were found for the proportion of genetic variance explained by few loci with large effects, so-called "sparse effects" (PGE in GEMMA) (Supplementary Table 6), suggesting that a larger proportion of the variation is explained by many loci with small effect sizes. Further, effect size for the SNP with the lowest p-value from each significant region and posterior mean estimates for the effect size parameters from Bayesian sparse linear mixed model are reported in Supplementary Text 1 and Supplementary Table 7.

For highly polygenic traits, we expect that the proportion of variance in bone length explained by each chromosome should be correlated with chromosome length (Yang et al. 2011a; Berenos et al. 2015). In our previous study of skull shape across the hybrid zone, we indeed found such a correlation (Pallares et al. 2014).

Hence, we conducted this analysis for each individual limb bone. We found significant positive correlations for humerus and femur, but not for the other bones (Figure 4 and Supplementary Figure 4).

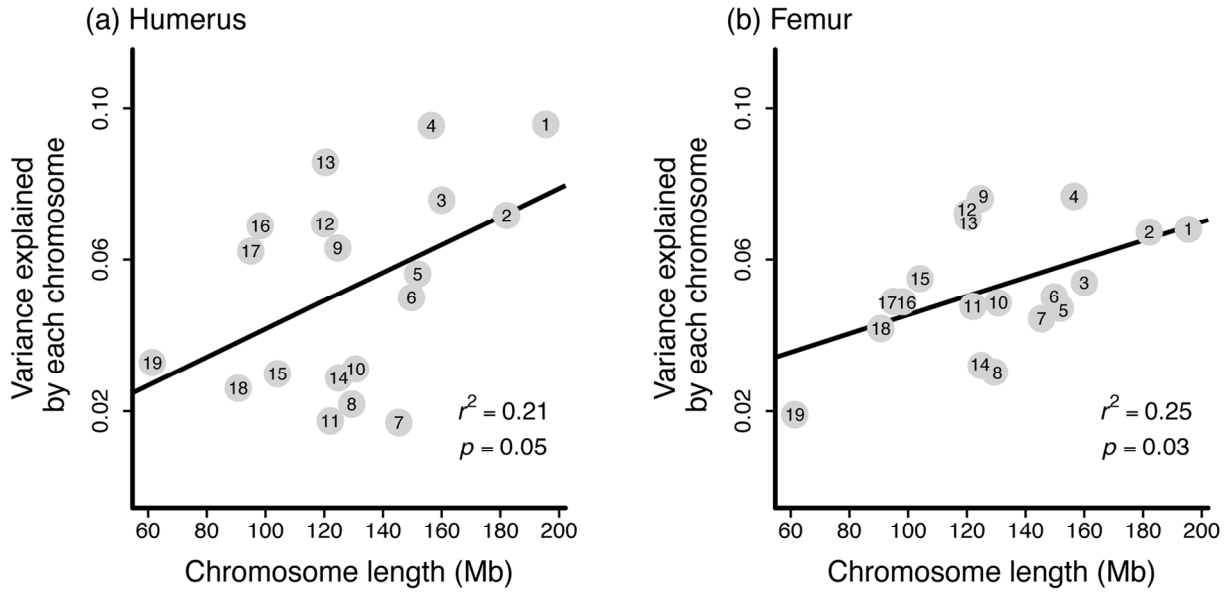


Figure 4. Correlation between variance explained and chromosome length. Relationship between phenotypic variance explained by each chromosome and chromosome length for (a) humerus, (b) femur bone.

Power to detect associations in the hybrid zone samples from simulation of additive model

Since our sample size is limited, we used simulations to assess the power of our dataset to detect significant SNPs at our chosen significance threshold. As detailed in Supplementary Text 2 (plus Supplementary Table 8), our power to detect at least one causal SNP as significant out of 10 (or 15 respectively) simulated is close to 100%. However, in none of the 1,000 simulated datasets did we identify all causal SNPs. In the majority of simulations two to three (i.e. 13-30%) causal SNPs were identified (i.e. significant SNP within 10 Mb of 'causal'). This suggests that the results reported above represent only a subset of all causal variation present in the hybrid zone, but, at the same time, confirm it is possible to detect significant associations in this population.

Candidate genes

To identify candidate genes influencing limb bone length in mice, we evaluated all annotated genes (coding and non-coding) overlapping the 13 LD regions containing SNPS with $p < 10^{-5}$. In Table 3, we highlight genes with known roles in any aspect of bone formation. Several genes were previously associated with limb phenotypes in other studies (see discussion for further details). It should be noted that, because we do not have sufficient mapping resolution to identify causative mutations, we cannot distinguish between regulatory or coding effects. However, three regions do not overlap any genes, suggesting the underlying mutation is a regulatory polymorphism affecting a nearby or distant gene. Hence, evaluation of candidate genes should be viewed as suggestive of potential underlying genetic causes.

Table 3. Genomic regions associated with limb bone length, identified by individual bone analysis

Bone	Region	N of SNPs[1]	p-value of the best SNP[2]	Size (Mb)[3]	Genes in the region[4]
Metacarpal	chr2: 35028134-35029436	1 (2)	**6.46×10^{-7}**	0.001 [0.114]	***Hc***, *AI182371*
Ulna	chr3: 82469792-82583792	2[5]	**5.60×10^{-7}**	[0.114]	*Npy2r*
Ulna	chr6: 67069424-67183424	1	9.56×10^{-6}	[0.114]	*E230016M11Rik, AK079709, AK039826*
Humerus	chr9: 46604715-46718715	1	**5.74×10^{-8}**	[0.114]	*Gm22805*
Tibia		1	1.17×10^{-6}		
Metatarsal	chr11: 55936031-55964675	2 (3)	3.29×10^{-6}	0.029 [0.114]	*Gm12239*
Femur	chr12: 5161413-5762255	2 (12)	2.88×10^{-6}	0.601	*Klhl29, 2810032G03Rik*
Humerus	chr12: 5792918-5907264	2 (12)	4.31×10^{-6}	0.114	*AK135963*
Ulna	chr12: 93777662-93851912	1 (3)	5.09×10^{-6}	0.074 [0.114]	/
Tibia	chr13: 117841280-117955280	1	8.69×10^{-6}	[0.114]	*Hcn1*
Ulna	chr15: 58538615-58676719	2 (9)	2.83×10^{-6}	0.138	*Fer1l6*
Metatarsal	chr16: 30263056-30263107	1 (2)	7.74×10^{-6}	5.10×10^{-5} [0.114]	*Cpn2,* ***Lrrc15,*** *Gp5, Atp13a3*
Metacarpal	chr16: 41712286-42136622	3 (33)	2.62×10^{-6}	0.424	***Lsamp***
Metatarsal		2 (33)	2.27×10^{-6}		
Metacarpal	chr16: 45984088-46202709	1 (16)	5.68×10^{-6}	0.219	*Cd96, Gm4737,* ***Plcxd2***[6]

[1]Number of SNPs in the region with $p < 10^{-5}$, in parentheses number of total SNPs that are in LD within the same region.
[2]p-values in bold represent the SNPs above the permutation-based threshold.
[3]Size of the LD region is provided. In cases where this was smaller than the median size of regions (0.114 Mb), we used the latter to search for annotated genes, indicated as [0.114] in the respective fields.
[4]Annotated genes overlapping with the region; genes involved in limb development, phenotype or expression are highlighted in bold.
[5]two SNPs within 6kb, but not in LD.

[6]Overlap with previous QTL study on limb length in Pavlicev et al. (2013).

355

Identifying genetic associations unique to individual bones by accounting for covariation with another bone

To identify possible rQTLs, we performed the bone length mapping with a second bone included as a covariate (see Methods for specific combinations of bones). This revealed 17 SNPs above the genome-wide permutation thresholds for the following bone combinations (the first capital letter refers to the bone tested, while the second 360 capital letter for the one used as a covariate): HcovU, FcovT, HcovF, UcovT, UcovF and FcovU (Table 4 and Supplementary Table 9). Manhattan plots are provided in Supplementary Figure 6 and the corresponding Q-Q-plots are provided in Supplementary Figure 7. None of the SNPs significant at genome-wide level and $p < 10^{-5}$ when mapping individual bone lengths were also significant in mapping analyses with a second bone as covariate. With significance threshold of $p < 10^{-5}$, we identified associations clustered in 39 LD regions with a median size of 118 kb 365 (min = 113bp, max = 1.4 Mb) for 12 bone combinations. Overlapping regions were found in the following combinations with a second bone as covariate: chr 1 (HcovF, HcovT, TcovH), chr6 (HcovF, FcovH), chr16 (HcovU, HcovF, HcovT) and chr19 (HcovU, UcovH) (Table 4 and Supplementary Table 9).

We are particularly interested in candidate regions identified by mapping with a covariate, which are specific to only one bone from a developmentally linked pair. These loci may be involved in breaking constraints from 370 correlated development, and thus would enable independent evolution of length for a single limb bone. We have therefore inspected these regions for candidate genes, using the same approach as for the individual bone analysis. The results are shown in Table 4 (see discussion for details).

Table 4. Genomic regions associated with limb bone length, identified by mapping with covariates

375

Bone[1]	Region	N of SNPs[2]	p-value of the best SNP[3]	Size (Mb)[4]	Genes in the region[5]
UcovF	chr1: 3625815-4718067	2 (87)	**5.63×10^{-7}**	1.092	***Sox17,*** *Rp1, Xkr4, AK149000*
FcovU	chr1: 9262814-9380814	1	5.92×10^{-6}	[0.118]	*Sntg1*
HcovF	chr1: 23690398-23716894	1 (3)	1.39×10^{-6}	0.026 [0.118]	/

FcovT	chr1: 84059655-84952620	2 (21)	**4.21 x 10^{-8}**	0.893	*Fbxo36, Pid1,* ***Dner*** [6]***,*** *Trip12*, Slc16a14, Mir6353, AK032919, AK036072*
FcovH	chr1: 119187233-119322362	1 (9)	8.77 x 10^{-6}	0.135	/
HcovF	chr1: 119320805-119438805	1	4.10 x 10^{-6}	[0.118]	***Inhbb,*** *AK039419*
HcovT	chr1: 128607010-128725010	1	7.48 x 10^{-6}	[0.118]	/
HcovF	chr1: 152359185-152477185	1	3.60 x 10^{-6}	[0.118]	*Tsen15, Colgalt2*
HcovF	chr1: 154018113-154036315	1 (3)	4.87 x 10^{-6}	0.018 [0.118]	*Gm29291, Gm28286, AK043564, AK154552*
HcovF	chr1: 154053011-154077631	2 (2)	5.20 x 10^{-6}	0.025 [0.118]	*AK043564, AK154552*
HcovT		2 (2)	3.29 x 10^{-6}		
TcovH		2 (2)	9.85 x 10^{-6}		
FcovH	chr2: 8781928-9002418	2 (10)	3.87 x 10^{-6}	0.22	/
MCcovMT	chr2: 33909493-34027493	1	5.15 x 10^{-6}	[0.118]	*AK162388, C230014O12Rik*
MCcovMT	chr2: 54244165-54348111	1 (3)	6.20 x 10^{-6}	0.104	/
TcovF	chr3: 100168735-100360505	1 (9)	6.63 x 10^{-6}	0.192	*Gdap2*
HcovT	chr3: 130829175-130833777	1 (2)	6.90 x 10^{-6}	0.005 [0.118]	/
HcovF	chr3: 154809225-154926757	1 (3)	8.46 x 10^{-6}	0.118	*Tnni3k*
UcovF	chr5: 8672123-8674096	1 (2)	2.48 x 10^{-6}	0.002 [0.118]	*Abcb1a, Rundc3b*
FcovT	chr6: 44025034-44505141	1 (3)	**4.17 x 10^{-8}**	0.48	/
FcovH	chr6: 102314940-102422804	1 (3)	7.91 x 10^{-6}	0.108	*4930587E11Rik,* ***Cntn3***
FcovT	chr6: 134066708-135439926	6 (37)	**6.77 x 10^{-9}**	1.373	*2810454H06Rik, Apold1, Gprc5d, Gm19434, Gsg1, Pbp2, Bcl2l14, Borcs5, Crebl2,* ***Cdkn1b,*** *Lockd, Ddx47, Gprc5a, Hebp1, Fam234b, Gsg1, Emp1,* ***Lrp6, Mansc1, Dusp16, Gpr19, Etv6***
FcovT	chr6: 139526562-139581085	3 (8)	**2.77 x 10^{-7}**	0.055 [0.118]	*Rergl, Pik3c2g*
HcovF	chr6: 144869562-144987562	1	1.03 x 10^{-6}	[0.118]	/
FcovH		1	8.08 x 10^{-6}		
TcovF	chr6: 146155662-146273662	1	7.51 x 10^{-6}	[0.118]	*Itpr2* [6]***
HcovF	chr6: 146457299-146575299	1	1.94 x 10^{-6}	[0.118]	*Ints13 (Asun), Itpr2* [6]***
TcovH	chr7: 34695176-34813176	1	4.19 x 10^{-6}	[0.118]	*Chst8*
FcovU	chr7: 43234461-43352461	1	**5.22 x 10^{-7}**	[0.118]	*Zfp715, Siglecf*
FcovT	chr8: 45416400-45534400	1	5.67 x 10^{-6}	[0.118]	***Sorbs2*** [6]********
TcovF	chr12: 67683133-68174899	1 (27)	3.11 x 10^{-6}	0.492	/
FcovH	chr12: 86591551-86709551	1	5.48 x 10^{-6}	[0.118]	*Vash1, Angel1*
TcovH	chr15: 44447489-44565489	1	3.51 x 10^{-6}	[0.118]	*Pkhd1l1*
FcovU	chr15: 59470046-59482266	1 (3)	**7.14 x 10^{-7}**	0.012 [0.118]	*Nsmce2, AK080559*
FcovU	chr15: 72720799-72733036	1 (2)	6.61 x 10^{-6}	0.012 [0.118]	***Trappc9***
UcovT	chr15: 75470303-76162911	2 (37)	3.29 x 10^{-6}	0.693	*Ly6h, Gpihbp1, Zfp41, Mafa, Gsdmd, Mroh6, Naprt, Tigd5, Pycrl, Tsta3, Zfp623, Ccdc166,* ***Mapk15,*** *Mir6952, Top1mt,* ***Rhpn1,*** *Eef1d, Zfp707, Fam83h,* ***Scrib*** [6]******,*** *Puf60, Nrbp2, Eppk1, BC024139, Zc3h3*
UcovT	chr15: 76148236-76266236	1	**8.84 x 10^{-8}**	[0.118]	*Mir1942, Grina, Mir6953, Plec, Parp10*
HcovF	chr16: 92988497-93106497	1	3.55 x 10^{-6}	[0.118]	/
HcovU	chr16: 93764330-94205101	4 (5)	**8.30 x 10^{-7}**	0.441	***Chaf1b, Cldn14, Dopey2, Morc3, Sim2, Hlcs,*** *AK009785*
HcovF		2 (5)	8.98 x 10^{-7}		
HcovT		3 (5)	1.72 x 10^{-6}		
HcovF	chr17: 30935073-31053073	1	**3.49 x 10^{-7}**	[0.118]	***Umodl1,*** *Glp1r, AK138161*

HcovF	chr17: 31860580-32474832	6 (17)	5.47×10^{-6}	0.614	***Rrp1b,*** *Ephx3, Gm4432, A530088E08Rik, Pglyrp2, Pdxk-ps,* ***Notch3***[6]******, Akap8,*** *Akap8l, Rasal3, Cyp4f39,* ***Hsf2bp,*** *Brd4, Wiz*
HcovU	chr19: 56931261-56931374	1 (2)	3.45×10^{-6}	1.13×10^{-4}	***Afap1l2, Vwa2***
UcovH		1 (2)	5.90×10^{-6}		

[1]Annotation: first capital letter stands for the bone tested, second capital letter for the one used as covariate.
[2]Number of SNPs in the region with $p < 10^{-5}$, in parentheses number of total SNPs that are in LD within the same region.
[3]p-values in bold represent the SNPs above the permutation-based threshold.
[4]Size of the LD region is provided. In cases where this was smaller than the median size of regions (0.118 Mb), we used the latter to search for annotated genes, indicated as [0.118] in the respective fields
[5]Annotated genes overlapping with the region; genes involved in limb development, phenotype or expression are highlighted in bold.
[6]Overlap with previous QTL studies in limb length: * Norgard et al. 2011; ** Pavlicev et al. 2013; *** Kenney-Hunt et al. 2006.

DISCUSSION

We have explored the use of animals from a natural hybrid zone for genome-wide association mapping of length differences in limb bones. Limb length variation is a model for understanding the evolution of correlated traits; therefore it is of particular interest to map not only loci that affect individual bone length, but also loci that are involved in influencing correlated variation. We were able to identify associated genomic regions for both of these components, despite our moderate sample size and the highly polygenic nature of these traits.

We report candidate regions significant at $p < 10^{-5}$, a more permissive level than the significance threshold obtained by permutations tests. For highly polygenic traits, permutation tests are overly conservative, especially if sample sizes are relatively small (Yang et al. 2010, 2011a). We note that significant SNP associations from mapping individual bones disappeared when including a correlated bone phenotype as a covariate. This pattern is not expected for spurious associations, and thus provides support for $p < 10^{-5}$ as an appropriate threshold.

In the single bone analyses we detected two loci that affected more than one bone, one on chr9 (humerus and tibia) and one on chr16 (metacarpal and metatarsal bones) (Table 3), implying these loci are involved in influencing covariation. Further such overlaps are detected under the relaxed threshold of $p < 10^{-4}$ (Supplementary Table 5), also indicating correlated responses across traits. Interestingly, the chr9 locus shows an additional association with ulna and femur at this lower p-value cutoff.

The phenotypic correlations among limb bones we find in mice from the hybrid zone are consistent with a hierarchical covariance structure, with higher correlations among homologous elements between limbs (stylopod and zeugopod) than correlations within limbs, and lower correlations between stylopod / zeugopod elements and autopod elements (Hallgrimsson et al. 2002; Young and Hallgrimsson 2005). Note that the bones of the autopod are shorter and have higher measurement error, which might reduce correlation estimates involving these bones. However, it is generally known that autopod elements can show larger morphological variation relative to proximal limb elements (Shubin et al. 1997; Capdevila and Belmonte 2000), apparently due to their interaction with the substrate and hence more room for plastic responses (see also below).

Hybrids show limb bone lengths intermediate between pure subspecies (Figure 2), in contrast to transgressive fertility phenotypes observed in some individuals from this mapping population (e.g. some hybrids have lower testis weight than either pure subspecies, Turner and Harr 2014), putatively caused by epistatic hybrid incompatibilities. Lack of transgressive bone length phenotypes suggests the developmental pathways leading to limb formation are not majorly affected by hybrid incompatibilities.

Our results are compatible with shared developmental processes determining lengths of correlated limb bones; we find high correlations among bone length phenotypes and identify candidate regions associated with multiple limb bones. However, these processes are not determined by a few controlling loci of large effect. Chromosomal partitioning of variance (Figure 4) and higher values of PVE (SNP heritability) vs. PGE (sparse effects) support a highly polygenic model for limb bone length.

Candidate genes

The molecular mechanisms of limb formation have been intensively studied in developmental biology, including identification and characterization of many underlying causative genes and interactions. The major developmental genes that have been implicated in regulating embryonic limb development include Hox genes, *Tbx3, Tbx4, Tbx5, Fgf4, Fgf8, Fgf10, Shh, Pitx1, Wnt2b, Meis1, Meis2, Wnt3a, Aldh1a2, Bmp2, Bmp4, Bmp7, Gli3, Hand2, Cyp26b1,*

Grem1, Grem2 (Logan 2003; Tickle 2006; Sheeba et al. 2016). Interestingly, none of these genes overlap with any candidate regions identified here (>1Mb away). These results suggest that the developmental genes that initiate embryonic limb growth and determine their identity may have little influence on natural variation in bone length.

In contrast to the lack of associations with previously described control genes, some candidate regions identified here overlap with previously reported QTL regions for limb length and proposed candidate genes within them. These include *Scrib* and *Notch3* (Kenney-Hunt et al. 2006), the region including *Dner*, *Trip12* and *Itpr2* (Norgard et al. 2011), as well as *Plcxd2* and *Sorbs2* (Pavlicev et al. 2013). However, not much is known about the roles these genes play in bone development, and only *Plcxd2* is known to be expressed in developing limbs. Interestingly, Pavlicev et al. (2013) found this gene as a rQTL candidate for forelimb / hind limb differences in mice, while we found it in the individual bone mapping of metacarpal length, a bone that was not included in their analysis.

Another gene implicated in metacarpal length is the hemolytic complement gene *Hc*. Although mostly known from the blood clotting cascade, complement genes were also found to be involved in bone homeostasis (Schoengraf et al. 2013). Complement proteins are known to be present in the zones of endochondral bone development with *Hc* being expressed in the hypertrophic zone of foetal tibiae and femurs (Andrades et al. 1996). Several SNPs associated with the autopodal bone length identify the gene *Lsamp* on chromosome 16. While *Lsamp* is mostly expressed in the cortical and subcortical regions of the neural limbic system (Zacco et al. 1990), it was also identified as a causative gene for osteosarcomas (Baroy et al. 2014). The Y2 receptor gene *Npy2r*, on chromosome 3, is associated with ulna length. Y2 receptor signaling is known to be important in neuropeptide Y (NPY)-mediated effects on energy homeostasis and bone physiology and *Npy2r* was found to regulate trabecular bone homeostasis (Shi et al. 2010).

The other genes identified in the individual bone analyses do not have annotated functions related to bone growth or homeostasis. However, given the double functions of such genes as *Hc* and *Lsamp*, it could be fruitful to investigate possible bone effects for these other candidate genes. Of particular interest are regions that code for recently evolved genes, which are generally thought to play roles in lineage specific adaptations (Tautz and Domazet-Loso 2011; Schlötterer 2015). We find quite a number of such loci within the mapped intervals, which are

so far only annotated as transcripts of unknown function (listed as Gm.. or AK.. numbers, or numbers ending with ...Rik in Table 3 and 4). On the other hand, some of these transcripts may be associated with a molecular function. For example, the region in chr9 that shows associations with more than one bone (see Table 3 - and discussion above) codes for a small nuclear RNA (snRNA - *Gm22805*). Such RNAs can be involved in regulating multiple other RNAs. Most interestingly, we find that the transcript is partially deleted in *M. m. musculus* populations (can be seen in the genome sequence data provided in Harr et al. (2016)), i.e. this could potentially constitute the causative *M. m. domesticus* / *M. m. musculus* polymorphism that is detected in the mapping.

Mapping analyses including a second bone as covariate also yielded a list of candidate genes, several of which have known involvement in bone growth. A region on chromosome 6, associated with femur length, contains two such genes; *Cdknb1* is broadly expressed and has a role in chondrocyte proliferation, among other functions (Cardelli et al. 2013). *Lrp6* is, together with *Lrp5*, an essential co-receptor of *wnt* signaling with a direct involvement in osteoblastogenesis and cartilage development (Joeng et al. 2011). Humerus length is associated with a region including *Morc3*, which is involved in calcium homeostasis and osteoblast differentiation (Jadhav et a. 2016).

Bone morphology is not only influenced by genetic and developmental processes, but can also be highly plastic due to changes in mechanical load. This can lead to an interdependence with other skeletal elements and the musculature (Tsutsumi et al. 2017). The extent to which plasticity is also under genetic constraints remains an open question, but it seems plausible that genes without direct roles in bone development may influence bone structures indirectly by moderating plasticity. On the other hand, Young (2013) has argued that macroevolutionary diversity of limbs may be dependent on constraints provided by the early developmental program. But subspecies may be too closely related to show such constraints, hence we would not have power to detect them here. Major changes in limb proportions, as well as digit losses, were described for the rodent superfamily Dipodoidea (including the bipedal Jerboas) (Moore et al. 2015), indicating that even macroevolutionary divergence can occur rapidly, provided the ecological conditions allow this. It has been suggested that both early and late developmental processes can contribute to digit losses in mammals (Cooper et al. 2014).

Our candidate gene list overlaps only partially with genes found in previous mapping efforts for the same phenotypes (e.g. Wolf et al, 2006; Norgard et al. 2008, 2009; Pavlicev et al. 2008; Pavlicev et al. 2013; Parmenter et al. 2016). However, these previous efforts were all based on *M. m. domesticus* mapping populations. Because our mapping population contains variation from two different subspecies of house mice, we may expect to find additional loci. We note that novel loci were also identified by mapping of skull shape phenotypes in this same hybrid mapping population (Pallares et al. 2014) relative to a *M. m. domesticus* panel (Pallares et al. 2015).

Conclusion

Revealing the genetic architecture of quantitative phenotypes of wild animals remains very challenging. Association mapping in natural populations has great promise to achieve this goal. Our results show that hybrid zones can be used to identify a subset of candidate loci for polygenic traits at relatively high resolution, even with a modest number of individuals.

Acknowledgements

We thank B. Poerschke and E. Blohm-Sievers for scanning the mice and A. Schunke for advice on scanning and 3D landmark digitalization. We thank X. Zhou for help with GEMMA, R. Bakarić for bioinformatics help, K. Delmore, K. Ullrich and G. Reeves for discussions on mapping. The study was funded by institutional support by the MPG to DT.

Data accessibility

Files for raw limb measures, age corrected measures, the kinship matrix and mapping files are deposited in Dryad under DOI: 10.5061/dryad.rg6k9

Files on Dryad include the PED file containing genotype data for 185 individuals (column 7 onwards) and MAP file with the original set of SNPs prior to LD-pruning (285,625 SNPs); binary ped file (BED) with genotypes of 172 individuals, extended MAP file (BIM), data after LD-pruning (156,183 SNPs with X chromosome) and phenotype information (FAM).

References

Andrades JA, Nimni ME, Becerra J, *et al.* (1996) Complement proteins are present in developing endochondral bone and may mediate cartilage cell death and vascularization. *Experimental Cell Research* **227**, 208-213.

Baroy T, Kresse SH, Skarn M, *et al.* (2014) Reexpression of LSAMP inhibits tumor growth in a preclinical osteosarcoma model. *Molecular Cancer* **13**.

Berenos C, Ellis PA, Pilkington JG, *et al.* (2015) Heterogeneity of genetic architecture of body size traits in a free-living population. *Molecular Ecology* **24**, 1810-1830.

Blake JA, Eppig JT, Kadin JA, *et al.* (2017) Mouse Genome Database (MGD)-2017: community knowledge resource for the laboratory mouse. *Nucleic Acids Research* **45**, D723-D729.

Bush WS, Moore JH (2012) Chapter 11: Genome-Wide Association Studies. *Plos Computational Biology* **8**.

Capdevila J, Belmonte JCI (2000) Perspectives on the evolutionary origin of tetrapod limbs. *Journal of Experimental Zoology* **288**, 287-303.

Cardelli M, Zirngibl RA, Boetto JF, *et al.* (2013) Cartilage-Specific Overexpression of ERR gamma Results in Chondrodysplasia and Reduced Chondrocyte Proliferation. *Plos One* **8**.

Cheverud JM (1988) A COMPARISON OF GENETIC AND PHENOTYPIC CORRELATIONS. *Evolution* **42**, 958-968.

Cheverud JM, Ehrich TH, Vaughn TT, *et al.* (2004) Pleiotropic effects on mandibular morphology II: Differential epistasis and genetic variation in morphological integration. *Journal of Experimental Zoology Part B-Molecular and Developmental Evolution* **302B**, 424-435.

Christians JK, Bingham V, Oliver F, Heath TT, Keightley PD (2003) Characterization of a QTL affecting skeletal size in mice. *Mammalian Genome* **14**, 175-183.

Claude J (2008). Morphometrics with R. Springer Verlag

Cooper KL, Sears KE, Uygur A, *et al.* (2014) Patterning and post-patterning modes of evolutionary digit loss in mammals. *Nature* **511**, 41-U537.

Deary IJ, Yang J, Davies G, *et al.* (2012) Genetic contributions to stability and change in intelligence from childhood to old age. *Nature* **482**, 212-215.

Drake TA, Hannani K, Kabo JM, *et al.* (2001) Genetic loci influencing natural variations in femoral bone morphometry in mice. *Journal of Orthopaedic Research* **19**, 511-517.

Eppig JT, Blake JA, Bult CJ, *et al.* (2012) The Mouse Genome Database (MGD): comprehensive resource for genetics and genomics of the laboratory mouse. *Nucleic Acids Research* **40**, D881-D886.

Flint J, Eskin E (2012) Genome-wide association studies in mice. *Nature Reviews Genetics* **13**, 807-817.

Flint J, Valdar W, Shifman S, Mott R (2005) Strategies for mapping and cloning quantitative trait genes in rodents. *Nature Reviews Genetics* **6**, 271-286.

Guenet JL, Bonhomme F (2003) Wild mice: an ever-increasing contribution to a popular mammalian model. *Trends in Genetics* **19**, 24-31.

Hall BK (1995) HOMOLOGY AND EMBRYONIC-DEVELOPMENT. *Evolutionary Biology, Vol 28* **28**, 1-37.

Hallgrimsson B, Jamniczky H, Young NM*, et al.* (2009) Deciphering the Palimpsest: Studying the Relationship Between Morphological Integration and Phenotypic Covariation. *Evolutionary Biology* **36**, 355-376.

Hallgrimsson B, Willmore K, Hall BK (2002) Canalization, developmental stability, and morphological integration in primate limbs. *Yearbook of Physical Anthropology, Vol 45* **45**, 131-158.

Hansen TF (2006) The evolution of genetic architecture. *Annual Review of Ecology Evolution and Systematics* **37**, 123-157.

Hansen TF, Armbruster WS, Carlson ML, Pelabon C (2003) Evolvability and genetic constraint in Dalechampia blossoms: Genetic correlations and conditional evolvability. *Journal of Experimental Zoology Part B-Molecular and Developmental Evolution* **296B**, 23-39.

Harr B, Karakoc E, Neme R*, et al.* (2016) Genomic resources for wild populations of the house mouse, Mus musculus and its close relative Mus spretus. *Scientific Data* **3**.

Jadhav G, Teguh D, Kenny J, Tickner J, Xu JK (2016) Morc3 mutant mice exhibit reduced cortical area and thickness, accompanied by altered haematopoietic stem cells niche and bone cell differentiation. *Scientific Reports* **6**.

Joeng KS, Schumacher CA, Zylstra-Diegel CR, Long FX, Williams BO (2011) Lrp5 and Lrp6 redundantly control skeletal development in the mouse embryo. *Developmental Biology* **359**, 222-229.

Kang HM, Sul JH, Service SK*, et al.* (2010) Variance component model to account for sample structure in genome-wide association studies. *Nature Genetics* **42**, 348-U110.

Kenney-Hunt JP, Vaughn TT, Pletscher LS*, et al.* (2006) Quantitative trait loci for body size components in mice. *Mammalian Genome* **17**, 526-537.

Kent WJ, Sugnet CW, Furey TS*, et al.* (2002) The human genome browser at UCSC. *Genome Research* **12**, 996-1006.

Kolarov NT, Ivanovic A, Kalezic ML (2011) Morphological Integration and Ontogenetic Niche Shift: A Study of Crested Newt Limbs. *Journal of Experimental Zoology Part B-Molecular and Developmental Evolution* **316B**, 296-305.

Lande R (1980) THE GENETIC COVARIANCE BETWEEN CHARACTERS MAINTAINED BY PLEIOTROPIC MUTATIONS. *Genetics* **94**, 203-215.

Laurie CC, Nickerson DA, Anderson AD*, et al.* (2007) Linkage disequilibrium in wild mice. *Plos Genetics* **3**, 1487-1495.

Lee SH, Yang J, Goddard ME, Visscher PM, Wray NR (2012) Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. *Bioinformatics* **28**, 2540-2542.

Logan M (2003) Finger or toe: the molecular basis of limb identity. *Development* **130**, 6401-6410.

Marchini M, Sparrow LM, Cosman MN*, et al.* (2014) Impacts of genetic correlation on the independent evolution of body mass and skeletal size in mammals. *Bmc Evolutionary Biology* **14**.

Martin-Serra A, Figueirido B, Perez-Claros JA, Palmqvist P (2015) Patterns of morphological integration in the appendicular skeleton of mammalian carnivores. *Evolution* **69**, 321-340.

Moore TY, Organ CL, Edwards SV*, et al.* (2015) Multiple Phylogenetically Distinct Events Shaped the Evolution of Limb Skeletal Morphologies Associated with Bipedalism in the Jerboas. *Current Biology* **25**, 2785-2794.

Norgard EA, Jarvis JP, Roseman CC*, et al.* (2009) Replication of long-bone length QTL in the F-9-F-10 LG,SM advanced intercross. *Mammalian Genome* **20**, 224-235.

Norgard EA, Lawson HA, Pletscher LS*, et al.* (2011) Genetic factors and diet affect long-bone length in the F-34 LG,SM advanced intercross. *Mammalian Genome* **22**, 178-196.

Norgard EA, Roseman CC, Fawcett GL*, et al.* (2008) Identification of quantitative trait loci affecting murine long bone length in a two-generation intercross of LG/J and SM/J mice. *Journal of Bone and Mineral Research* **23**, 887-895.

Pallares LF, Carbonetto P, Gopalakrishnan S*, et al.* (2015) Mapping of Craniofacial Traits in Outbred Mice Identifies Major Developmental Genes Involved in Shape Determination. *Plos Genetics* **11**.

Pallares LF, Harr B, Turner LM, Tautz D (2014) Use of a natural hybrid zone for genomewide association mapping of craniofacial traits in the house mouse. *Molecular Ecology* **23**, 5756-5770.

Pallares LF, Turner LM, Tautz D (2016) Craniofacial shape transition across the house mouse hybrid zone: implications for the genetic architecture and evolution of between-species differences. *Development Genes and Evolution* **226**, 173-186.

Parmenter MD, Gray MM, Hogan CA*, et al.* (2016) Genetics of Skeletal Evolution in Unusually Large Mice from Gough Island. *Genetics* **204**, 1559-+.

Pavlicev M, Kenney-Hunt JP, Norgard EA*, et al.* (2008) Genetic variation in pleiotropy: Differential epistasis as a source of variation in the allometric relationship between long bone lengths and body weight. *Evolution* **62**, 199-213.

Pavlicev M, Wagner GP (2012) A model of developmental evolution: selection, pleiotropy and compensation. *Trends in Ecology & Evolution* **27**, 316-322.

Pavlicev M, Wagner GP, Noonan JP, Hallgrimsson B, Cheverud JM (2013) Genomic Correlates of Relationship QTL Involved in Fore- versus Hind Limb Divergence in Mice. *Genome Biology and Evolution* **5**, 1926-1936.

Phifer-Rixey M, Nachman MW (2015) Insights into mammalian biology from the wild house mouse Mus musculus. *eLife* **4**.

Price AL, Zaitlen NA, Reich D, Patterson N (2010) New approaches to population stratification in genome-wide association studies. *Nature Reviews Genetics* **11**, 459-463.

Purcell S, Neale B, Todd-Brown K*, et al.* (2007) PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics* **81**, 559-575.

Rieseberg LH, Buerkle CA (2002) Genetic mapping in hybrid zones. *American Naturalist* **159**, S36-S50.

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rolian C, Lieberman DE, Hallgrimsson B (2010) THE COEVOLUTION OF HUMAN HANDS AND FEET. *Evolution* **64**, 1558-1568.

Schlötterer C (2015) Genes from scratch - the evolutionary fate of de novo genes. *Trends in Genetics* **31**, 215-219.

Schmidt M, Fischer MS (2009) MORPHOLOGICAL INTEGRATION IN MAMMALIAN LIMB PROPORTIONS: DISSOCIATION BETWEEN FUNCTION AND DEVELOPMENT. *Evolution* **63**, 749-766.

Schoengraf P, Lambris JD, Recknagel S*, et al.* (2013) Does complement play a role in bone development and regeneration? *Immunobiology* **218**, 1-9.

Schunke AC, Bromiley PA, Tautz D, Thacker NA (2012) TINA manual landmarking tool: software for the precise digitization of 3D landmarks. *Frontiers in Zoology* **9**.

Sheeba CJ, Andrade RP, Palmeirim I (2016) Getting a handle on embryo limb development: Molecular interactions driving limb outgrowth and patterning. *Seminars in Cell & Developmental Biology* **49**, 92-101.

Shi YC, Lin S, Wong IPL*, et al.* (2010) NPY Neuron-Specific Y2 Receptors Regulate Adipose Tissue and Trabecular Bone but Not Cortical Bone Homeostasis in Mice. *Plos One* **5**.

Shubin N, Tabin C, Carroll S (1997) Fossils, genes and the evolution of animal limbs. *Nature* **388**, 639-648.

Sul JH, Bilow M, Yang WY*, et al.* (2016) Accounting for Population Structure in Gene-by-Environment Interactions in Genome-Wide Association Studies Using Mixed Models. *Plos Genetics* **12**.

Tautz D, Domazet-Loso T (2011) The evolutionary origin of orphan genes. *Nature Reviews Genetics* **12**, 692-702.

Tickle C (2006) Making digit patterns in the vertebrate limb. *Nature Reviews Molecular Cell Biology* **7**, 45-53.

Tsutsumi R, Tran MP, Cooper KL (2017) Changing While Staying the Same: Preservation of Structural Continuity During Limb Evolution by Developmental Integration. *Integrative and Comparative Biology* **57**, 1269-1280.

Turner LM, Harr B (2014) Genome-wide mapping in a house mouse hybrid zone reveals hybrid sterility loci and Dobzhansky-Muller interactions. *Elife* **3**.

Turner LM, Schwahn DJ, Harr B (2012) REDUCED MALE FERTILITY IS COMMON BUT HIGHLY VARIABLE IN FORM AND SEVERITY IN A NATURAL HOUSE MOUSE HYBRID ZONE. *Evolution* **66**, 443-458.

Visscher PM, Hemani G, Vinkhuyzen AAE*, et al.* (2014) Statistical Power to Detect Genetic (Co)Variance of Complex Traits Using SNP Data in Unrelated Samples. *Plos Genetics* **10**.

Wheeler HE, Shah KP, Brenner J*, et al.* (2016) Survey of the Heritability and Sparse Architecture of Gene Expression Traits across Human Tissues. *Plos Genetics* **12**.

Wolf JB, Pomp D, Eisen EJ, Cheverud JM, Leamy LJ (2006) The contribution of epistatic pleiotropy to the genetic architecture of covariation among polygenic traits in mice. *Evolution & Development* **8**, 468-476.

Wray NR, Yang J, Hayes BJ*, et al.* (2013) Pitfalls of predicting complex traits from SNPs. *Nature Reviews Genetics* **14**, 507-515.

Yang H, Ding YM, Hutchins LN*, et al.* (2009) A customized and versatile high-density genotyping array for the mouse. *Nature Methods* **6**, 663-U655.

Yang J, Manolio TA, Pasquale LR*, et al.* (2011a) Genome partitioning of genetic variation for complex traits using common SNPs. *Nature Genetics* **43**, 519-U544.

Yang JA, Benyamin B, McEvoy BP*, et al.* (2010) Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics* **42**, 565-U131.

Yang JA, Lee SH, Goddard ME, Visscher PM (2011b) GCTA: A Tool for Genome-wide Complex Trait Analysis. *American Journal of Human Genetics* **88**, 76-82.

Yezerinac SM, Lougheed SC, Handford P (1992) MEASUREMENT ERROR AND MORPHOMETRIC STUDIES - STATISTICAL POWER AND OBSERVER EXPERIENCE. *Systematic Biology* **41**, 471-482.

Young NM (2013) Macroevolutionary Diversity of Amniote Limb Proportions Predicted by Developmental Interactions. *Journal of Experimental Zoology Part B-Molecular and Developmental Evolution* **320**, 420-427.

Young NM, Hallgrimsson B (2005) Serial homology and the evolution of mammalian limb covariation structure. *Evolution* **59**, 2691-2704.

Zacco A, Cooper V, Chantler PD*, et al.* (1990) ISOLATION, BIOCHEMICAL-CHARACTERIZATION AND ULTRASTRUCTURAL ANALYSIS OF THE LIMBIC SYSTEM-ASSOCIATED MEMBRANE-PROTEIN (LAMP), A PROTEIN EXPRESSED BY NEURONS COMPRISING FUNCTIONAL NEURAL CIRCUITS. *Journal of Neuroscience* **10**, 73-90.

Zhou X, Carbonetto P, Stephens M (2013) Polygenic Modeling with Bayesian Sparse Linear Mixed Models. *Plos Genetics* **9**.

Zhou X, Stephens M (2012) Genome-wide efficient mixed-model analysis for association studies. *Nature Genetics* **44**, 821-U136.

Zhou X, Stephens M (2014) Efficient multivariate linear mixed model algorithms for genome-wide association studies. *Nature Methods* **11**, 407-409.