

# Phenotypic variability and genetic architecture of limbs in populations and strains of the house mouse (*Mus musculus*)

**Dissertation**

Zur Erlangung des Doktorgrades

“Doctor rerum naturalium“

der Mathematisch-Naturwissenschaftlichen Fakultät

der Christian-Albrechts-Universität zu Kiel

vorgelegt von

Neva Škrabar

Max-Planck-Institut für Evolutionsbiologie

Plön, Juni 2018

Erster Gutachter: Prof. Dr. Diethard Tautz  
Zweiter Gutachter: Prof. Dr. Thomas C. G. Bosch  
Tag der mündlichen Prüfung: 14.08.2018

# Table of Contents

Zusammenfassung .....	5
Summary .....	7
Introduction .....	9
Bone development .....	9
Limb bone evolution .....	11
Phenotypic variability .....	11
Canalization and developmental stability .....	12
Morphological integration .....	12
Complex traits .....	13
Genome-wide association studies .....	14
Origin of the house mouse .....	14
Hybrid zone .....	15
Inbred strains .....	16
Outbred stocks .....	17
Goals of the study .....	18
CHAPTER I .....	20
Fluctuating asymmetry in <i>Mus musculus</i> .....	20
Introduction .....	20
Methods .....	23
Mouse samples .....	23
Phenotypic measurements .....	24
Measurement error .....	26
Fluctuating asymmetry (FA) indices .....	27
Directional asymmetry (DA) .....	28
Normality and antisymmetry of the data .....	28
Character size .....	29
Differences between age and sex .....	29
Comparisons between populations / strains and groups .....	29
Results .....	30
Measurement error .....	30
Preliminary analyses per population / strain and sex .....	30

DA and antisymmetry per population / strain and hybrid subgroup .....	32
DA and antisymmetry per group .....	33
Differences in bone length .....	36
Comparisons of differences in FA .....	38
Differences between the bones per population / strain and group .....	38
Differences between populations / strains and groups for the same bone .....	40
Discussion .....	46
Overview of differences in bone length .....	46
Differences in FA level .....	46
Studies of FA in limb bones .....	49
Conclusions .....	50
CHAPTER II .....	52
Morphological integration in <i>Mus musculus</i> limb elements .....	52
Introduction .....	52
Methods .....	54
Mouse samples .....	54
Statistical analysis .....	55
Matrix repeatabilities .....	55
Matrix correlations .....	56
Partial correlations .....	57
Index of integration .....	57
Correction for variation due to size .....	58
Results .....	59
Limb bone lengths .....	59
Matrix repeatability .....	61
Correlation and covariance matrices .....	62
Similarities between matrices .....	64
Partial correlations in populations / strains and groups .....	65
Patterns of integration .....	67
Discussion .....	71
Conclusions .....	73
CHAPTER III .....	75
Using the <i>Mus musculus</i> hybrid zone to assess covariation	



and genetic architecture of limb bone lengths .....	75
Introduction.....	76
Methods .....	78
Mapping population .....	78
Phenotype measurements .....	78
Association mapping .....	78
Statistical analysis .....	79
Genetic correlation .....	80
Phenotypic variance explained by each chromosome .....	80
Results .....	80
Bone length measures and correlations .....	80
Genomewide association mapping and genetic architecture.....	80
Power to detect associations in the hybrid zone samples from simulation of additive model .....	81
Candidate genes .....	81
Identifying genetic associations unique to individual bones by accounting for covariation with another bone .....	81
Discussion .....	82
Candidate genes .....	83
Conclusion .....	86
Contributions to the thesis .....	90
Perspectives .....	91
Acknowledgments .....	92
References .....	94
Digital supplement .....	109
Curriculum vitae .....	112
Affidavit .....	113

## Zusammenfassung

Die Variabilität von Phänotypen ist ein generelles Phänomen in Tieren und Pflanzen und bildet die Basis für die evolutionäre Veränderung von Organismen. Die vorliegende Arbeit fokussiert sich auf die Variabilität und die Genetik von Längen Unterschieden in den einzelnen Knochen der Vorder- und Hinterfüße von Mäusen aus verschiedenen Populationen und Stämmen. Es wurden Mäuse von vier Auszucht Populationen verwendet (drei zugehörig zur Subspezies *M. m. domesticus* und eine zur Subspezies *M. m. musculus*), sowie von einem kommerziell erhältlichen Auszucht Stamm (CD1 - Charles River). Diese werden insgesamt als "Auszucht Gruppe" bezeichnet. Inzucht Tiere werden durch zwei Stämme repräsentiert, nämlich PWD (ein von *M. m. musculus* Wildtieren abgeleiteter Stamm) und C57BL/6J (ein klassischer Inzucht Stamm ursprünglich abgeleitet von *M. m. domesticus*). Weiterhin wurden Individuen aus der ersten Generation von Tieren verwendet, die in einer Hybridzone zwischen den Subspezies gefangen wurden. Um den Einfluss von Umweltbedingungen auf die Variabilität zu reduzieren wurden die Tiere vor der Analyse unter gleichen Bedingungen gehalten.

Die beiden ersten Kapitel widmen sich hauptsächlich der Frage der entwicklungsbiologischen Architektur. Kapitel 1 behandelt den Faktor fluktuierende Asymmetrie (FA). Diese wird als Indikator der entwicklungsbiologischen Stabilität angesehen, ausgedrückt darin, dass sich in einem Organismus der gleiche Phänotyp unter den gleichen Umwelt und genetischen Bedingungen auf beiden Seiten des Körpers entwickeln sollte. Die Ergebnisse meiner Analysen zeigen, dass der erste Knochen der Vorderfüße (der Humerus) generell die höchste FA zeigt. In Bezug auf die Länge der Knochen findet sich zwar die niedrigste Varianz in Inzucht Stämmen, aber zumindest der PWD Stamm zeigt gleichzeitig die höchsten Werte für FA. Für die Hybridtiere fand ich, dass der Grad der Hybridisierung FA nicht beeinflusst und gleichzeitig zeigen sie auch die geringsten Werte für FA. Diese Resultate legen einen Vorteil von Heterozygotität in den Hybridtieren nahe, während die größeren Abweichungen in der Inzucht Gruppe eine Konsequenz der höheren Homozygotität sein könnte.

Das zweite Kapitel befasst sich mit den Ko-Variationsmustern zwischen den Längen der einzelnen Knochen, eine Analyse die ein Maß für die morphologische Integration darstellt. Es wird angenommen, dass natürliche, oder nahezu natürliche Populationen auf Grund stabilisierender Selektion einen höheren Grad an Integration zeigen sollten, da diese die Ko-Varianz Struktur beeinflussen sollte. Meine Daten zeigen einen signifikant höheren Grad an Integration in den Hybrid Tieren im Vergleich zu den Inzucht und Auszucht Gruppen. Mit Ausnahme des Auszucht Stamms CD1 findet sich kein Unterschied in der

Integration zwischen den Auszucht und Inzucht Gruppen. Hingegen gibt es unterschiedliche Grade der Integration zwischen den Gruppen der Hybrid Tiere. Teilweise kann das auf unterschiedliche Größe als Faktor zurück geführt werden. Die Frage ob es eine höhere Integration zwischen entwicklungsbiologisch verwandten Knochen gibt, konnte für die Auszucht Gruppe und die Hybrid Tiere positiv beantwortet werden.

Das dritte Kapitel widmet sich einem Ansatz genetische Faktoren zu kartieren die die Variation der Gliedmaßen beeinflussen und stellt gleichzeitig die Frage ob genetische Variation mehr als einen Charakter beeinflusst. In diesem Teil der Studie wurden nur die Tiere aus der Hybridzone verwendet, da für diese schon vorher gezeigt wurde, dass sie sich für einen Genomweite Assoziationsstudie eignen. Basierend auf den Ergebnissen des Kapitels 2, in dem eine hohe phänotypische Korrelation gefunden wurde, stellte sich die Frage, ob Charakter die entwicklungsbiologisch und funktionell zusammen hängen auch gemeinsame genetische Varianten haben, die diese komplexen Strukturen generieren. Dementsprechend wurde nach Varianten gesucht, die sowohl die Länge einzelner Knochen beeinflussen, wie auch die Korrelation zwischen den Knochen. Die Ergebnisse zeigen generell eine hohe Erbllichkeit, basierend auf den genotypisierten Markern, sowie eine polygene Architektur. Ich konnte Kandidatengene identifizieren für die bereits früher ein Zusammenhang mit Gliedmaßen und Knochen Entwicklung gezeigt wurde, aber auch Varianten die nicht in früheren Kartierungsstudien für den untersuchten Phänotyp gefunden wurden. Ich konnte auch genetische Regionen (Loci) identifizieren die mit der Länge mehrerer Knochen und ihrer Korrelation assoziiert sind, also für seriell homologe Strukturen. Interessanterweise konnte aber für keines der Gene die an der Embryonalentwicklung der Gliedmaßen beteiligt sind eine Assoziation mit der Entwicklung im adulten Tier festgestellt werden.

## Summary

Variability of phenotypes is an ubiquitous phenomenon in animals and plants that is considered to be the basis on which organisms undergo evolutionary changes. The present thesis addresses the variability and genetic basis of limb bone length phenotypes in different populations and strains. I included mice from four outbred populations (three from the subspecies *M. m. domesticus* and one from the subspecies *M. m. musculus*), as well as the commercially available outbred stock CD1 (Charles River) called cumulatively the "outbred group". The inbred strains are represented by PWD (a wild derived inbred strain from *M. m. musculus*) and C57BL/6J (a classical inbred strain originally derived from *M. m. domesticus*) cumulatively called the "inbred group". Further I used individuals from the first-generation offspring of mice captured in a natural hybrid zone between the two *Mus musculus* subspecies. To reduce environmental influences on the variability, all animals were kept under the same environmental conditions before their analysis.

The first two chapters investigate mainly questions of developmental architecture. Chapter one deals with degrees of fluctuating asymmetry (FA). This is an indicator of developmental stability, reflected in the possibility of an organism to ensure the same phenotypic expression under the same genetic and environmental conditions on both sides of the body. I found that the first bone of the forelimb, the humerus, shows generally the highest level of FA. The lowest variance of bone length measures was found in inbred strains, but at least one of them (PWD) showed at the same time the highest level of FA. For the hybrid group I found the level of FA was not affected by the degree of hybridization and that they showed the highest level of stability. My results suggest a potential benefit of heterozygosity in the hybrid group, whereas the larger deviation observed in the inbred group might be a consequence of higher homozygosity.

The second chapter examines covariation between bone lengths, which are considered to reflect levels of morphological integration. Close to natural populations are assumed to show greater level of integration due to a stronger influence of stabilizing selection which should affect the covariance structure. I found a higher degree of integration in the hybrid animals which was significant compared to the inbred and outbred groups. Outbred populations and inbred strains did not have larger differences except significantly higher integration in one population that belongs to the outbred stock (CD1). Moreover, different levels of integration could be noticed among populations of the hybrid group. Influence of size was also found as an important factor in shaping the overall integration in each observed group. In

addition, I investigated whether stronger connections could be found between developmentally related bones, which was supported in outbred and hybrid groups.

The third chapter constitutes an approach to map genetic factors that generate limb variation and considers genetic variation that can affect multiple traits. In this part of the study, I used only mice from the hybrid zone, since these were previously shown to be suitable to conduct a genome wide association study with them. Based on the results from the second chapter which showed high phenotypic correlations, I asked whether traits that are developmentally and functionally related could have common genetic variants underlying these complex structures. Accordingly, special interest is given to genomic regions that underly individual bone length, as well as correlated variation between the bones. Overall, these traits revealed high heritability explained by genotyped markers, as well as a polygenic genetic architecture. Candidate genes previously described in limb and bone formation were identified together with genetic variants that were not previously reported in QTL studies of this phenotype. Further I found genetic regions (loci) associated with different bones, as well as high genetic correlations between the bones that share developmental mechanisms, i.e. serially homologous structures. Most interestingly, none of the genes known to be required for embryonic development of the limbs showed up as a factor shaping the adult development.

# Introduction

Survival and reproduction represent the two most important components in the evolution of all organisms. Therefore, numerous morphological, physiological, and behavioral adaptations are needed, as a response to different selection pressures. Many studies investigated adaptations and evolution of complex morphological traits, including genetics, developmental and ecological aspects. Interconnection between the genetic and developmental perspectives, proposed in the Atchley-Hall model (Atchley & Hall 1991) provided insights into the variation of quantitative traits. Variation represents the driving force in evolution and influences all phenotypic characters. It was proposed that it is of great importance to investigate mechanisms that generate and regulate variability in natural populations, as well as their interaction with the genetic architecture, which is commonly investigated in genotype – phenotype mapping.

Long bones have been shown to be a very suitable model system in evolutionary and developmental studies due to the great diversity of these structures, as well as their dependencies on each other (e.g. Kronenberg 2003; Farnum 2007; Butterfield et al. 2010). Different modes of locomotion in mammals are generated through structural and functional requirements which may lead either to dissociation of characters, or to developing strong connections among them. Studying the origin and evolution of phenotypic traits is oriented towards understanding mechanisms that underlie these modifications by intersecting genetic and developmental approaches.

## **Bone development**

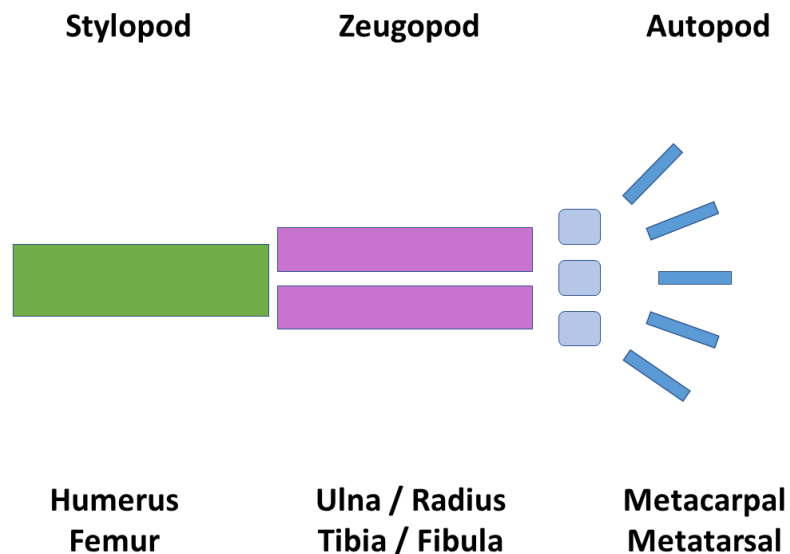
Multicellular organisms develop from many single cells through differentiation, followed by positioning of tissues and organs in different regions of the organism (Gilbert 2000a).

Generation of the skeleton includes three distinct lineages. The axial skeleton (the part of the skeleton that consists of the bones of the head and trunk of a vertebrate) is formed by the somites. The appendicular skeleton (limb skeleton) is generated from the lateral plate mesoderm, while the branchial arch, craniofacial bones and cartilage are made of the cranial neural crest (Gilbert 2000b).

Osteogenesis, or bone formation undergoes two processes which involve transformation of the preexisting mesenchymal tissue into bone tissue. Intramembranous ossification proceeds through direct conversion of the mesenchymal tissue into bone and it occurs in the bones of the skull. On the other hand,

endochondral ossification occurs in long bones and requires an additional step of the mesenchymal cell differentiation into cartilage, which is replaced by the bone cells (Gilbert 2000b).

The general skeletal architecture of the tetrapod limb encompasses a proximal stylopod (humerus / femur), a medial zeugopod (radius and ulna / fibula and tibia), and a distal autopod (carpals, tarsals, metacarpals, metatarsals and digits) (Young & Hallgrímsson 2005, Hall 2007) (Figure 1). Limb skeletal elements are formed in proximo-distal direction from the stylopod to the autopod. Fore- and hindlimbs represent serial homology in their structures, and these repeated parts are thought to share their developmental architecture (Hall 1995).



**Figure 1.** Schematic presentation of limb elements and bones within each corresponding element. The picture is modified from the original: Stylopod-zygopod-autopod.png from Prof. Dr. Peter Uetz. <https://commons.wikimedia.org/wiki/File:Stylopod-zygopod-autopod.png>

During embryonic development, patterning of the limb bud involves three main axes: proximo-distal (shoulder to finger tips), dorso-ventral (back of hand to palm), and antero-posterior (thumb to little finger). Abnormalities in proximal or distal parts can be caused by mutations in Hox genes (Wellik & Capecchi 2003).

Initiation of limb development depends on the expression of the T-box transcription factors, *Tbx5* (in forelimb) and *Tbx4* (in hindlimb) (Gibson-Brown et al. 1996; Tamura et al. 1999). Differential expression of these genes can result in morphological variability of fore- and hindlimbs (Gibson-Brown et al. 1996).

Further skeletal development involves several key transcription factors which are necessary for regular chondrogenesis and osteogenesis, among these are multiple members of the SOX family (Wright et al. 1995; Ng et al. 1997; Zhao et al. 1997) and the HOX homeobox transcriptional regulators (in chondrogenesis) (Dolle et al. 1993; Small & Potter 1993; Davis & Capecchi 1994, 1996). *Hox* genes were found to be more important in the later regulation of longitudinal growth of the individual elements (Hall 2007).

### **Limb bone evolution**

Tetrapods are evolutionarily derived from sarcopterygian fish (Cloutier & Ahlberg 1996) with the proximal parts of the limb (stylo- and zeugopod) thought to represent a plesiomorphic state, i. e. derived from the sarcopterygian fins, while the distal parts of the limbs (autopod) are thought to constitute evolutionary new traits (apomorphic) (Hall 2007). Development of the autopodium differs from the other limb segments due to its distal position in comparison to stylo- and zeugopodium (reviewed in Wagner & Chiu 2001).

Cellular and molecular mechanisms of early limb and fin development are fairly conserved. Similarities are found in initial phases of the development, starting with the bud formation. Further diversification between bony fishes and higher vertebrates involves differences in the bud patterning due to the individual developmental system of the cell-cell interaction (Hall 2007). Among vertebrates, chick and mouse show an overlap in early stages of limb bud development, as well as in signaling molecules and transcription factors involved in initiating limb bud development in embryos (Hall 2007).

### **Phenotypic variability**

Phenotypic variability reflects the potential of an organism to vary (Wagner & Altenberg 1996), with canalization, developmental stability and morphological integration being the main components of the variability (Hallgrímsson et al. 2002; Willmore et al. 2007). Other processes also involved in variability are phenotypic plasticity, heterochrony, and heterotopy (Willmore et al. 2007). Variability produces a range of possible outcomes, while developmental processes and their interactions confine expression of this variability. On the other hand, changes in the phenotype are generated through alterations in developmental mechanisms, caused by mutation, recombination and tissue interaction (Willmore et al. 2007).



### *Canalization and developmental stability*

The concept of canalization describes limitation in phenotypic variation among individuals (Wagner et al. 1997). A certain degree of this variation is preserved under different genetic and environmental conditions (Willmore et al. 2007). There are differences between genetic and environmental canalization, revealed by phenotypic robustness to the effects of mutations or environmental perturbations, respectively (Wagner et al. 1997).

Developmental stability is reflected in phenotypic consistency within individuals (Willmore et al. 2007). The main difference between canalization and developmental stability is in the type of variation they buffer (Willmore et al. 2007).

Processes, such as developmental noise, can affect frequency distribution of differences between the right and the left sides (right-minus-left) through the cumulative effects of small, random environmental perturbations (Waddington 1957; Lewontin 1983). In contrast, developmental stability reflects processes that minimize disturbance and enable development of a close to ideal phenotype (Palmer & Strobeck 2003). Developmental noise and stability both contribute to developmental instability which is shown as within-individual variation (Dongen 2006).

The fore mentioned processes are observed through the following patterns: 1) fluctuating asymmetry (FA), shown as random deviations from symmetry in bilaterally symmetrical organisms (Ludwig 1932 in Palmer & Strobeck 2003), with the variation between the right and the left side normally distributed around a mean of zero (Palmer & Strobeck 2003). 2) directional asymmetry (DA) is reflected in a tendency towards a bias in one direction. 3) antisymmetry with a bimodal distribution i.e. a pattern with stronger right side effect in half of the population and a stronger left side effect in the other half (Palmer & Strobeck 1986; Palmer & Strobeck 1992). Examples of FA are numerous in the literature, DA is mostly seen in human (mammalian) heart and brain, while antisymmetry is less prevalent and most prominently studied in male fiddler crabs, where one of the claws is always enlarged (Palmer & Strobeck 1992; Palmer 1996).

### *Morphological Integration*

Morphological integration observes different structures that are interconnected due to their common function and / or development, which are expected to benefit the overall stability of the organism (Willmore et al. 2007). This field of exploration was inspired by concepts of Olson and Miller (1958), where characters with shared developmental and functional influence form integrated units. Covariation

patterns among morphological traits were introduced through estimation of the level of covariation or correlation among the structures (Olson & Miller 1958). Cheverud (1996a) proposed three different levels of integration: individual, which involves functional and developmental integration (as described by Olson and Miller 1958); population level depicted in genetic integration based on pleiotropy or linkage disequilibrium; and evolutionary level, reflected in structures that evolve in a coordinated manner. Variability might be constrained in characters that are tightly connected, because change in one part of the structure would need to be beneficial for the other parts of the same structure, hence selection will act to reduce variability of individual parts (Wagner & Altenberg 1996).

### **Complex traits**

Variation caused by many genes and their interaction with environmental factors can shape complex traits, which are commonly investigated in studies of quantitative traits or medical diseases. Revealing the genetic basis of complex traits provided closer insights into genes, gene variants and molecular mechanisms that are related to one or more of these traits (Visscher et al. 2012). Numerous examples included studies of height (Turner et al. 2011; Wood et al. 2014) and weight (Locke et al. 2015), psychiatric disorders (Collins & Sullivan 2013; Otowa et al. 2016), autoimmune disease (Kochi 2016; Ramos et al. 2015) behavioral traits (Parker et al. 2016; Sanchez-Roige et al. 2018), educational attainment (Okbay et al. 2016) and many others. Evolution of complex traits was observed as well in different traits, such as wings of *Drosophila* (Kingsolver & Koehl 1994) and horns in beetles (Emlen et al. 2005). Skeletal elements, such as skull and mandibles, scapula, vertebrae, and limbs have also been much studied with quantitative trait approaches (Leamy et al. 1998; Leamy et al. 2002; Klingenberg et al. 2004; Kenney-Hunt et al. 2006; Norgard et al. 2008; Norgard et al. 2009). Combining morphometric statistics with complex trait statistics allowed even to approach studies of the shape of whole craniofacial skeleton of mice (Boell & Tautz 2011; Boell 2013; Boell et al. 2013; Pallares et al. 2014, 2015, 2016).

## Genome-wide association studies

In recent years, complex traits are mostly analyzed in genome-wide association studies (GWAS), which enable understanding of their genetic architecture, and can be used to investigate the genetic basis of natural variation. These studies are based on the identification of a large number of common genetic variants (usually single-nucleotide polymorphisms (SNPs)) across the genome, which are further tested for the association between each variant and a phenotype of interest (Bush et al. 2012). The genetic variation captured by SNPs or markers is usually termed as the 'narrow-sense heritability' (Wray et al. 2013). Nevertheless, variance explained by the genotyped SNPs is found to be lower than the proportion of phenotypic variance due to additive genetic variance (Yang et al. 2010).

In contrast to previous quantitative trait loci (QTL) studies, which usually reveal only large QTL regions that contain hundreds of genes (Gonzales & Palmer 2014), GWAS provided higher resolution and identification of individual genes (Flint & Mackay 2009; Parker & Palmer 2011). Most heritable traits are polygenic, i.e. controlled by many loci with minor phenotypic effects (Rockman 2012) and express continuous variation (Comeault et al. 2014). Large number of loci, each with small effects are explained with the infinitesimal model for the inheritance of quantitative traits (Barton et al. 2017, Boyle et al. 2017). This model assumes a combined effect of genetic and environmental components to shape quantitative traits, with a genetic component showing a normal distribution in offspring traits around the mean value of their parents (Barton et al. 2017).

Mammalian limbs are complex structures that express coordinated interactions between structure, function, and development. Several studies investigated the heritability in limb elements (reviews by Cock 1966; Thorpe 1981) and quantitative genetics of skeletal traits in mice (Leamy 1974, 1975, 1977; Leamy & Bradley 1982). High heritability of limb element lengths was described in Leamy (1974). Although limb heritability is higher than in skull or body traits, evolution of limb traits is still slower (on microevolutionary time scales) in comparison to dental or skull traits (Leamy & Sustarsic 1978).

## Origin of the house mouse

The present study focuses on the variation of limb elements in house mice, including wild derived populations. The house mouse (*Mus musculus*) is a small mammal of the order Rodentia. The genus *Mus* has originated in Southern Asia and is divided into several subspecies which have started to diverge about 0.5 million years ago (Sage 1981; Bonhomme et al. 1984). A reanalysis of the Eurasian collection of wild

mice based on microsatellite loci suggested the Iranian plateau as the main region for the initial house mouse diversification (Hardouin et al. 2015). Together with humans, it has spread across the globe during the past few thousand years. *Mus musculus musculus* has spread into Eastern Europe and Northern Asia, with a territory from Vladivostok to Scandinavia and Central Europe. *M. m. domesticus* is found in Western Europe, North Africa, and Near East. In the past few hundred years, it has further spread to the Americas, Australia and Oceania (Searle et al. 2009; Hardouin et al. 2010; Gabriel et al. 2011; Jones et al. 2011). *Mus musculus castaneus* inhabits Southeastern Asia including Japan, where it forms a hybrid zone with *M. m. musculus*. This hybridization is thought to have resulted in a hybrid subspecies called *M. m. molossinus*.

### *Hybrid zone*

*Mus m. domesticus* and *M. m. musculus* form a narrow hybrid zone from Scandinavia to the Black Sea. The range of the hybrid zone is more than 2500 km long and encompasses Denmark, Northern Germany (Schleswig-Holstein), Eastern Germany (Saxony), North-eastern Bavaria and Western Bohemia, Southern Bavaria and North-western Austria (Oberösterreich), as well as Eastern Bulgaria (see Baird & Macholán 2012 for a recent review).

Barton and Hewitt (1985) described hybrid zones as tension zones with a balance between dispersal of parental types and selection against hybrids (Haldane 1948; Fisher 1950). This was later confirmed for the mouse hybrid zone (Payseur et al. 2004; Raufaste et al. 2005; Macholan et al. 2007). There are various limitations to gene flow, with the prediction that less fit populations are more influenced (Barton & Hewitt 1985). Some studies reported mostly directed gene flow across the hybrid zone from *M. domesticus* into *M. musculus* populations (Vanlerberghe et al. 1988; Tucker et al. 1992; Fel-Clair et al. 1996; Boissinot & Boursot 1997; Raufaste et al. 2005).

Since hybridization in stable hybrid zones results in many recombinational events, natural hybrid zones represent a good tool to map traits associated with speciation (Rieseberg & Buerkle 2002). Among the most studied questions so far is the question of the genetics of hybrid sterility (Payseur & Hoekstra 2005). In the mouse hybrid zone, mapping hybrid sterility loci revealed lower mean values for traits in hybrids compared to pure subspecies (Turner et al. 2012). Studies with wild-caught hybrids (Albrechtova et al. 2012; Turner et al. 2012) and in *M. m. musculus* – *M. m. domesticus* hybrids generated in the laboratory (Britton-Davidian et al. 2005, reviewed in Good et al. 2008) have found reduced male fertility.

Hybridization may also produce novel phenotypes, called transgressive segregation. In this case phenotypic variation in offspring is expanded beyond the range of both parental species, which may be explained through complementary gene action (Rieseberg et al. 1999). In this phenomenon, phenotypes of hybrids after the F1 generation can be produced through additive effect of alleles at quantitative trait loci that generally have opposite sign in their parents (Rieseberg et al. 1999; Stelkens & Seehausen 2009). Transgressive segregation has been described for different morphological traits: skull morphology of cichlid fish (Albertson & Kocher 2005); shape in natural sculpine hybrids (Nolte & Sheets 2005); physiological traits (salt tolerance in *Helianthus* sunflowers (Lexer et al. 2003)); life history traits (flowering time in *Arabidopsis* (Clarke et al. 1995)); behavioural traits (mating behavior of *Drosophila* (Ranganath & Aruna 2003)) and behavioral traits in mice (Hiadlovská et al. 2013).

### *Inbred strains*

The first reported data with using mouse as laboratory animal came from Robert Hooke's study in 1664. This was further continued in the study of inheritance in mice, by William Castle in 1902 (Morse 1978). Afterwards, C. C. Little developed the first inbred mouse strain in 1909 (Holmes 2003, cited in Casellas 2011), which was followed with more inbred strains (in early 1920s) that found wide utility in scientific studies (Festing 1996, cited in Casellas 2011). Inbred strains are defined as colonies that have been sibling-mated (brother-sister) for at least twenty generations, which results in genetic uniformity within each strain (Flint & Eskin 2012). Inbred strains express high level of inbreeding (~98.6%), although a small proportion of genetic variance (~2%) still persists (Casellas 2011). Festing (1979) described the ideal inbred strain as isogenic and homozygous, phenotypically uniform, with high stability and easy to identify due to its genetic profile. Nevertheless, mutations can still arise and influence the genetic background of inbred strains. Inbred mice were used to generate several strains whose application subsequently expanded genetic researches (Flint & Eskin 2012). In order to detect sequence variants involved in phenotypic variation, a genetic cross was developed from the offspring of two inbred strains which is further mated to each other or to a progenitor strain (Flint & Eskin 2012). Common use of genetic cross and recombinant inbred strains provided advantages in quantitative trait loci studies and identification of the genetic basis of traits and diseases. However, requirements for higher mapping resolution and statistical power, as well as greater genetic diversity for genome-wide association studies developed new inbred strains such as the Hybrid Mouse Diversity Panel and Collaborative Cross (Flint & Eskin 2012). Another design with classical inbred strains generated heterogeneous stock mice (Valdar et al. 2006) and the diversity outbred

mice (Svenson et al. 2012) which belong to outbred mice. Mapping resolution was increased with this approach, while phenotyping and reproducibility were reduced in comparison with inbred strains (Flint & Eskin 2012). Inbred strains can be divided in two groups: classical inbred (with *M. m. domesticus* origin) and wild-derived strains (more recently derived, including some with origin from other subspecies).

### *Outbred stocks*

Outbred stocks are widely used in studies based on the assumption about their genetic similarity to natural mouse or human populations (Rice & O'Brien 1980). Their application is reported in toxicology, pharmacology and genetic researches (Chia et al. 2005), such as quantitative trait loci (QTLs) studies (Woods 2014), mapping of complex traits and genome-wide association studies (Aldinger et al. 2009; Yalcin et al. 2010). An outbred stock represents a population of genetically variable animals which has been closed for at least four generations. They are bred to preserve maximum heterozygosity (Festing 1993) and to reduce genetic change (Chia et al. 2005). Outbred stocks in mice descended from two male and seven female mice of the noninbred albino stock of André de Coulon (the Centre Anticancereux Romand in Lausanne, Switzerland). These mice were imported to the US by Clara J. Lynch in 1926 (Lynch 1969). One of the most reported outbred stocks include colonies of Swiss mice. Variation in the Swiss mice is lower compared to wild mice populations. On the other hand, these mice have shown more similarities than differences with natural populations (Rice & O'Brien 1980). In addition, classical outbred stocks and inbred strains differ considerably from wild mice in genetic architecture, size and productivity (Chia et al. 2005). Among Swiss mice, the CD1 mouse outbred stock is commonly used and it was first produced by Charles River Laboratories in 1959 (Rice & O'Brien 1980). Further, CD1 produces more offspring than (any) inbred strains and shows conspicuous differences in size (Poiley 1972; Chapin et al. 1993). Their main advantage over inbred strains is greater genetic and phenotypic variability (Falconer 1981, cited in Chia 2005).

Recent studies in commercial outbred colonies have estimated higher haplotype contribution from classical inbred strains than from wild-derived strains (Mott et al. 2000). Mice genetically similar to inbred strains could be involved in the origin of outbred stocks (Yalcin et al. 2010). Locus-specific and genome-wide sequencing approaches have shown that most of genetic variation (~about 95% of the polymorphisms) in outbred colonies derived from classical inbred strains (Yalcin et al. 2010).

Some of the important findings based on the genetic architecture of outbred stocks encompass large variation between colonies; limited number of segregating alleles, i.e., the same genetic variants could be

found in classical inbred strains; lower mapping resolution in comparison with human populations (Yalcin et al. 2010). Commercial outbreds could be used in association studies because these colonies showed low linkage disequilibrium, lower genetic diversity and heritable phenotypic variation. However, they cannot be used as a replacement for outbred population due to their limited genetic diversity (Yalcin et al. 2010).

## **Goals of the study**

The goals of this study are focused towards the exploration of natural genetic variability in the limb bones of the house mouse (*Mus musculus*), as well as using a GWAS approach to map loci that are involved in generating this variability. There are two major concepts of natural variability in the literature, one deals with left-right asymmetry, the other with co-variation between different elements. I approach both of these in my study. Both aim to approach a better understanding of the complex interactions between development, function, genetics and evolution. These questions require examination of a variety of organisms, as well as different characters in order to provide a broader picture of the numerous possible outcomes that exist in nature. My study contributes to this using a classic model trait - the limbs of tetrapods.

In the first chapter I observe within-individual variation, reflected in non-directional differences between right and left sides of paired bilateral characters (Thoday 1958). This question was investigated in different populations coming from outbred and inbred individuals, cumulatively named "outbred groups" and "inbred strains". Further, these groups (outbred and inbred) were compared to mice from the hybrid zone between *M. m. musculus* and *M. m. domesticus* (collected and previously studied in Turner et al. 2012; Pallares et al. 2014; Turner & Harr 2014). Differences were observed in each respective population / strain, as well as in groups. The main interest is in revealing which groups, as well as populations and strains are more or less affected by the disturbance in developmental processes.

The second chapter addresses the question of covariation between limb elements, which is of relevance for the underlying developmental and genetic factors that may structure this variation (Hallgrímsson 2002). The degree of covariation is thought to influence the evolutionary potential of biological organisms which is usually inferred from the strength of interactions between characters (Olson & Miller 1958; Van Valen 1965; Cheverud 1996a; Wagner & Altenberg 1996). For this analysis I used the same mouse samples as in chapter one to investigate the overall level of correlation among sets of traits.

The third chapter connects the developmental and genetic background of limb bones based on higher phenotypic correlations that were detected in the previous chapter. Cheverud (1984) found large similarity between phenotypic and genetic correlation patterns and stressed the importance of correlated responses to selection which further direct coordinated evolution (Lande 1979; Falconer 1981; Cheverud 1982, 1984). To get closer insight into the genetic architecture of the length of limb bones, a genome-wide association study was performed in the mice from the hybrid zone. These mice represent a valuable source for phenotype – genotype mapping due to larger phenotypic and genotypic variation and greater power to detect potential causal loci.



# CHAPTER I

## Fluctuating asymmetry in *Mus musculus*

### INTRODUCTION

The two sides of individuals in bilaterally symmetric animals share the same developmental program and hence are expected to develop symmetrically. On the other hand, one can often find small morphological differences between the right and the left side in symmetric organisms (Van Valen 1962). This type of asymmetry can differ between the individuals of the same population and was called fluctuating asymmetry (FA). It has been widely used as a measure of developmental stability (Mather 1953; Soule & Cuzin-Roudy 1982; Leary et al. 1985), with higher level of fluctuating asymmetry supposed to reflect lower stability (Van Valen 1962; Leary et al. 1985; Palmer & Strobeck 1986).

Apart of the influence of environmental factors, two genetic factors are thought to play an important role in the genetic basis of developmental stability: heterozygosity, associated with genetic variance in individuals and populations (Tomkins & Kotiaho 2002) and genomic co-adaptation, reflecting variances in interactions between loci (Graham & Felley 1985).

Morphological variation received special attention in evolutionary studies, since it could reveal mechanisms of its genetic and developmental basis (Lande 1979; Atchley et al. 1982; Boag 1983; Cheverud et al. 1983; Mukai & Nagano 1983). Different scenarios proposed possible explanations for the character variation, such as individuals with higher heterozygosity should express lower fluctuating asymmetry in bilateral characters (Lerner 1954; Van Valen 1962) while higher asymmetry is assumed in more homozygous individuals (Soule & Cuzin-Roudy 1982). Support for these hypotheses was found in the populations of salmonid fishes, where higher developmental stability was described in individuals with greater heterozygosity at protein loci (Leary et al. 1983, 1984). In contrast, this pattern was absent in many other studies (Wooten & Smith 1986; Patterson & Patton 1990; Clarke et al. 1992; Yampolsky & Scheiner 1994). Research with natural and laboratory-produced interspecific hybrids of salmonid fishes showed increased levels of fluctuating asymmetry (and hence reduced developmental stability) in hybrids in comparison to parental groups from the same environment (Leary et al. 1985).

Developmental stability in hybrids can be affected by changes in coadapted gene complexes. Dobzhansky (1950) hypothesized that genomes with the same evolutionary background express positive correlation

between heterozygosity at enzyme loci and developmental stability. Further, alteration in dominance relationship between alleles can be a consequence of changes in the genome and impact the developmental program. Therefore, alleles with deleterious effects on developmental stability that are not shown in intraspecific heterozygotes can be expressed in (interspecific) hybrids (Leary et al. 1985), e. g. in studies with *Drosophila* and fishes (Dobzhansky & Spassky 1968; Leslie & Vrijenhoek 1980).

Higher FA in homozygous individuals can also be due to a loss of genetic variance (Tomkins & Kotiaho 2002). Different examples reported positive correlation between inbreeding and FA (Robertson & Reeve 1952; Bader 1965; Leamy 1984, 1992) and a negative one among FA and allozyme heterozygosity in wild populations (Mitton & Grant 1984; Clarke 1993; Markow 1995). Nevertheless, results in various taxa also showed lack or even negative association between FA and inbreeding (mentioned in Tomkins & Kotiaho 2002). Different predictions are made in cases of hybridization. For instance, in the presence of inbreeding depression, i. e. reduced fitness due to inbreeding, hybridization should provide higher developmental stability (Tomkins & Kotiaho 2002). On the other side, hybridization between differentially adapted populations might disrupt coadapted gene complexes and increase levels of fluctuating asymmetry (Graham 1992).

The study of mandible variation in mice has yielded insights into the genetic background of asymmetries (Leamy et al. 1997, Leamy 1999). The authors found more QTLs for directional rather than fluctuating asymmetry. Moreover, mandibular characters from the backcross mouse population of two inbred strains (M16i and CAST) did not reveal any genetic basis for FA (Leamy et al. 2000). This result is consistent with the previous studies about QTLs (Leamy et al. 1997) and the heritability of FA in mouse mandibles (Leamy 1999). Further, FA in mouse mandibles might also depend on the epistatic interactions among some QTLs (Leamy et al. 2000). The developmental model used by Klingenberg & Nijhout (1999) implied that epistasis together with dominance could affect genetic variation of FA. Leamy et al. (2002) found a higher number of QTL interactions than single-locus QTLs that affect asymmetry in mandible size in a F2 population of mice produced from an intercross of the Large (LG/J) and Small (SM/J) inbred strains. This was confirmed in Leamy et al. (2015) with an F3 intercross of these strains and revealed influence of epistasis in the genetic architecture of asymmetry in mandible size.

In contrast to the mandible studies, QTLs for FA were reported in discrete skeletal characters in mice (Leamy et al. 1998), which shows the importance of the choice of characters for such an analysis. However, the number of these QTLs was rather small with a cumulative effect spanning from 0 - 15.3% of the total variation (Leamy et al. 1998). Estimated broad-sense heritability for FA of these traits averaged 0.065 and

confirmed an overlap with the narrow-sense heritability estimate (0.04) from the previous study on eight different discrete skeletal characters in random-bred mice (Leamy 1997).

Meta-analysis of fourteen species from different taxa, including plants, insects, reptiles, birds, fishes and mammals reported the mean estimate of FA heritability ( $h^2 = 0.27$ ) that was considered relatively low in comparison to the heritability of other traits (Moller & Thornhill 1997). More detailed analysis of different traits with a substantial statistical power, yielded an even lower heritability estimate ( $\sim 0.1$ ) (Leamy 1997; Whitlock & Fowler 1997).

Previous studies with mice from the Danish transect across the mouse hybrid zone between *M. m. musculus* and *M. m. domesticus* found lower FA in molars in mice from the center of the zone (Alibert et al. 1994), which was confirmed with laboratory hybrids of these two subspecies (Alibert et al. 1997). Similar results were found for the shape of the dorsal part of the skull (Debat et al. 2000), which indicated a possible advantage of heterotic effect that might influence the development of the whole skull. On the other hand, FA differences in size among hybrid and parental groups were not reported (Debat et al. 2000). In another study, lower FA in shape was not detected in the ventral side of the skull in mice from the Central European part of the hybrid zone (Mikula & Macholán 2008).

Conclusions about the influence of hybridization on developmental instability should be based on more traits, because fluctuating asymmetry can depend on the choice of the traits (Polak et al. 2003). Examples with differences in FA between traits are shown in studies of hybrids (Alibert & Auffray 2003). Further, an example with higher level of FA in one trait, but lower in other traits was reported in lab-bred hybrids between two chromosome races of *Sceloporus grammicus* lizards (Dosselman et al. 1998).

Functional traits are assumed to express lower level of FA (Balmford et al. 1993; Swaddle 1997; Brommer et al. 2003). Locomotion has an important role in the general performance of individuals with respect to predator pressure, with more symmetrical individuals being advantageous in survival (Galeotti et al. 2005). An example of the woodmice that are preyed by tawny owls discovered a higher level of asymmetry in prey individuals in comparison to individuals that survived. Although differences were not found in the mean size of hindlimb bones between the two groups of woodmice, mean asymmetry of hindlimb bones showed significant differences (Galeotti et al. 2005).

The goal of this chapter is to explore random deviations in limb elements between the two sides of individuals (FA) in different outbred populations and two inbred strains (see general introduction), as well as in comparison with the hybrid group. These differences are studied between different bones per population / strain or group and between groups for the same bone. Further, I examine the organism-

wide asymmetry based on information from all four traits pooled. This is observed in different populations / strains and groups. The main interest is in differences among populations / strains and between hybrid, outbred and inbred groups. From the skull examples reported in the literature, we would expect to find more instability in inbred groups and less in the hybrid, since we are using hybrids between two *Mus musculus* subspecies. However, limb bones and measurements of length, rather than shape may yield different outcomes.

## METHODS

### *Mouse samples*

Individuals included in this study include two general groups. The first consists of frozen specimens of first-generation male offspring of wild-caught mice collected by Leslie Turner in the hybrid zone in Bavaria in 2008 (Turner et al. 2012). The second group included five outbred populations (three from *M. m. domesticus*, one from *M. m. musculus*, as well as one population from the outbred stock (CD1)) and two inbred strains (classical inbred strain and wild-derived strain). The sampling procedure and breeding of the mice from the hybrid zone was previously described in Turner et al. (2012) and Turner and Harr (2014). The first generation of mice from the hybrid zone, as well as all other strains were raised under standard laboratory conditions to reduce the environmental effect on the traits (at the mouse facility of the Max-Planck Institute in Plön, Germany with the permit from the Veterinärämamt Kreis Plön: 1401-144/PLÖ-004697). They were sacrificed by CO<sub>2</sub> asphyxiation, followed by cervical dislocation at the age of 196 days (28 weeks) in outbred populations and inbred strains. I included also a few individuals with the following deviating age in days: 195 (1), 197 (7), 199 (3), 200 (2), 203 (1), 204 (1), 205 (5) and 206 (1). Mice from the hybrid zone were between 9 and 12 weeks of age.

The hybrid zone data set was divided into six groups based on the percentage of *M. m. domesticus* alleles found in each individual. The average percentage of *M. m. domesticus* alleles of the parents based on 37 diagnostic SNPs of the two subspecies of house mouse, *M. m. musculus* and *M. m. domesticus* (Turner et al. 2012) was used to estimate the percentage of *M. m. domesticus* alleles in the hybrid group (Pallares et al. 2016). The genotype data were provided by Leslie Turner and are listed as PairMid in Table S1.1, with the following division: H1 (8.3 – 18.3%), H2 (18.31 – 24%), H3 (24.01 -28.3%), H4 (28.31 – 38.3%), H5 (38.31 – 58.3%) and H6 (58.31 – 99.3%). The last two groups encompass data from the second half of the hybrid index (> 50%), because the majority of individuals included in the following analyses were grouped

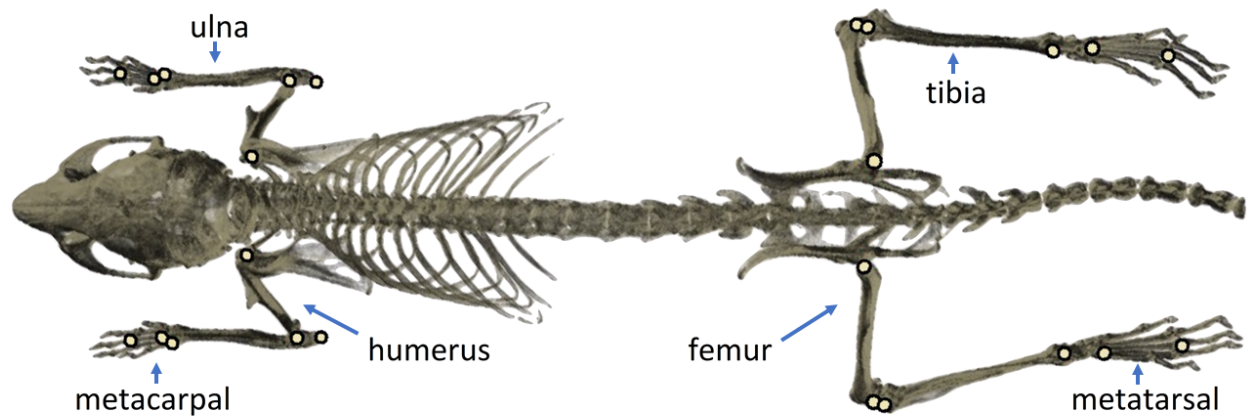
into the first half. Moreover, the second and the third groups were initially represented as one group, which was subsequently divided into two separate groups (H2 and H3) due to a larger number of individuals in this group, which also enabled similar sample sizes per group for further analyses.

The outbred samples from the second group consist of two subspecies of the house mouse, *M. m. musculus*, with a population KH (Almaty / Kasachstan) and *M. m. domesticus*, with CB (Köln / Bonn), MC (Massif Central) and AH (Ahvaz / Iran) populations. Further, I included also the commercially available outbred stock CD1 (Charles River). The inbred strains are represented by PWD (a wild derived inbred strain from *M. m. musculus*) and C57BL/6J (a classical inbred strain originally derived from *M. m. domesticus*).

### *Phenotypic measurements*

Mice were scanned with a computer tomograph (micro-CT-vivaCT 40, Scanco, Bruettisellen, Switzerland) with the following settings - energy: 70 kVp, intensity: 114  $\mu$ A, voxelsize: 38  $\mu$ m. I 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 hindlimb (femur, tibia and metatarsal bone) (Figure 1.1) using the TINA landmarking tool (Schunke et al. 2012). Two landmarks were used per left and right limb bone and linear measurements were obtained from distances between each pair of landmarks, calculated as the square root of the sum of the squared differences between each coordinate in three dimensions (Claude 2008). These distances were further multiplied by the scale factor which is expressed as the voxel size (0.038).

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; 3<sup>rd</sup> 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, 3<sup>rd</sup> 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 1.1. In this chapter only four bones were used: humerus, ulna, femur and tibia.



**Figure 1.1.** CT-scan of a mouse showing the limb bones measured in this study. The approximate positions of landmarks are indicated by yellow dots. Note that the actual landmarks were set in 3D representations of the skeleton. This figure corresponds to Figure 1 in Skrabar et al. (2018).

Presence of outliers was tested in each separate trait with visual inspection of scatter plots, by contrasting the difference between replicate measurements of the same bone (i. e. plots between the first and the second measurement within one side per bone (e. g. right humerus)); as well as between the right and the left side per trait where replicate measurements were averaged first (Palmer & Strobeck 2003). Data which showed larger differences between two measurements, due to blurred images and recording errors were excluded before all further analyses. Moreover, inspection of data measured once is also important, as greater differences between the two sides can arise due to injury, wear, disturbance of development which is not associated with developmental instability etc. (Palmer & Strobeck 2003). Further, these measurements can confound the analyses and cause spurious results. Therefore, differences between two sides were examined by visual inspection of histograms (Figures S1.1 - S1.2). Where larger deviation in one of the bones was found, the whole individual was excluded from further analysis. The number of individuals excluded per sex per group and their differences between raw right – left sides expressed in millimeters is shown in supplementary list 1.1. Among outliers, the largest differences between the right and the left sides were found in MC males in femur = -0.572, AH males in ulna = -0.695 and CD1 females in femur = -0.414. The highest number of individuals is excluded in CD1 females (10) and AH males (7).

### *Measurement error*

Measurement error was estimated from double measurements of the same image in the right and the left bone separately. 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). Number of individuals per group which were measured twice: hybrid (49), KH males (12), KH females (14), CB males (9), CB females (12), MC females (26), AH males (12), AH females (18). Double measurements were averaged in further calculations.

The two-way, mixed model ANOVA, with sides as fixed and individuals as random effects was used to test the significance of fluctuating asymmetry relative to measurement error (Palmer & Strobeck 1986, Palmer 1994, Palmer & Strobeck 2003). These tests were performed for each trait separately in the subset of individuals measured twice. Directional asymmetry and individual variation were tested from two-way ANOVA with repeated measurements of each side following the protocol from Palmer & Strobeck (2003) and Claude (2008). The analysis of fluctuating asymmetry can be conducted when between-sides variation ( $MS_{\text{interaction}}$ ) is significantly larger than the measurement error ( $MS_m$ ). Further, effect of the side ( $MS_s$ ) and individual variation ( $MS_i$ ) are tested against the interaction ( $MS_s / MS_{\text{interaction}}$  and  $MS_i / MS_{\text{interaction}}$ ) (Palmer 1994).

Additionally, this procedure allows computation of measurement error (ME3), where the average difference between replicate measurements is shown as the percent of average difference between sides (Palmer & Strobeck 2003). Repeatability, expressed with ME5 explains FA variation as a proportion of the total between sides variation, which includes measurement error. ME5 ranges between -1 to +1. Both estimates, percent of measurement error (ME3) and repeatability (ME5) are independent of the units of the measurement (Palmer & Strobeck 2003).

$$ME3 = \frac{MS_m}{MS_{\text{interaction}}} \times 100$$

$$ME5 = \frac{MS_{\text{interaction}} - MS_m}{MS_{\text{interaction}} + (n-1)MS_m}$$

Mean squares were obtained from a two-way ANOVA:  $MS_{\text{interaction}}$  explains sides by individual interaction mean squares;  $MS_m$  is the measurement error based on the variance of repeated measurements;  $n$  represents the number of repeat measurements.

These analyses were performed in each group, with all individuals measured twice before excluding the outliers based on larger differences between right and left side, since we were interested in testing

possible differences when all data are included. An exception was made with one hybrid individual which was the only one with a higher difference between sides found in femur bone and hence excluded prior to further analyses.

### *Fluctuating asymmetry (FA) indices*

Several indices are commonly used as descriptors of the level of FA in a sample, such as absolute asymmetry values ( $FA1 = \text{mean}|R-L|$ ); the average deviation of raw individual differences between two sides around the mean ( $FA4a = 0.798 \sqrt{\text{var}(R - L)}$ ); size adjusted index expressed through the absolute differences between natural logarithms of the right and the left side ( $FA8a = \text{mean}|\ln(R/L)|$ ) (Palmer & Strobeck 2003).

FA8a is computed from the ratio between two sides, however, this yields a skew in the frequency distribution of R/L which is further corrected by log transformation of the data (Palmer & Strobeck 2003). This index is directly comparable to an asymmetry estimate scaled for character size and it reflects FA as a proportion of the trait mean (Palmer & Strobeck 2003). FA8a facilitates comparison between traits of different sizes and it results in a dimensionless index of FA. This index is very sensitive to the presence of antisymmetry, directional asymmetry and measurement error. It will be denoted as FA8 in further text.

To estimate FA based on multiple traits, differences in FA due to trait size should be removed (Palmer & Strobeck 2003). Moreover, log transformations have been shown as an appropriate method in standardizing the variances between the traits (Lewontin 1966). A composite measure of FA per individual can be further calculated from averaging size-scaled values  $|\ln(R_j/L_j)|$  across the total number of traits (FA17).  $R_j$  and  $L_j$  represent measurements of the right and left side for trait  $j$  (Palmer & Strobeck 2003). The advantage of this index is a possibility to estimate an organism-wide developmental instability (DI), because it includes the information from multiple traits, while effects depending on size are excluded.

Developmental instability (DI) can be concluded only from traits which show ideal FA, i.e. when DA and departure from normality are absent. Regular tests for DA and antisymmetry include observation from subsamples of the data, as well as when data are pooled to a higher level (Palmer & Strobeck 2003). In this study, populations and sexes within population represent subsamples, while groups reflect the next level in the hierarchy.



### *Directional asymmetry (DA)*

It is of great importance to examine whether directional asymmetry (DA) can influence departure from symmetry (Palmer 1994, Palmer & Strobeck 2003). DA was tested with one sample *t*-test in each bone per group to determine if the mean values of right minus left side (R-L) differed significantly from zero. Due to small sample sizes in analyses per sex in each group, data were tested also with non-parametric Wilcoxon signed rank test (Dixon & Mood 1946). These tests were performed per group, as well as with pooled data (from all groups). Corresponding histograms of (R-L) distribution are shown in Figures S1.3 – S1.10. In addition, DA was estimated from log transformed data as  $DA_{\ln} = \ln(R/L)$ .  $DA_{\ln}$  corrects for trait size, with the positive values showing larger right side, while negative values represent larger left side (Palmer 1994, Palmer & Strobeck 2003). Departure from expected mean of zero was tested in  $DA_{\ln}$  values as well with one sample *t*-test and Wilcoxon signed rank test in groups with smaller sample sizes. Following Haag (2016) and Palmer & Strobeck (2003), I have calculated the average deviation around the mean (R-L) expressed as FA4a. Effects of DA, expressed as the mean (R-L) are considered removed, when traits show lower values of this estimate in comparison to FA4a.

### *Normality and antisymmetry of the data*

Distribution of (R-L) differences around a mean of zero was observed with one-sample Kolmogorov-Smirnov test (KS) with using standard deviation of (R-L) as the spread of these differences around that mean (Palmer & Strobeck 2003). Additional tests for skew (D'Agostino test for skewness in normally distributed data) and kurtosis (Anscombe-Glynn test of kurtosis for normal samples) were performed in R package 'momments' (Komsta & Novomestky 2015). Kurtosis shows information about the shape of the distribution with the value of 3 for univariate normal distribution. Values below 3 are considered as platykurtic, while above this value are leptokurtic (Liang et al. 2008). Skew is a measure of asymmetry of the distribution of a variable around its mean. Skew is zero in case of normal distribution. Negative values show distribution skew to the left, while positive values indicate on a right skew (Arnold & Groeneweld 1995).

Values for kurtosis were also compared with the table from Palmer & Strobeck (2003) for critical values of the kurtosis test statistic for deviations of frequency distributions from normality in the direction of platykurtosis and leptokurtosis for the appropriate sample size range. The table from Palmer & Strobeck (2003) is originally taken from Pearson & Hartley (1966) and D'Agostino (1986). For the purpose of the comparison with the table, a value of 3 was subtracted from the calculated data for kurtosis.

### *Character size*

Spearman and Kendall rank correlation coefficients were used to test associations between trait asymmetry  $|R-L|$  and trait size  $(R+L)/2$  (Palmer & Strobeck 2003) within each sex, population and group. Non-parametric tests avoid assumptions about homogeneity of variances, hence are more appropriate for FA data (Palmer & Strobeck 2003). In addition, differences in bone length between four traits, as well as between populations and groups for the same trait were tested with Kruskal-Wallis rank sum test. Separate tests were performed with average values between two sides for each sex, population / strain and group.

### *Differences between age and sex*

Hybrid individuals differed in age, with a range between 62 to 86 days old (nineteen age groups in total) and most of the individuals between 81 to 86 days old (93 individuals). Differences in unsigned deviations between different age groups were tested as absolute differences between sides ( $|R-L|$  and  $|\ln(R/L)|$ ) over age expressed in days with Spearman and Kendall rank correlation coefficients.

Differences in FA values between sexes and among traits in each population and group were tested with two-way ANOVA. In addition, Mann-Whitney-Wilcoxon test (Wilcoxon rank sum test) was performed for each bone within population and group between two sexes. Tests included both FA estimates (FA1 and FA8). Within populations / strains, tests were performed between all males and females (without outliers for larger differences between sides). Further, when data were pooled to a higher level (groups), differences between sexes were estimated from the data which were included (i.e. if we missed a bone of a certain sex in this group, it was not considered).

### *Comparisons between populations / strains and groups*

Differences in level of FA between traits per population / strain and group, as well as between populations / strains or group for each bone were tested with non-parametric Kruskal-Wallis analysis of variance, because asymmetry distributions  $|R-L|$  and  $|\ln(R/L)|$  (FA1 and FA8) are truncated at zero and skewed to the right (Palmer & Strobeck 1986, 2003). Descriptors of FA estimate the variance, therefore, tests for heterogeneity of variance such as F-test have been shown as the most appropriate for testing the differences in FA between individuals, traits and samples (Palmer 1994; Palmer & Strobeck 2003).

Additional tests for differences in FA between populations or groups and traits were performed with two-way ANOVA. All statistical analyses were performed in R version 3.2.5 (R Core Team 2016).

## RESULTS

### *Measurement error*

The measurement error (ME) of each separate bone length estimated with ANOVA was on average 0.295% across all comparisons included (measurement error of the right and of the left side in all bones) with the minimum found in left ulna in hybrids (0.043%) and the maximum in left ulna of KH males (1.120%) (Table S1.2).

Two-way, mixed model ANOVA in a subset measured twice revealed significantly larger between-sides variation than the measurement error ( $p < 0.001$ ), which supported further analyses of asymmetry variation (see Methods). Contribution of the error variance in the total between-sides variance (ME3) was between 2.301% in ulna of MC females, up to 10.469% in ulna of KH males (overall comparison included all bones in each tested group). Repeatability of FA (ME5) followed these results (Table S1.2).

Directional asymmetry (DA) tested with this approach revealed significant differences in humerus of KH males ( $p = 0.048$ ) and CB females ( $p = 0.040$ ), ulna of KH ( $p = 0.007$ ) and MC females ( $p = 0.022$ ) and tibia of hybrids ( $p = 0.021$ ). After including individuals with only one measurement, presence of DA with one-sample  $t$ -test for bones in these groups was found only in MC females, however, more groups with potential DA appeared (Table S1.3).

An additional test for significance of fluctuating asymmetry relative to measurement error was performed without individuals with larger deviation between sides. Between-sides variation was confirmed to be higher than the measurement error, although contribution of the error variance was slightly higher, with the highest values reported in femur of AH males (13.113%) and AH females (11.753%). Directional asymmetry was found in the same bones as reported, except for ulna of CB females (data not shown).

### *Preliminary analyses per population / strain and sex*

After removing outliers based on larger differences between two sides (see Methods and Supplementary List 1.1 and Figures S1.1 – S1.2), preliminary analyses were first performed in all populations and sexes.

This observation revealed significant values in DA tests for humerus of CD1 and C57BL/6J, both males and females, ulna of MC and PWD females, femur and tibia of AH females, while marginal DA was found in humerus of MC females. Close to marginal DA was found in humerus of CB males, ulna of CB females, femur of MC males, tibia of KH males (Wilcoxon signed test revealed significance) and PWD females (Table S1.4a and Figures S1.3 - S1.5). Significant departure from normality was detected in humerus of CD1 females and C57BL/6J, both females and males, ulna PWD females, femur in MC males and AH females, tibia of KH males. Skew was marginal in ulna of PWD females and close to marginal in ulna of KH females and tibia in KH, both males and females. Kurtosis was found in humerus of CD1 females and tibia of PWD males; close to marginal in humerus of AH males and PWD females (Table S1.5a and Figures S1.3 - S1.5). Association of FA to character size was present in humerus of CB males, ulna and tibia of MC females, femur C57BL/6Jf and tibia CD1 males (Table S1.6a).

After Bonferroni correction for multiple tests across four traits in respective groups, significant DA remained in humerus C57BL/6J males and females, ulna of PWD females and femur of AH females. Further, departure from normality was detected in ulna of PWD females and tibia of KH males, while kurtosis was present in tibia of PWD males.

Both FA estimates (FA1 and FA8) were first compared for differences between sexes with two-way ANOVA, which revealed close to marginal significances in MC and AH populations with unscaled estimates, while data calculated with size-scaled FA ( $|\ln(R/L)|$ ), indicated differences in KH and AH populations, although only marginal (Table S1.7). Neither group showed a significant interaction term, except marginal in PWD population with unscaled FA. This would be interpreted as differences in FA between sexes depending on the trait (Palmer & Strobeck 2003). Further, separate analyses per bone in each population and group with Mann-Whitney-Wilcoxon test, yielded significant differences between sexes in humerus of AH, femur of MC and tibia of PWD with  $|R-L|$  values. Similar results were found with scaled differences, with additional marginal influence noted for ulna in PWD, while tibia of PWD was close to marginal (Table S1.8). However, significance was not detected after Bonferroni correction for multiple tests. In the hybrid group, only males were included, therefore, they were omitted from these tests.

The hybrid group was tested for differences in FA due to age. An influence was found in ulna and marginal in humerus of the first hybrid subgroup in both data, tested as FA1 and FA8 over age (Table S1.9). However, the data were not significant after Bonferroni correction for multiple tests.

After tests for differences in FA between age groups in hybrids and between sexes in each population / strain and group, no major differences appeared, therefore the data were pooled and used in further analyses.

Due to marginal and significant differences which arose per group when data were pooled for sex per group, which directed even larger differences when data were pooled to outbred and inbred groups (data not shown), further analyses included data only from sexes per bone which did not express larger deviations in distribution of raw (R-L) values. Therefore, the following bones per sex were excluded (with abbreviation for the type of asymmetry, DA refers to directional asymmetry, KS stands for departure from normality tested with Kolmogorov-Smirnov): humerus in AH males (kurtosis), CB males (DA close to marginal), MC females (marginal DA), CD1 males and females (DA in both, with KS in females), C57BL/6J males and females (DA, KS), PWD females (kurtosis); ulna in PWD females (DA, KS); femur in AH females (DA, KS), MC males (close to marginal DA, KS); tibia in KH males (DA with Wilcoxon signed rank test, KS, skew), PWD males (kurtosis) (Tables S1.4a and S1.5a).

#### *DA and antisymmetry per population / strain and hybrid subgroup*

Analyses per population / strain with pooled sexes revealed significant DA in ulna of MC, although non-significant after Bonferroni correction (Table 1.1a and Figure S1.6). Among hybrids, the second group showed significant differences (DA) in humerus and femur, as well as ulna in the sixth group, marginal influence was detected in humerus and femur of the first group (Table S1.4b and Figure S1.8). Significant departure from normality was found in ulna of MC and tibia of C57BL/6J, as well as in humerus of the first two hybrid groups. Significant kurtosis was seen in femur of CD1 and close to marginal in ulna of PWD strain and tibia in the first hybrid group (Table 1.2a and Figures S1.7 – S1.8). However, departure from normality and antisymmetry were not present after Bonferroni correction per population and strain. Relation with character size was found in ulna and tibia of MC, tibia of CD1, ulna and femur of the second and humerus in the fifth hybrid group, close to marginal in tibia of the third and humerus in the fourth group (Tables S1.6b and S1.6c).

Bonferroni correction across four traits per relevant group among hybrids revealed the following differences: marginal DA in humerus and femur in the second group. This result was further reflected in significant DA when all hybrid groups were pooled together for humerus and femur (data not shown). Departure from normality (KS) was found significant in humerus of the second group, as well as in humerus and femur in pooled data. Kurtosis was also found significant in pooled data for humerus (data

not shown). Due to larger deviations detected in humerus and femur, mostly from the first two groups, these were omitted from further analyses. In addition, data from the third group were excluded, because preliminary analyses with pooled data still showed departures in these two bones.

#### *DA and antisymmetry per group*

When populations / strains were pooled to groups, significant DA was found in femur of hybrid and tibia of the outbred group (Table 1.1b). Departure from normality in femur of the hybrid and humerus of an outbred group, as well as kurtosis in femur of the outbred group (Table 1.2b and Figure S1.9). Afterwards, all data for each bone were pooled together (Tables 1.1b and 1.2b); departure from normality was found only in humerus, while significant kurtosis was detected in femur and marginal in humerus (Figure S1.10). Kurtosis in femur remained significant after Bonferroni correction. Character size was found significant for the femur of an outbred group, which was reflected also when data were pooled together for this bone (Table S1.6d).

DA and  $DA_{in}$  values were also compared to FA4a index, which showed lower values than the index in all observations with pooled populations and groups, hence DA did not bias the tested data.

Preliminary analyses of DA and antisymmetry with data used in index comparison did not reveal a major influence of these confounding factors. Spurious conclusions caused by greater signed differences between sides (R-L) could bias FA analyses and inflate the variation (Palmer &Strobeck 2003).

**Table 1.1.** Tests for directional asymmetry (DA) with raw (R-L) and scaled  $\ln(R/L)$  values per population / strain (a) and group (b). Data with excluded bones that showed larger deviation per sex in each population / strain. Significant values are highlighted in blue, while marginal values are shown in bold.

(a)

Group	Population Strain	Bone	n	DA = (R-L)			DA <sub>ln</sub> = $\ln(R/L)$			FA4a
				Mean	t	p	Mean	t	p	
Outbred	KH	Humerus	65	-0.018	-1.458	0.150	$-1.64 \times 10^{-3}$	-1.432	0.157	0.078
		Ulna	65	0.016	1.420	0.160	$1.23 \times 10^{-3}$	1.462	0.149	0.071
		Femur	65	0.002	0.164	0.871	$9.29 \times 10^{-5}$	0.142	0.887	0.061
		Tibia	30	0.003	0.173	0.864	$1.54 \times 10^{-4}$	0.165	0.870	0.065
	CB	Humerus	31	-0.009	-0.564	0.577	$-8.39 \times 10^{-4}$	-0.597	0.555	0.073
		Ulna	61	0.010	0.909	0.367	$6.37 \times 10^{-4}$	0.875	0.385	0.066
		Femur	61	0.009	0.745	0.459	$5.93 \times 10^{-4}$	0.794	0.430	0.074
		Tibia	61	0.017	1.681	0.098	$9.65 \times 10^{-4}$	1.680	0.098	0.063
	MC	Humerus	27	-0.001	-0.050	0.960	$-9.51 \times 10^{-5}$	-0.054	0.957	0.085
		Ulna	55	0.023	2.350	<b>0.022</b>	$1.63 \times 10^{-3}$	2.347	<b>0.023</b>	0.057
		Femur	28	-0.009	-0.610	0.547	$-5.86 \times 10^{-4}$	-0.596	0.556	0.063
		Tibia	55	0.005	0.520	0.605	$2.78 \times 10^{-4}$	0.463	0.646	0.062
	AH	Humerus	29	-0.023	-1.693	0.102	$-1.80 \times 10^{-3}$	-1.679	0.104	0.058
		Ulna	52	-0.009	-0.624	0.536	$-6.24 \times 10^{-4}$	-0.662	0.511	0.082
		Femur	23	0.019	1.097	0.285	$1.08 \times 10^{-3}$	1.084	0.290	0.066
		Tibia	52	0.022	1.882	<b>0.066</b>	$1.14 \times 10^{-3}$	1.858	<b>0.069</b>	0.067
CD1	Ulna	48	0.003	0.261	0.795	$1.73 \times 10^{-4}$	0.211	0.834	0.067	
	Femur	48	-0.003	-0.189	0.851	$-1.97 \times 10^{-4}$	-0.192	0.849	0.092	
	Tibia	48	0.006	0.414	0.681	$3.44 \times 10^{-4}$	0.442	0.661	0.079	
Inbred	PWD	Humerus	33	-0.004	-0.203	0.841	$-3.46 \times 10^{-4}$	-0.222	0.826	0.079
		Ulna	33	0.026	1.471	0.151	$2.04 \times 10^{-3}$	1.459	0.154	0.082
		Femur	60	0.006	0.504	0.616	$3.69 \times 10^{-4}$	0.482	0.632	0.069
		Tibia	27	-0.032	-1.787	0.086	$-2.12 \times 10^{-3}$	-1.783	0.086	0.075
	C57BL/6J	Ulna	57	-0.013	-1.345	0.184	$-9.11 \times 10^{-4}$	-1.352	0.182	0.057
		Femur	57	-0.007	-0.515	0.609	$-4.35 \times 10^{-4}$	-0.510	0.612	0.082
		Tibia	57	0.024	1.913	<b>0.061</b>	$1.36 \times 10^{-3}$	1.895	<b>0.063</b>	0.077

(b)

Group	Bone	n	DA = (R-L)			DA <sub>ln</sub> = $\ln(R/L)$			FA4a
			Mean	t	p	Mean	t	p	
Hybrid	Humerus	77	-0.005	-0.553	0.582	$-5.02 \times 10^{-4}$	-0.629	0.531	0.0635
	Ulna	155	-0.004	-0.689	0.492	$-2.98 \times 10^{-4}$	-0.681	0.497	0.0591
	Femur	77	-0.023	-2.271	<b>0.026</b>	$-1.52 \times 10^{-3}$	-2.241	<b>0.028</b>	0.0703
	Tibia	155	-0.009	-1.411	0.160	$-5.66 \times 10^{-4}$	-1.412	0.160	0.0659
Outbred	Humerus	152	-0.014	-1.849	<b>0.066</b>	$-1.23 \times 10^{-3}$	-1.828	<b>0.070</b>	0.075
	Ulna	281	0.009	1.757	<b>0.080</b>	$6.55 \times 10^{-4}$	1.807	<b>0.072</b>	0.069
	Femur	225	0.003	0.491	0.624	$1.84 \times 10^{-4}$	0.477	0.634	0.072
	Tibia	246	0.012	2.156	<b>0.032</b>	$6.28 \times 10^{-4}$	2.090	<b>0.038</b>	0.067
Inbred	Humerus	33	-0.004	-0.203	0.841	$-3.46 \times 10^{-4}$	-0.222	0.826	0.079
	Ulna	90	0.002	0.171	0.865	$1.70 \times 10^{-4}$	0.250	0.803	0.068
	Femur	117	-0.001	-0.064	0.949	$-2.27 \times 10^{-5}$	-0.040	0.968	0.075
	Tibia	84	0.006	0.567	0.572	$2.40 \times 10^{-4}$	0.376	0.708	0.079
All together*	Humerus	262	-0.010	-1.805	<b>0.072</b>	$-9.06 \times 10^{-4}$	-1.830	<b>0.068</b>	0.072
	Ulna	526	0.004	1.080	0.281	$2.91 \times 10^{-4}$	1.120	0.263	0.066
	Femur	419	-0.003	-0.614	0.539	$-1.86 \times 10^{-4}$	-0.642	0.522	0.073
	Tibia	485	0.004	0.995	0.321	$1.79 \times 10^{-4}$	0.785	0.433	0.069

Table 1.1. (continued): Group – defines a group to which a population / strain belongs (a) and data for pooled populations / strains (b). Population / Strain – defines used population / strain. Bone – limb bone used in the test. n – number of individuals used per bone per sample. Mean – mean value of (R-L) and ln(R/L) tested for departure from zero. Test for DA with one-sample t-test (t) and Wilcoxon signed rank test (V). FA4a – the average deviation around the mean (R-L) (Palmer and Strobeck 2003).

\* All data used per bone (i. e. complete data for each bone pooled from all three groups).

**Table 1.2.** Tests for departure from normality, skew and kurtosis in raw (R-L) data per population / strain (a) and group (b). Significant values are highlighted in blue, while marginal values are shown in bold.

(a)

Group	Population Strain	Bone	n	Kolmogorov - Smirnov		Skew		Kurtosis	
				D	p	Skew	p	Kurtosis	p
Outbred	KH	Humerus	65	0.126	0.237	0.143	0.609	2.307	0.166
		Ulna	65	0.135	0.174	-0.356	0.210	2.715	0.836
		Femur	65	0.062	0.954	0.207	0.460	2.460	0.377
		Tibia	30	0.136	0.591	0.734	<b>0.071</b>	2.610	0.924
	CB	Humerus	31	0.137	0.560	0.341	0.373	2.430	0.642
		Ulna	61	0.103	0.504	-0.131	0.648	3.484	0.261
		Femur	61	0.100	0.543	-0.042	0.884	2.308	0.187
		Tibia	61	0.150	0.115	-0.113	0.694	3.372	0.329
	MC	Humerus	27	0.174	0.347	0.613	0.140	2.601	0.947
		Ulna	55	0.190	<b>0.033</b>	-0.278	0.358	2.747	0.942
		Femur	28	0.094	0.945	-0.223	0.573	2.780	0.831
		Tibia	55	0.068	0.944	0.160	0.594	2.271	0.179
	AH	Humerus	29	0.201	0.169	0.040	0.918	2.592	0.909
		Ulna	52	0.107	0.553	0.111	0.718	2.359	0.313
		Femur	23	0.157	0.571	-0.119	0.778	2.451	0.799
		Tibia	52	0.124	0.373	0.324	0.297	2.480	0.502
CD1	Ulna	48	0.071	0.953	0.295	0.357	2.712	0.924	
	Femur	48	0.120	0.460	0.087	0.783	2.002	<b>0.026</b>	
	Tibia	48	0.080	0.891	0.023	0.942	2.273	0.230	
Inbred	PWD	Humerus	33	0.107	0.804	-0.028	0.939	2.133	0.208
		Ulna	33	0.168	0.276	-0.148	0.689	1.992	<b>0.086</b>
		Femur	60	0.060	0.973	0.214	0.460	2.434	0.367
		Tibia	27	0.216	0.139	0.277	0.490	1.978	0.132
	C57BL/6J	Ulna	57	0.141	0.188	0.228	0.442	2.535	0.559
		Femur	57	0.091	0.701	0.122	0.680	2.311	0.214
		Tibia	57	0.179	<b>0.046</b>	-0.320	0.285	2.843	0.911



(b)

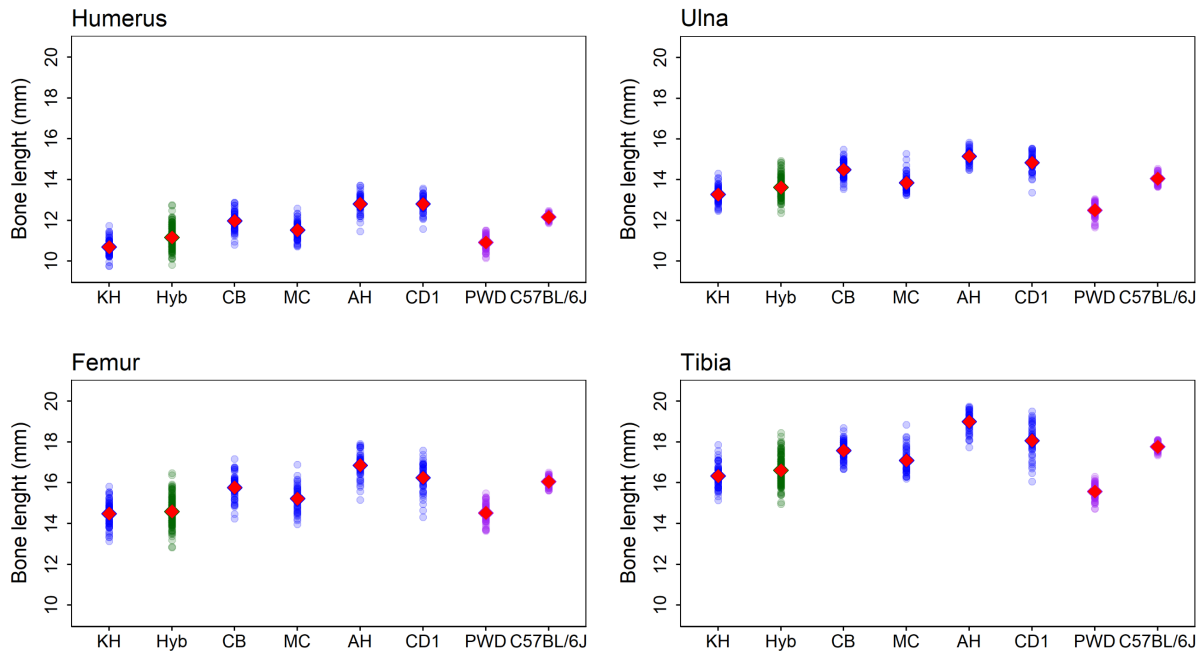
Group	Bone	n	Kolmogorov - Smirnov		Skew		Kurtosis	
			D	p	Skew	p	Kurtosis	p
Hybrid	Humerus	77	0.094	0.478	-0.143	0.582	2.351	0.163
	Ulna	155	0.051	0.823	-0.182	0.339	2.850	0.870
	Femur	77	0.157	<b>0.039</b>	0.161	0.535	2.593	0.544
	Tibia	155	0.069	0.448	-0.090	0.632	2.686	0.476
Outbred	Humerus	152	0.116	<b>0.033</b>	0.327	0.093	2.641	0.385
	Ulna	281	0.070	0.132	-0.136	0.341	2.798	0.561
	Femur	225	0.041	0.849	0.018	0.910	2.462	<b>0.038</b>
	Tibia	246	0.070	0.175	0.149	0.328	2.636	0.214
Inbred	Humerus	33	0.107	0.804	-0.028	0.939	2.133	0.208
	Ulna	90	0.058	0.900	0.225	0.356	2.375	0.144
	Femur	117	0.061	0.775	0.121	0.573	2.435	0.139
	Tibia	84	0.078	0.659	-0.119	0.632	2.365	0.152
All together*	Humerus	262	0.093	<b>0.021</b>	0.166	0.263	2.518	<b>0.054</b>
	Ulna	526	0.042	0.302	-0.053	0.614	2.765	0.268
	Femur	419	0.051	0.223	0.080	0.500	2.471	<b>0.004</b>
	Tibia	485	0.030	0.777	0.022	0.845	2.661	<b>0.088</b>

*Group* – defines a group to which a population / strain belongs (a) and data for pooled populations / strains (b). *Population / Strain* – defines used population / strain. *Bone* – limb bone used in the test. *n* – number of individuals used per bone per sample. *One-sample Kolmogorov-Smirnov test* for departure from normality (*D* – test statistics and *p* – value). Tests for *skew (D’Agostino)* and *kurtosis (Anscombe-Glynn)* with corresponding *p*-values.

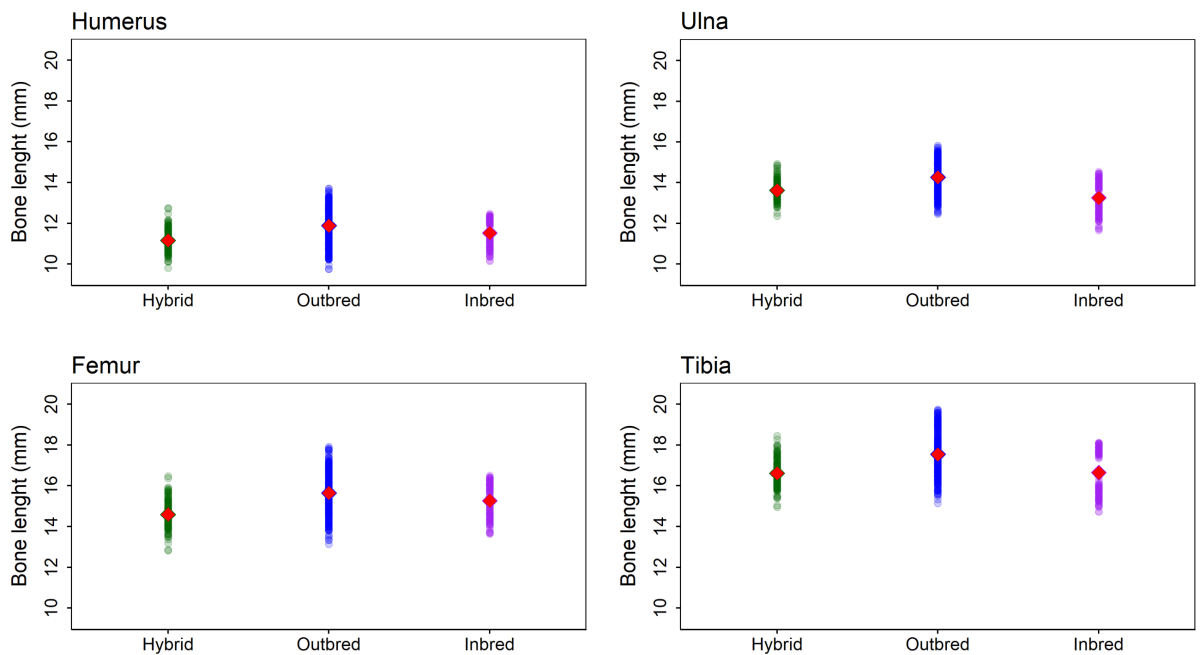
\* All data used per bone (i. e. complete data for each bone pooled from all three groups).

### *Differences in bone length*

Due to significant differences in bone length of the same bone between different populations / strains and groups (data for all males and females per population / strain, with hybrid group included (553 individuals)), and in the length of different bones (Figures 1.2 - 1.3 and S1.11) (Kruskal-Wallis rank sum test,  $p < 2.2 \times 10^{-16}$ ), the size-adjusted index is a better option for comparing the differences in FA levels between populations / strains and groups for the same trait and between different traits. Only in the hybrid group differences for the same bone were not found between H1, H2 and H3, but these three groups differed from H5 and H6 which were not significant among themselves. H4 showed differences with H2 in ulna, femur and tibia; H3 did not differ from H6 in femur (data not shown). Descriptive statistics for each bone per population / strain and group are shown in Table S1.10.



**Figure 1.2.** Bone length averaged between the right and the left side in forelimb bones (humerus, ulna, upper graphs) and hindlimb bones (femur, tibia, lower graphs) in outbred: *M. m. musculus* (KH), hybrid group (with hybrid subgroups pooled), *M. m. domesticus* (CB, MC, AH, CD1) and two inbred strains (PWD, C57BL/6J). The red diamond shows the mean bone length value in each population / strain.



**Figure 1.3.** Bone length averaged between the right and the left side in forelimb bones (humerus, ulna, upper graphs) and hindlimb bones (femur, tibia, lower graphs) in the hybrid group (with hybrid subgroups pooled), outbred and inbred groups (with populations / strains pooled per group). The red diamond shows the mean bone length value in each group.

### *Comparisons of differences in FA*

Two-way ANOVA between groups and among traits performed with size-scaled  $|\ln(R/L)|$  data showed significant variability among traits in hybrid and outbred group, whereas only marginal differences in the inbred group, which implies different levels of asymmetry between traits. Differences between populations were found in an outbred group and among two inbred strains, as well when data were pooled together to hybrid, outbred and inbred group. Interaction term (Group x Trait) was significant in an outbred and inbred group, which explains different levels of asymmetry between populations / strains, depending on information from multiple traits (Table S1.11). Therefore, additional tests with dividing mean squares of the group over the mean squares of the interaction were performed, although no significant result was found for differences between populations in each respective group.

Previous tests with Spearman and Kendall correlation coefficient did not detect dependence of FA on trait size between individuals from different groups in each separate trait (Table S1.6), therefore, additional tests with unscaled data were performed as well (Table S1.11). Significant differences were present only in the interaction term of outbred and inbred groups and between groups in all data pooled together. A marginal difference between populations was found only among outbred groups.

### *Differences between the bones per population / strain and group*

Separate tests were performed between the traits within each group with Kruskal-Wallis rank sum test in size-scaled data (Table 1.3). Results showed significant differences between the bones in all outbred populations, as well as in the PWD strain from an inbred group. Pairwise comparison using Wilcoxon rank sum test with Bonferroni correction showed the following differences (the first letter stands for a corresponding bone): in KH population: H-F, H-T; CB pop: H-T and marginal in H-U; MC pop: H-U, H-T; AH: U-T; CD1: marginal between U-F (Figure S1.12), while significance was absent in PWD after correction (Figure S1.13). Among hybrid subgroups, the first three groups were tested only for differences between ulna and tibia with Wilcoxon rank sum tests, while other groups were tested for differences between all four bones with the aforementioned test. Nevertheless, none of the hybrid subgroups showed significant differences between the bones (Figure S1.14).

Populations and strains were further pooled to groups, this comparison yielded significant differences in hybrids between H-U and H-T, in the outbred group between H-U, H-F, H-T, U-T and F-T, while the inbred group showed significance in H-T and marginal in H-F across four traits (Table 1.3 and Figure 1.4). When

data from all groups were pooled together, significant differences were reflected between all bones, except between U-F (Table 1.3 and Figure S1.15). Additional tests with unscaled data found differences only in CD1 between U-F, as well as when all groups were pooled together (Table S1.12).

**Table 1.3.** Comparison of differences in fluctuating asymmetry levels between bones with Kruskal-Wallis test and Mann-Whitney U test (when only two bones were included).

Group	Subgroup / Population / Strain	N of ind per bone H / U / F / T	Kruskal-Wallis rank sum test on  ln(R/L)		
			W		p
Hybrid	H1 <sup>1</sup>	25 / 25	363		0.335
	H2 <sup>1</sup>	31 / 31	490		0.900
	H3 <sup>1</sup>	22 / 22	280		0.382
			$\chi^2$	df	p
Hybrid	H4	20 / 20 / 20 / 20	4.106	3	0.250
	H5	25 / 25 / 25 / 25	1.689		0.640
	H6	32 / 32 / 32 / 32	6.138		0.105
Outbred	KH	65 / 65 / 65 / 30	12.979	3	<b>0.005*</b>
	CB	31 / 61 / 61 / 61	13.699		<b>0.003*</b>
	MC	27 / 55 / 28 / 55	9.967		<b>0.019*</b>
	AH	29 / 52 / 23 / 52	8.902		<b>0.031*</b>
Inbred	CD1	48 / 48 / 48	7.031	2	<b>0.030</b>
	PWD	33 / 33 / 60 / 27	8.462	3	<b>0.037</b>
	C57BL/6J	57 / 57 / 57	1.898	2	0.387
Hybrid	Humerus = 77 Ulna = 155 Femur = 77 Tibia = 155		16.195	3	<b>0.001*</b>
Outbred	Humerus = 152 Ulna = 281 Femur = 225 Tibia = 246		38.112		<b>2.68 x 10<sup>-8</sup>*</b>
Inbred	Humerus = 33 Ulna = 90 Femur = 117 Tibia = 84		7.700		<b>0.053*</b>
All together**	Humerus = 262 Ulna = 526 Femur = 419 Tibia = 485		56.853		<b>2.76 x 10<sup>-12</sup>*</b>

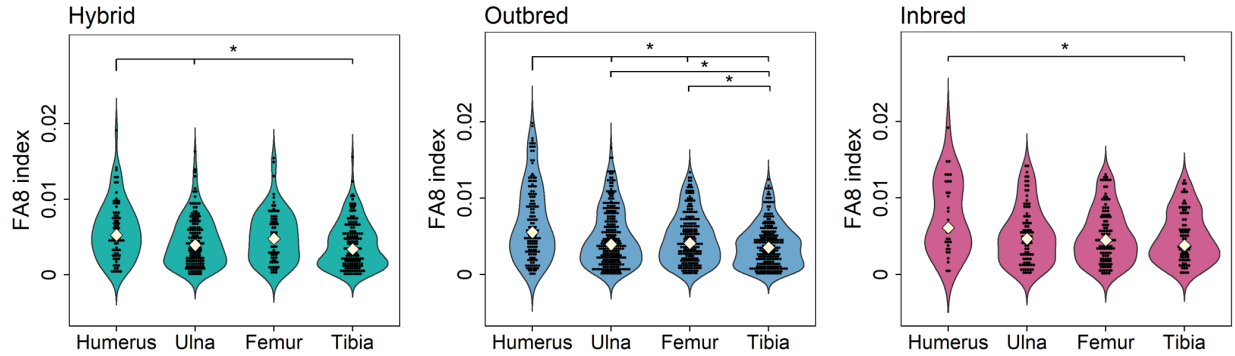
<sup>1</sup>In the first three hybrid subgroups only ulna and tibia were included, therefore, Wilcoxon rank sum test was performed in comparison between these two bones. W – test statistics for the Mann-Whitney U test.  $\chi^2$  – test statistics for the Kruskal-Wallis test. df – degree of freedom. p - value < 0.05

\*Significant after Bonferroni correction: KH: humerus – femur (p = 0.007), humerus – tibia (p = 0.039); CB: humerus – tibia (p = 0.002), marginal in humerus – ulna (p = 0.057); MC: humerus – ulna (p = 0.040), humerus – tibia (p = 0.025); AH: ulna – tibia (p = 0.036); CD1: marginal between ulna – femur (p = 0.056); hybrids pooled to one group: humerus – ulna (p = 0.022), humerus – tibia (p = 0.002); outbred pooled to one group: humerus – ulna (p = 5.20 x 10<sup>-4</sup>), humerus – femur (p = 1.18 x 10<sup>-3</sup>), humerus – tibia (p = 8.40 x 10<sup>-9</sup>), ulna – tibia (p = 0.043), femur – tibia (p = 0.019); inbred pooled to one group: humerus – tibia (p = 0.043)

Table 1.3. (continued):

and marginal between humerus – femur ( $p = 0.071$ ). All data pooled together: humerus – ulna ( $p = 9.30 \times 10^{-7}$ ), humerus – femur ( $p = 6 \times 10^{-5}$ ), humerus – tibia ( $p = 9.60 \times 10^{-13}$ ), ulna – tibia ( $p = 0.039$ ), femur – tibia ( $p = 0.002$ ).

\*\* All data used per bone (i. e. complete data for each bone pooled from all three groups).



**Figure 1.4.** Difference in FA levels (FA8 index) between four bones in hybrid (subgroups pooled), outbred and inbred group (populations / strains pooled). Significant differences between bones are marked with an asterisk. Longer line bar shows significance between corresponding bone with other bones marked with a shorter line bar. The yellow diamond shows the median FA value per bone in each group.

#### *Differences between populations / strains and groups for the same bone*

Another comparison included differences in median FA values between populations / strains and groups for the same trait (Table 1.4), these tests were based on scaled FA values  $|\ln(R/L)|$ . Significant difference was detected in femur between outbred populations and ulna between PWD and C57BL/6J strains when observed within the group. Differences in outbred populations were not significant after Bonferroni correction (Figure 1.5). When data were compared between all tested populations and strains, with hybrids pooled to one group (155 individuals), a significant difference remained in ulna, although nothing was significant after Bonferroni correction (Table 1.4). Comparison between the groups with pooled populations / strains per group found significant difference only in tibia between hybrid, outbred and inbred groups. After Bonferroni correction significance was found between outbred and inbred group (Figure 1.6).

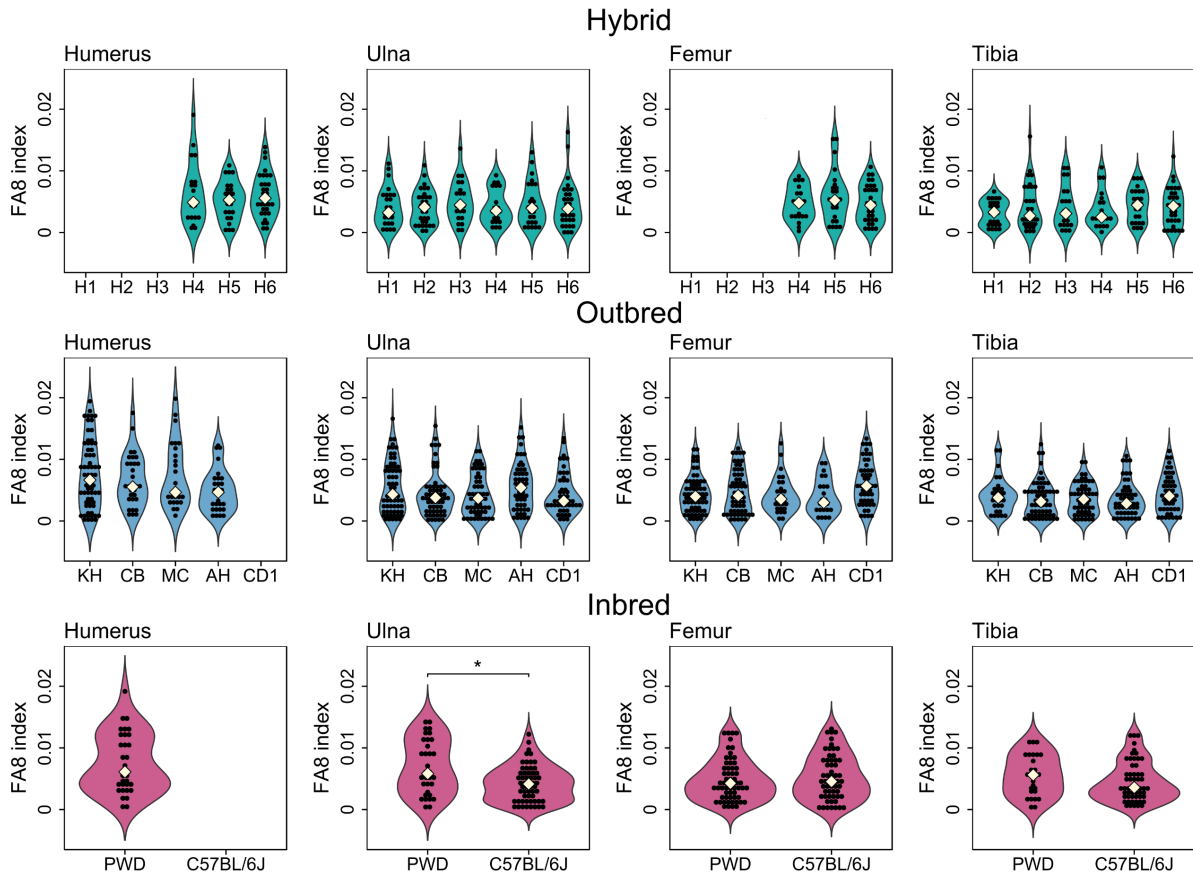
**Table 1.4.** Comparison of differences in fluctuating asymmetry levels between populations / strains and groups for the same bone with Kruskal-Wallis test and Mann-Whitney U test (when only two populations were included).

Group	Bone	n	Kruskal-Wallis rank sum test on $ \ln(R/L) $		
			$\chi^2$	df	p
Hybrid	Humerus	77	0.585	2	0.746
	Ulna	155	1.573	5	0.905
	Femur	77	0.669	2	0.716
	Tibia	155	2.788	5	0.733
Outbred	Humerus	152	4.610	3	0.203
	Ulna	281	5.299	4	0.258
	Femur	225	9.576		<b>0.048</b>
	Tibia	246	4.486		0.344
			<b>W</b>		<b>p</b>
Inbred <sup>1</sup>	Ulna	90	609		<b>0.006</b>
	Femur	117	1809		0.591
	Tibia	84	622		0.159
<b>Comparison per population / strain and subgroup (hybrid subgroups separated)</b>					
Group	Bone	n	$\chi^2$	df	p
Hybrid - Outbred - Inbred	Humerus	262	7.602	7	0.369
	Ulna	526	16.592	12	0.166
	Femur	419	11.438	9	0.247
	Tibia	485	15.245	12	0.228
<b>Comparison per population / strain and group (hybrid subgroups pooled)</b>					
Group	Bone	n	$\chi^2$	df	p
Hybrid - Outbred - Inbred	Humerus	262	7.151	5	0.210
	Ulna	526	15.261	7	<b>0.033</b>
	Femur	419	10.660		0.154
	Tibia	485	12.353		0.090
<b>Comparison per group</b>					
Group	Bone	n	$\chi^2$	df	p
Hybrid - Outbred - Inbred	Humerus	262	2.087	2	0.352
	Ulna	526	2.293		0.318
	Femur	419	0.736		0.692
	Tibia	485	6.205		<b>0.045*</b>

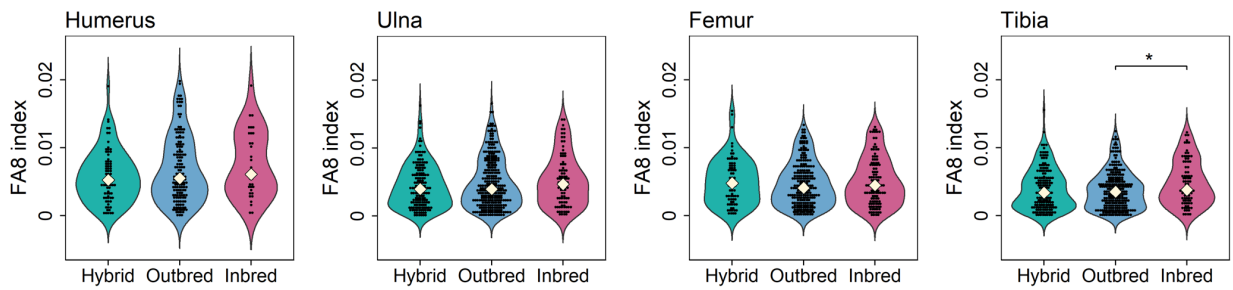
<sup>1</sup>In inbred group, only PWD males have included humerus.

$\chi^2$  – test statistics for the Kruskal-Wallis test. *df* – degree of freedom. *W* – test statistics for the Mann-Whitney U test. p - value < 0.05

\* Significant after Bonferroni correction was found in tibia between outbred and inbred groups (p = 0.045).



**Figure 1.5.** Difference in FA levels (FA8 index) between populations / strains per bone in hybrid, outbred and inbred groups. Significant differences between bones are marked with an asterisk. The yellow diamond shows the median FA value per bone in each hybrid subgroup and population / strain.



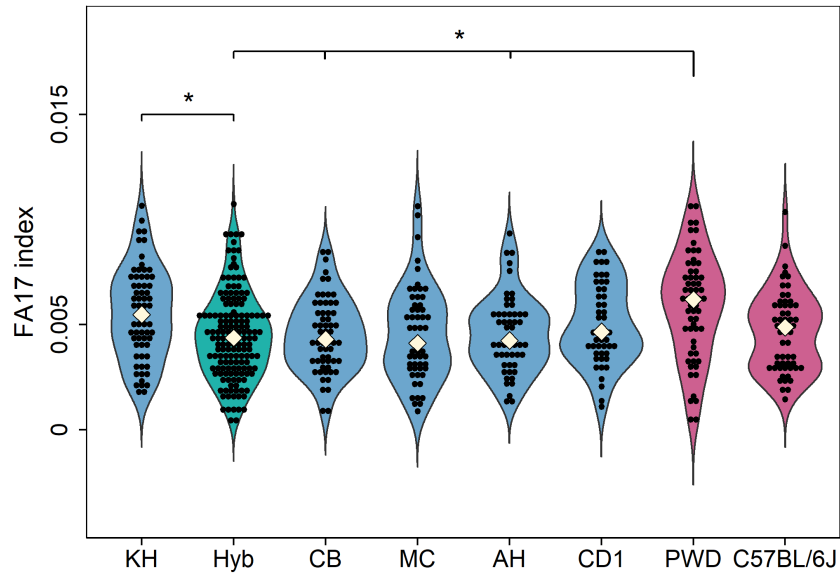
**Figure 1.6.** Difference in FA levels (FA8 index) between groups per bone in hybrid, outbred and inbred groups. Significant differences between bones are marked with an asterisk. The yellow diamond shows the median FA value per bone in each group.

Additional analyses were done with unscaled FA values (Table S1.13), which yielded similar results as with scaled, with significant differences in femur of an outbred group, between KH and CD1, as well as close to marginal between MC and CD1. This difference in femur bone was further reflected between KH and CD1 when all data are tested together, with hybrids being pooled in one group.

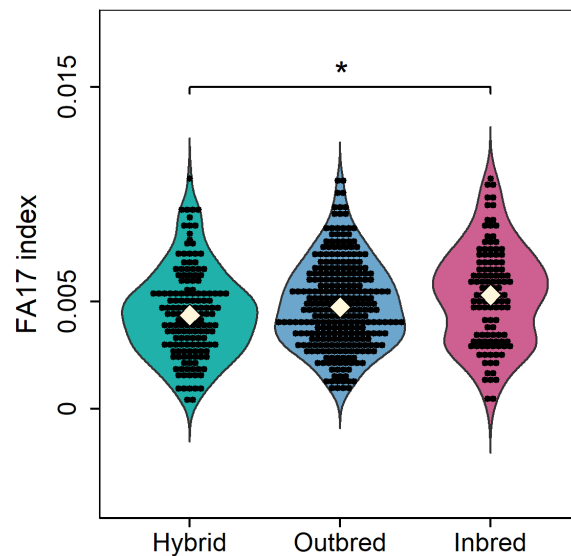
Moreover, I tested data (FA8 values) with all males and females included (Table S1.14) and with non-filtered data (i. e. all individuals before exclusion for larger differences between two sides) (Table S1.15). Interestingly, when all males and females were included in the analysis, the same pattern was found as in the main analysis, with only higher significance values, particularly in PWD, and additional significance in the first two hybrid subgroups (between all four bones). A similar outcome was found in comparison between populations and groups for the same bone, with additional significance in tibia between two inbred strains, then in tibia between different populations and ulna in comparison between the groups. Similarities were found even with the non-filtered data set, although these differences were higher and probably biased with other possible types of asymmetry.

Organism-wide developmental instability was examined with the composite index (FA17). This was calculated based on the number of traits that were available per individual (e.g. among hybrids, DI in the first three subgroups was estimated only based on ulna and tibia, while other three groups included information from all four bones). Differences based on the FA17 index were detected with the Kruskal-Wallis rank sum test ( $\chi^2 = 28.346$ ,  $df = 7$ ,  $p = 1.90 \times 10^{-4}$ ). After pairwise comparisons using Wilcoxon rank sum test with Bonferroni correction, significant values were found between KH to hybrid group (pooled); PWD to CB, AH and hybrid group (pooled) (Figure 1.7). Marginal significance was detected between PWD to MC and C57BL/6J. Comparison between the groups was found significant as well ( $\chi^2 = 9.465$ ,  $df = 2$ ,  $p = 0.009$ ), although significance was found only between inbred and hybrid group (Figure 1.8). Indices per populations and group are shown in Tables 1.5 and S1.16.





**Figure 1.7.** Differences in organism-wide DI (FA17 index) between populations in outbred: *M. m. musculus* (KH), hybrid group (with hybrid subgroups pooled), *M. m. domesticus* (CB, MC, AH, CD1) and two inbred strains (PWD, C57BL/6J). Significant differences between populations are marked with an asterisk. The longer line bar shows significance between corresponding population with other populations. The yellow diamond shows the median FA17 value for each population.



**Figure 1.8.** Differences in organism-wide DI (FA17 index) between hybrid group (with hybrid subgroups pooled), outbred and inbred groups (with populations pooled per group). Significant differences between groups are marked with an asterisk. The yellow diamond shows the median FA17 value for each group.

**Table 1.5.** FA1 and FA8 indices for each bone and composite measure of FA (FA17) per population / strain (a) and group (b).

(a)

Group	Pop / Strain	Bone	n	FA1		FA8		FA17	
				Mean	Median	Mean	Median	Mean	Median
Outbred	KH	Humerus	65	0.0806	0.0712	0.0076	0.0067	0.0055	0.0055
		Ulna	65	0.0721	0.0564	0.0054	0.0043		
		Femur	65	0.0627	0.0563	0.0043	0.0040		
		Tibia	30	0.0671	0.0606	0.0042	0.0038		
	CB	Humerus	31	0.0771	0.0655	0.0066	0.0055	0.0045	0.0043
		Ulna	61	0.0639	0.0544	0.0044	0.0038		
		Femur	61	0.0746	0.0625	0.0047	0.0041		
		Tibia	61	0.0615	0.0553	0.0035	0.0031		
	MC	Humerus	27	0.0858	0.0573	0.0073	0.0047	0.0046	0.0041
		Ulna	55	0.0594	0.0525	0.0043	0.0037		
		Femur	28	0.0638	0.0545	0.0042	0.0035		
		Tibia	55	0.0631	0.0596	0.0037	0.0035		
	AH	Humerus	29	0.0620	0.0555	0.0049	0.0047	0.0046	0.0042
		Ulna	52	0.0850	0.0838	0.0056	0.0054		
		Femur	23	0.0672	0.0530	0.0039	0.0030		
		Tibia	52	0.0686	0.0541	0.0036	0.0028		
CD1	Ulna	48	0.0669	0.0483	0.0045	0.0032	0.0050	0.0047	
	Femur	48	0.0988	0.0933	0.0061	0.0057			
	Tibia	48	0.0811	0.0727	0.0044	0.0041			
Inbred	PWD	Humerus	33	0.0827	0.0672	0.0075	0.0061	0.0059	0.0062
		Ulna	33	0.0878	0.0748	0.0069	0.0058		
		Femur	60	0.0704	0.0605	0.0048	0.0043		
		Tibia	27	0.0847	0.0849	0.0055	0.0056		
	C57BL/6J	Ulna	57	0.0593	0.0567	0.0042	0.0041	0.0047	0.0049
		Femur	57	0.0844	0.0714	0.0053	0.0045		
		Tibia	57	0.0804	0.0633	0.0045	0.0036		

(b)

Group	Bone	n	FA1		FA8		FA17	
			Mean	Median	Mean	Median	Mean	Median
Hybrid	Humerus	77	0.0667	0.0622	0.0059	0.0053	0.0045	0.0044
	Ulna	155	0.0596	0.0534	0.0044	0.0039		
	Femur	77	0.0754	0.0727	0.0051	0.0048		
	Tibia	155	0.0663	0.0572	0.0040	0.0034		
Outbred	Humerus	152	0.0772	0.0629	0.0068	0.0055	0.0049	0.0047
	Ulna	281	0.0693	0.0566	0.0049	0.0039		
	Femur	225	0.0742	0.0636	0.0048	0.0041		
	Tibia	246	0.0678	0.0600	0.0038	0.0035		
Inbred	Humerus	33	0.0827	0.0672	0.0075	0.0061	0.0053	0.0053
	Ulna	90	0.0698	0.0626	0.0052	0.0047		
	Femur	117	0.0772	0.0674	0.0051	0.0045		
	Tibia	84	0.0818	0.0669	0.0048	0.0038		
All data together*	Humerus	262	0.0748	0.0631	0.0066	0.0055	0.0048	0.0047
	Ulna	526	0.0665	0.0561	0.0048	0.0040		
	Femur	419	0.0753	0.0671	0.0049	0.0044		
	Tibia	485	0.0697	0.0600	0.0041	0.0036		

*Table 1.5. (continued): Group* – defines a group to which a population / strain belongs (a) and data for pooled populations / strains (b). *Population / Strain* – defines used population / strain. *Bone* – limb bone used in calculation. *n* – number of individuals used per bone per sample. Mean and median values are shown per index in each bone.

\* All data used per bone (i. e. complete data for each bone pooled from all three groups).

## DISCUSSION

### *Overview of differences in bone length*

Forelimb bones are shorter than the hindlimb bones, with ulna being longer in the forelimb and tibia in the hindlimb. Among outbred populations, AH and CD1 have the longest and KH the shortest bones. Mice from the hybrid zone show intermediate length between the two subspecies. The inbred group shows differences between two strains, with bones from PWD being smaller than C57BL/6J. Standard deviations are usually larger in hindlimb bones, except in inbred strains (PWD and C57BL/6J males and females), which showed similar variances between bones, with the smallest values in C57BL/6J; this pattern was reflected in the coefficient of variation as well. When populations were pooled for sex, the smallest standard deviations could still be seen in C57BL/6J. Comparison among groups found the longest bones in an outbred group, while the inbred group showed shorter forelimbs, but longer hindlimbs than the hybrid group. Variances in the bone length were on average smaller in the hybrid group in comparison to the other two groups (Table S1.10).

### *Differences in FA level*

Larger differences between sides were found in forelimbs, with humerus showing on average larger deviations. Among outbred populations, KH showed the highest FA values for humerus. These were significant in comparison to femur and tibia, although non-significant when compared to other outbred populations for the same bone. CB and MC populations showed similar differences, while AH population differed between ulna and tibia. In CD1 animals, data for humerus were excluded beforehand due to indication toward other types of asymmetry (Figure S1.12). Inbred strains and hybrid subgroups did not show significant differences between the bones (Figures S1.13, S1.14). However, ulna in PWD females found departure from normality and tibia in PWD males showed platykurtosis, while humerus in C57BL/6J population showed indication towards DA (Tables S1.4, S1.5). Therefore, these data were not included in

the FA analyses. In the hybrid group, larger deviations in asymmetry were found in humerus and femur of the first two groups (Tables S1.4b, S1.5b). Similarities in FA indices occurred with humerus of these hybrid subgroups which are closer to *M. m. musculus* and also showed larger deviation. Median FA8 values in for humerus in the first three hybrid groups (0.0072, 0.0068, 0.0079) (Table S1.16b), KH (*M. m. musculus*) (0.0067) and PWD (wild-derived strain with *M. m. musculus* origin) (0.0061) (Table 1.5a) were higher than in other populations / strains and subgroups for this bone. Mean FA8 values were higher than median and followed described differences in these populations, with also including humerus of MC (0.0073). On the other hand, more individuals were included for humerus in KH (65) in comparison to other populations, however, the pattern remained when data were examined with all males and females included; all hybrids pooled (155 ind., median FA8 = 0.0064), PWD (60 ind., median FA8 = 0.0072) (Table S1.14c) and even with non-filtered data (Table S1.15c). Differences in FA were compared in each group, humerus confirmed more deviation than other bones, in the hybrid group differed from ulna and tibia, while in inbred group only from tibia. Nevertheless, the outbred group showed significant differences among humerus and all other bones, additional finding was in tibia which differed from ulna and femur (Figure 1.4). Further, data from all groups were pooled per bone, which showed the same pattern between bones as in an outbred group (Figure S1.15).

Results for humerus could be biased due to smaller bone size, although the measurement error was comparable to other bones (Table S1.2). Another explanation could be the influence of other types of asymmetry in this bone and hence, it should be completely excluded from FA observation, although preliminary analyses did not find strong DA even when all males and females per population were included (tested as a comparison of raw mean asymmetry (R-L) to a FA4a index), but presence of platykurtosis, or tendency toward bimodality (Van Valen 1962; Palmer & Strobeck 2003) was obvious in some groups (e.g. in hybrid group when all data for humerus were included, in CD1 and PWD with all males and females pooled). Sample size, which was smaller for humerus was also investigated with inclusion of all males and females after removing possible outliers based on larger differences between sides, however, final results remained unchanged. Another bone with platykurtosis was tibia in PWD males, although, final examination in all four bones (i. e. when all data per bone were pooled) found this type of kurtosis in femur bone.

Observation with different populations / strains and hybrid subgroups per bone did not yield major differences, except one comparison in the inbred group where significance existed between two strains

for ulna. Groups did not differ for the same bone as well, only significance was found in tibia between outbred and inbred groups.

Composite index based on multiple traits per individual showed the PWD strain as being statistically different from CB, AH populations and pooled hybrid group, while KH differed from the pooled hybrid group. The lowest median FA17 index is reported in MC (0.0041) and the highest in PWD (0.0062) (Table 1.5a). Mean values for the FA17 index were similar, with the lowest in CB (0.0045) and highest in PWD (0.0059). Among groups, developmental instability was higher in inbred animals (0.0053) and lower in the hybrid group (0.0044) (Table 1.5b), which is reflected in significant differences between them.

Overall, my result in limb bones for the hybrid group found an overlap with the previous study of craniofacial shape transition across the house mouse hybrid zone (Pallares et al. 2016), which confirmed that the level of fluctuating asymmetry was not affected by the degree of hybridization. However, mean FA values in the hybrid group were on average smaller in comparison to the outbred group which consists of pure subspecies, as also noticed in organism-wide DI. This was comparable with Alibert and Renaud (1994) who showed decrease in FA in lower molars of the hybrids between *M. m. musculus* and *M. m. domesticus*. Although development of dental characters (in their study) was shown to be more stable in these hybrids, an opposite result was found for reproductive fitness in hybrids (Vanlerberghe et al. 1986, 1988; Tucker et al. 1992; Dod et al. 1993), also higher parasite load was described in *M. musculus* hybrids (mentioned in Alibert & Renaud 1994). Continuation of the study from Alibert and Renaud (1994) in laboratory hybrids between *M. m. musculus* and *M. m. domesticus* did not show larger differences in FA levels among F1, F2 and backcross hybrids, hence recombination between the genomes did not influence major perturbations in development, which is reflected in preserved heterosis in the backcross and F2 hybrids for dental characters (Alibert et al. 1997). Additional observation between laboratory and wild samples revealed lower FA levels in laboratory *M. m. musculus* and *M. m. domesticus* individuals in comparison to wild house mice from the hybrid zone in Denmark, although this difference was only slight (Alibert et al. 1997). In addition, F1 hybrid sterility did not show any connection with the level of FA (Alibert et al. 1997), which was consistent with the study on wild populations (Alibert et al. 1994).

Lower level of FA in the hybrid group might be also influenced by age, with mice from the hybrid zone being much younger than all other populations (mean age in the hybrid group ~ 79 days old). Impact of age cannot be completely ruled out because I did not have any hybrid individual at later age, nor younger individuals from the other two groups (outbred and inbred) were included. Preliminary analyses for FA differences in age did not show correlation with this factor in the hybrid group. Comparison with other

populations from an outbred group also did not show significant differences, hence age influence was not considered as a bias in the analysis.

Less stability is found in the inbred group, with significantly greater deviation from the hybrid group. However, this conclusion is mostly biased due to PWD. On the other hand, similar results were reported in inbred strains of mice, with more FA in osteometric traits than the hybrids (Leamy 1984, 1992). Carter et al. (2009) observed a connection between genetic stress (inbreeding) and the values of FA in the wing of two outbred unrelated laboratory populations of *Drosophila melanogaster*. Higher levels of FA were found in inbred lines, whereby he confirmed previous studies in more homozygous populations (Waldmann 1999; Schaefer et al. 2006). Moreover, inbreeding may express various effects in different populations (Lens et al. 2000).

Among other examples reported in the literature, reduced developmental stability, measured as increased FA or higher frequency of developmental abnormalities was described in different organisms, such as birds (unpublished results cited in Graham 1992), fish (Graham & Felley 1985; Leary et al. 1985), frogs (Szymura & Barton 1986), insects (Ross & Robertson 1990) and plants (Manley & Ledig 1979). Example of inter-genera hybrids between bison (*Bison bonasis*) and domestic cattle (*Bos taurus*) (unpublished results cited in Graham 1992), reported an effect on reproduction but not on developmental stability. Higher stability was also reported for several characters in hybrids between two subspecies of sagebrush in comparison to their parents (Freeman et al. 1995).

#### *Studies of FA in limb bones*

Hallgrímsson et al. (2003) investigated changes in the level of FA in limb skeletal structures of CD1 mice during prenatal development, which resulted in FA differences based on the timing of developmental events. The study found decreased FA in size during fetal growth (between 14 and 17.5 days old), an opposite result was shown in skeletal structures of the three primate species during postnatal growth (Hallgrímsson 1993, 1998, 1999).

In humans, upper limb bones express higher asymmetry, mostly towards the right side, whereas lower limb bones are more symmetrical (reviewed in Auerbach & Ruff 2006). Across human populations, directional asymmetry is prominent in diaphysis, then in articular dimensions, while bone lengths show the lowest level (Auerbach & Ruff 2006; Reeves et al. 2016). Similar results were shown in a study conducted in two Dutch archaeological populations, greater asymmetry was found in midshaft diameter

of the long bones, while lengths of the bones showed lower indices, although the highest value was reported for femur bone (Hagg 2016).

Due to quadrupedal locomotion and in general greater limb integration in the cotton-top tamarin (*Saguinus oedipus*), differences in FA between the upper and lower limbs are expected to be reduced or even absent. Result of this study found similarity in a pattern of asymmetry between tamarin long bones and those in humans (Auerbach & Ruff 2006). In contrast to humans, higher level of DA was detected in the lower limb in tamarins (Reeves et al. 2016). The reported results might be biased due to the choice of the traits that were used in the analysis of asymmetry, i. e. measurements obtained from the limb bones may differ between two studies and include different parts of the bone, e.g. width, length, articular surfaces etc.

In our study, differences could not be detected when populations and groups were compared for the same bone, which supports difficulties in estimation FA based from only one trait. Further, comparison between bones found different pattern in these groups, with most of the differences in an outbred group.

Many difficulties obscure analyses of FA, nevertheless these questions are still of great importance for developmental and evolutionary studies, especially associations between FA/DI with fitness and inbreeding. Attention should be given also to observation of different traits, which can improve current methods and increase the reliability of their choice as a measure of DI (Dongen 2006).

## **CONCLUSIONS**

My results suggest a potential benefit of heterozygosity in the hybrid group, which was seen in lower asymmetry, whereas the larger deviation observed in the inbred group might be a consequence of higher homozygosity in these individuals. However, only one of the two inbred strains used in this study showed major differences, hence investigation with more inbred strains is needed for stronger support of this hypothesis. Further, for the hybrid group I did not find larger differences between subgroups, in both per bone and among four bone analysis. However, I could not completely compare the patterns between the center and the edges of the hybrid zone, due to the lack of samples and pooling the second half of the data to only two groups. Still, one subgroup (H5) could represent the center, but the results in this group were similar to the last one (H6). Environmental influence is also very reduced, since these were lab-bred

mice, which may explain lower possibility to detect differences between the sides, a similar outcome was mentioned in Alibert et al. (1997).

Comparisons between the bones when all data were pooled detected significant FA differences between four bones, although, these were mostly influenced by the outbred group. The bone with the highest mean FA value is found in humerus, then femur which was very similar in FA to ulna, and the lowest value in tibia. Limb bone lengths were found in previous studies to show lower level of asymmetry than measurements of the middle part of the long bones (Auerbach & Ruff 2006; Hagg 2016; Reeves et al. 2016). Further, differences in size are usually smaller than the ones in shape, as reported for the skull (Debat et al. 2000).



## CHAPTER II

### Morphological integration in *Mus musculus* limb elements

#### INTRODUCTION

Morphological integration is influenced by complex processes of covariation among traits (Olson & Miller 1958). This field of study is oriented towards understanding the principles of cohesion and coordination among different parts of the body (Klingenberg 2008). One of the main interests of integration involves the strength of covariation, which is inferred from the distribution of variation across phenotypic dimensions (Wagner 1984; Cheverud et al. 1989). Another question considers the pattern of covariation, observed in common coordinated changes of different parts of the structure (Klingenberg 2008). Integration is closely related to modularity which refers to separation of developmental systems into modules, i. e. components that share stronger internal connections, rather than between other parts of the system (Klingenberg 2008; Hallgrímsson et al. 2009).

Different patterns of integration are investigated in the limb bones of mammals due to the shared developmental pathways, which further induce higher correlation between these structures (Young & Hallgrímsson 2005). Developmental interactions between traits can be explained through covariation, because developmental pathways include many processes which require coordinated variation between them in order to produce a morphological trait (Klingenberg 2008). Precursors of limb bones in tetrapods are formed in consecutive processes of mesenchymal cell proliferation and division and represent a much studied example of the fore mentioned mechanisms (Mariani & Martin 2003). A hierarchical model based on shared developmental factors among limbs explains the structured covariation across the overall body size, between limbs, within limbs, between homologous elements and within elements (Hallgrímsson et al. 2002).

Integration can arise due to genetic effects on multiple traits, i.e. pleiotropy. The impact of genes on different traits can be transmitted through developmental pathways, in direct interactions, or in parallel influence of genes on separate pathways (Klingenberg 2008). Genetic covariance structure can be produced by combined pleiotropic gene effects, which further imposes limitation in responses to selection and direct the evolution of phenotypic traits (Lande 1979; Cheverud 1984; Schluter 1996). However, variation in gene effects on multiple traits could exist due to epistasis (Wright 1968; Mayr 1970). This was

explained in a model of the evolution of pleiotropy by epistatic interactions (selection–pleiotropy–compensation (SPC model)) (Pavlicev & Wagner 2012). Pavlicev et al. (2013) investigated individuation between the fore- and the hindlimbs in inbred mouse strains and found that the genetic basis of traits is mainly influenced by specific genetic interactions, rather than individual genes.

Evolution of morphological traits might depend on functional and developmental connections between traits which impact covariance structure (Cheverud 1982, 1984, 1995). Due to similarity between genetic and phenotypic correlations, phenotypic covariances can be included in revealing the genetic architecture of traits (Cheverud 1988). Effects of stabilizing selection are also important in patterning covariance structures (Lande 1980; Cheverud 1984) and directing the evolution of correlated traits.

Studies of the adult limb skeleton reported covariation within limbs, between homologous elements and between fore- and hindlimbs, which is influenced by shared genetic factors (Hallgrímsson et al. 2002; Young & Hallgrímsson 2005; Reno et al. 2008). Further, epigenetic effects on integration of the limbs in the laboratory house mouse was investigated in Young et al. (2009). Interestingly, a major impact of locomotory activity on fore- and hindlimb lengths was not detected, while body mass showed influence on bone length in limbs (Young et al. 2009). Previously reported examples of integration found different patterns, such as higher within-limb integration in adult (random-bred) house mouse (Leamy 1977). This was further confirmed in hens (Van Valen 1965). Study in strepsirrhine primate limbs reported similar covariation in both, homologous elements and within limbs (Villmoare et al. 2011).

Integration patterns in nonquadrupedal species are expected to yield reduced covariation in their limbs due to different functions, this was shown mainly in flying vertebrates, such as bats (Young & Hallgrímsson 2005), birds and pterosaurs (Bell et al. 2011). Divergence in limb bones in rodents is found in the lesser Egyptian jerboa (*Jaculus jaculus*) which have bipedal locomotion (Cooper 2011). Although it has a common ancestor with a laboratory mouse, differences are described in limb morphologies, as well as in various locomotory functions (Moore et al. 2015).

The main goal of this chapter is the exploration of the relationship between traits and integration patterns in outbred mouse populations and two inbred strains. Further, mice from the hybrid zone between the two subspecies may serve as a potential for revealing differences in more natural populations. Comparison between the hybrid, outbred and inbred group provide more information about general patterns of variability among these groups which might be caused by differences in their genetic architecture.

Due to quadrupedal locomotion and shared function, developmental and genetic network, we would assume to find high covariation between limb bones in the house mouse. Following this expectation, homologous elements as well as bones from the same segment would show better correlations. Different evolutionary histories of the outbred, inbred and hybrids might alter the level of integration. Inbred mice are expected to show greater differences in integration because of the reduced genetic variation and thus higher tendency toward changes in covariance structure which might be influenced by a single mutation (Hallgrímsson et al. 2009). In contrast, covariation structure in natural populations was found to be more stable in previous studies (Steppan 1997; Ackermann & Cheverud 2000; Marroig & Cheverud 2001; Jamniczky & Hallgrímsson 2009).

## METHODS

### *Mouse samples*

Populations from two outbred mouse subspecies, two inbred strains and the first generation of mice from the hybrid zone are described in Chapter I. The same sample size per population and group was used for an outbred and inbred group as in the previous chapter. Further, the same individuals with larger differences between the right and the left side were excluded beforehand for these two groups. Sample sizes with pooled sexes were: KH (65), CB (61), MC (55), AH (52), CD1 (48), PWD (60), C57BL/6J (57). An exception was made with the mice from the hybrid zone, where all available male mice were included in the study (197 individuals). Therefore, division of the hybrid zone data set based on the percentage of *M. m. domesticus* alleles found in each individual has changed compared to the first chapter by including one more group (H7); H1 (8.3 – 18.3%), H2 (18.31 – 24%), H3 (24.01 – 28.3%), H4 (28.31 – 38.3%), H5 (38.31 – 58.3 %), H6 (58.31 – 69.3%) and H7 (69.31 – 99.3%). Four bones, humerus, ulna, femur and tibia were used as well in this chapter.

In this part of the study, the average measurement between the right and left side of the same bone was used, as the variation within a specimen, reflected as the asymmetry between two sides was not investigated in this chapter. However, data from an outbred and inbred group were restricted to the number of individuals which did not show large differences between the two sides per bone in each group (Supplementary List 1.1 and Figures S1.1 and S1.2). Additional measurements were provided for the mice from the hybrid zone by including the individuals with either a complete right (N=18) or a complete left

(N=23) side, i.e. only data from one side that was available were included in the analysis. The number of individuals from the hybrid zone with both sides averaged counted 156 individuals (note that I have included one more individual with both sides that was excluded in the previous chapter). Partially damaged bones occurred in some individuals due to previous handling of the samples, therefore, all bones from the affected side for the individuals with only one measurement per side were excluded from the analysis.

Additional comparison was conducted with all available individuals from the outbred and inbred groups, regardless the differences between the sides and results are shown in the supplementary material.

### *Statistical analysis*

In order to pool data per population and strain (for outbred and inbred groups), each separate population and strain was first corrected for the influence of sex with a linear model, where the length of each bone was regressed over sex and residual values were used for further analysis. The influence of sex per bone in each population and strain is shown in Table S2.1 The sample of mice from the hybrid zone was corrected for the variation due to age, also with a linear model, where all individuals per bone were first regressed over age and residual values for each individual were subsequently used for the division to hybrid subgroups (H1 – H7). Further, data pooled to an outbred / inbred group were corrected for the variation due to sex and population / strain respectively. Comparison between mice from the hybrid zone, outbred and inbred groups were performed after all three groups were pooled together and regressed over age, sex and population / strain factors. Residual values after this correction were further divided to the hybrid, outbred and inbred group. Normality of the distribution for each bone length residual values per population / strain, hybrid subgroup, as well as in pooled data to the hybrid, outbred and inbred group was assessed with the inspection of Q-Q-plots (Figures S2.1 – S2.3). Pearson correlation and covariance matrices were computed from residual values of each respective data set. All analyses were performed in R version 3.2.5 (R Core Team 2016).

### *Matrix repeatabilities*

The error of the matrix estimation should be considered before further analysis through calculation of the repeatability of the matrix. This approach is based on sampling the individuals with replacement from the same population and calculation of a covariance or correlation matrix (Houle & Meyer 2015). The estimate

of the repeatability is provided as the mean value of the correlation between the random sample matrix and original estimated matrix, i. e. the observed correlation is corrected by the maximum value of the correlation between matrices from the same data set (Melo et al. 2015). The value of one would mean high similarity between the observed and corrected matrices. Repeatabilities of matrices were performed with data from residual values using a Mantel test with 10,000 replicates for correlation matrices, while covariance matrices were examined with the random skewers method and the same number of replicates. The random skewers method is more appropriate for comparison of covariance matrices, which are measured on a different scale from correlation matrices and therefore, not appropriate for randomization tests used in correlation matrix similarity (Cheverud 1989). The random skewers method is based on a series of random selection vectors which are multiplied with two matrices, and the correlation of response vectors to the same selection vector between two matrices is compared (Cheverud 1996b; Cheverud & Marroig 2007; Melo et al. 2015).

#### *Matrix correlations*

Comparison of the overall correlation structure between populations, strains and hybrid subgroups were performed with Mantel's test, while similarities in covariances matrices were assessed with the random skewers method. Following Marroig and Cheverud (2001), Young and Hallgrímsson (2005), similarities in correlation and covariance matrices were observed with adjusted values for the respective matrix, i. e. observed matrix correlations ( $R_{obs}$ ) were divided by an estimate of the maximum correlation ( $R_{max}$ ). With this approach, the impact of the sampling error on matrix correlations is taken into account with adjusted matrix correlations ( $R_{adj}$ ), which include the maximum correlation, estimated from repeatabilities of compared matrices (Marroig & Cheverud 2001). This procedure included first the calculation of the correlation matrix repeatability with Monte Carlo simulation (1000 replicates). Further, estimation in similarities of correlation matrices from resampled data was performed with Mantel's test (1000 replicates). This is a permutation test which compares the matrices by preserving one matrix constant while randomly permuting the second matrix. Significance of the test was considered when observed correlation exceeded 95% quantile of the permuted distribution (Bell et al. 2011; Kolarov et al. 2017; Melo et al. 2015). Matrix repeatabilities and correlations between matrices were performed with *EvoLQG* package in R (Melo et al. 2015).

### *Partial correlations*

Examination of the patterns of limb correlations was observed with partial correlations, which measure the strength of the relationship between two variables, while the effect of other variables in the correlation matrix is excluded. This approach has been used in previous studies of appendicular skeleton (Marroig & Cheverud 2001; Young & Hallgrímsson 2005; Bell et al. 2011; Kolarov et al. 2017). Correlation coefficients were computed as Pearson's product-moment coefficients and statistical significance was determined with a *t*-test implemented in the R package ppcor (Kim 2015). Significance of these correlation coefficients was also tested with edge exclusion deviance (EED).  $EED = -N \ln(1 - \rho_{ij \cdot \{k\}}^2)$ , where *N* reflects the sample size;  $\rho_{ij \cdot \{k\}}$  is the partial correlation of variables *i* and *j* with all other elements held constant (Magwene 2001). Conditional independence between two measures that are compared is considered when an EED value is less than 3.84 (this corresponds to  $p = 0.05$ ,  $df = 1$  from the  $\chi^2$  distribution) (Magwene 2001). Similarities in the overall structure of matrices from partial correlations were performed with Mantel's test of the observed matrix correlations, without the adjustment.

### *Index of integration*

The overall level of integration was estimated from correlation matrices, by using two common indices, the average coefficient of determination ( $r^2$ ) (Cheverud et al. 1989), which refers to the magnitude of integration and it is calculated as the average of the squared correlations (Shirai & Marroig 2010). Higher values for this would indicate greater integration, while lower values reflect an opposite trend. The second index widely used in studies of phenotypic integration is based on the dispersion of the eigenvalues from the correlation matrix (Wagner 1984; Cheverud et al. 1989). Eigenvalues depict the amount of variance associated with each corresponding eigenvector (Pavlicev et al. 2009), while eigenvectors represent axes along which linear transformation acts. Greater differences between eigenvalues are shown when the variation is concentrated in a few dimensions which is further interpreted as higher integration. This pattern is confirmed in traits with higher correlation, while lower variance of eigenvalues is detected in uncorrelated traits (Pavlicev et al. 2009). The average squared deviation of the eigenvalues from the mean eigenvalue defines eigenvalue variance (Pavlicev et al. 2009). The mean eigenvalue of a correlation matrix is calculated from the trace of the correlation matrix, thus it equals one.

$Var(\lambda) = \frac{\sum_{i=1}^N (\lambda_i - 1)^2}{N}$ , will be denoted as EV in further explanations.

Further, this measure can be divided by the maximum eigenvalue variance for the corresponding number of traits which is referred as the relative eigenvalue variance and ranges from zero to one (Pavlicev et al. 2009).

$\text{Var}_{\text{rel}}(\lambda) = \frac{\text{Var}(\lambda)}{N-1}$ , will be denoted as  $\text{EV}_{\text{rel}}$ .

Shirai and Marroig (2010) showed an overlap between coefficient of determination ( $r^2$ ) and index from Pavlicev (2009), which is also seen in our study.

Differences between populations / strains, hybrid subgroups, as well as general comparison between groups were calculated from resampled data as the number of times a population / strain with lower eigenvalue variance (EV) exceeds other population eigenvalue variance (EV), divided by the number of iterations (replicates) (10,000) (Young & Hallgrímsson 2005; Kolarov et al. 2017). Significant differences were observed at  $p < 0.05$ . To control for the sample size across different populations / strains and groups, each bootstrap (with replacement) procedure included the equal number of random individuals from the whole sample per respective group.

The index of integration from the variance-covariance matrix should account for the scale-dependence in the data, reflected as larger variances and covariances in larger individuals. The index is based on the coefficient of variation of the eigenvalues (ICV) (Wagner 1984) and it is computed as the standard deviation of eigenvalues divided by the mean eigenvalue (Shirai & Marroig 2010).

$$\text{ICV} = \frac{\sigma(\lambda)}{\bar{\lambda}}$$

#### *Correction for variation due to size*

Effects of general size variation have been shown as an important factor in studies of morphological integration (Marroig et al. 2009; Porto et al. 2013). Therefore, to account for the influence of size in each data set, I followed the approach from Marroig et al. (2004) and Shirai & Marroig (2010), where residual matrices were calculated after the effect of size was removed from the corresponding correlation and covariance matrix. Previous studies considered PC1 obtained from covariance and correlation matrices of analyzed traits as potential factor of size variation. This was concluded when the examined characters are correlated with general size and the specimens differ in overall body size variation (Jolicoeur 1963). Therefore, principal component analysis was used to extract the first eigenvector from the matrix and for further calculation of the size related factor. The residual matrix was computed as  $R = P-V'V$ ; R stands for

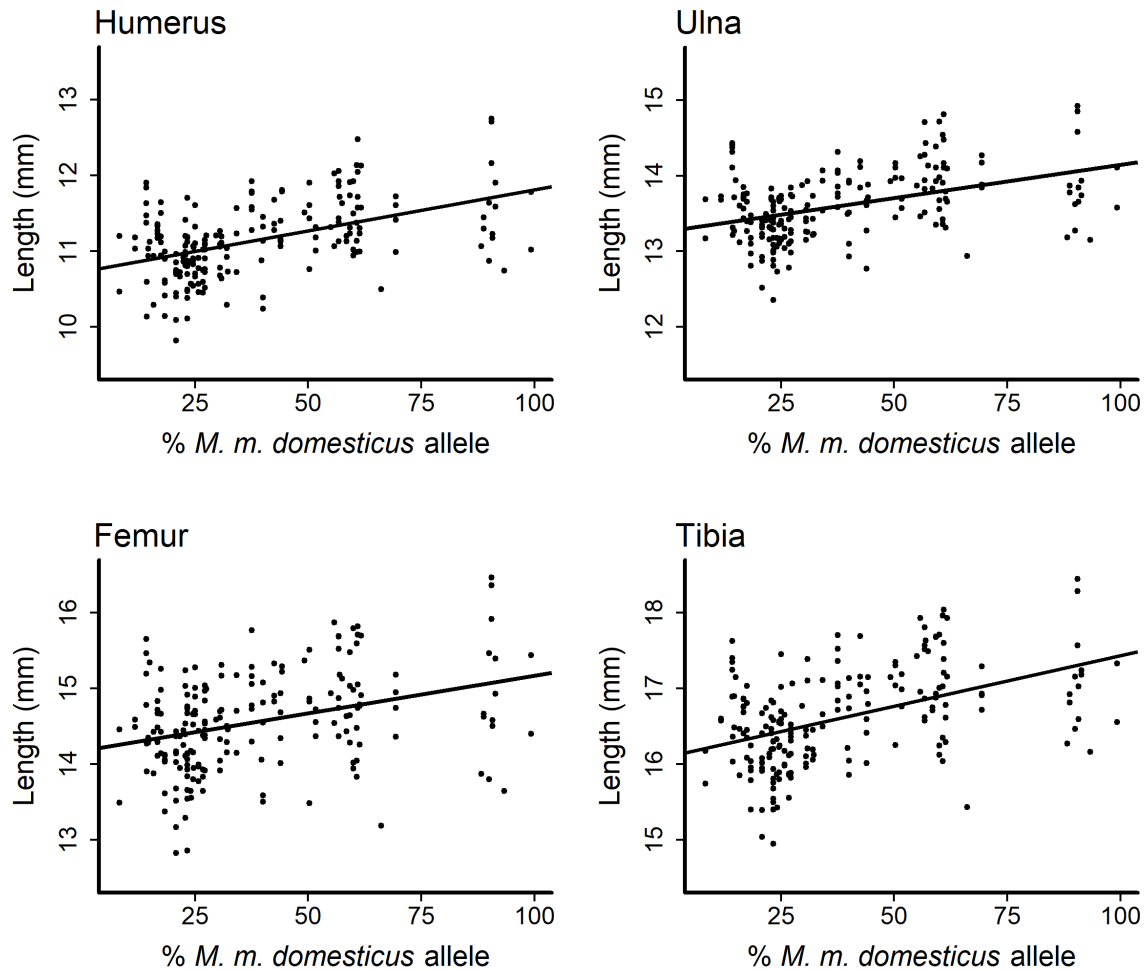
the residual matrix,  $P$  is the original correlation or covariance matrix,  $V$  is the size related eigenvector, while  $V'$  is transposed size eigenvector (Shirai & Marroig 2010). The effect of size was removed from matrices of each population / strain (pooled and already corrected for sex) and hybrid subgroups (previously corrected for the influence of age) and used to recalculate the index of integration. Proportion of variance explained by PC1 was estimated for both correlation and covariance matrices and shown in Table S2.2.

## RESULTS

### *Limb bone lengths*

Data from the raw limb element lengths per sex (mean, standard deviation, standard error of the mean and coefficient of variation) for each population and strain are presented in Table S1.10a of the previous chapter. Females have on average shorter limb bones than males from the respective population / strain, except in the CD1 outbred stock, where females are on average larger than the males. Another exception was seen in the femur and tibia in females of the C57BL/6J strain which are longer in comparison to males. Further, data were pooled for sex per population / strain and average measurements per bone are shown in Table S2.3a and Figure 1.2 in the previous chapter. Information about outbred and inbred groups is provided in the Table S2.3b and Figure 1.3 in the previous chapter. Note that data shown here represent average measurements between the two sides, as well as the information for all bones per population / strain is included. The hybrid group included more individuals in this part of the study, therefore, descriptive statistics per hybrid subgroup is shown in Table S2.3c. Corresponding plots for the length of each limb bone over the percentage of *M. m. domesticus* allele are shown in Figure 2.1.





**Figure 2.1.** Length of each limb bone in millimeters in the mice from the hybrid zone. Regression of the humerus ( $r^2 = 0.230$ ,  $p < 1.067 \times 10^{-12}$ ), ulna ( $r^2 = 0.180$ ,  $p < 5.010 \times 10^{-10}$ ), femur ( $r^2 = 0.115$ ,  $p < 1.051 \times 10^{-6}$ ) and tibia ( $r^2 = 0.212$ ,  $p < 9.72 \times 10^{-12}$ ) over the percentage of *M. m. domesticus* alleles.

All bone lengths showed normal distribution in raw (data not shown) and residual values, smaller deviation was noticed in residual values for tibia of the AH and ulna of the MC population, corresponding Q-Q-plots are shown in Figures S2.1 – S2.3. Individuals from each population and strain were regressed over sex due to a larger influence of this factor, ranging from  $r^2 = 0.071$  ( $p = 0.049$ , slope = -0.313) in femur of the MC population to  $r^2 = 0.745$  ( $p < 0.001$ , slope = -0.418) in ulna of the C57BL/6J strain. Impact of sex was not found only in ulna of CD1 and marginal influence in tibia of the AH population ( $p = 0.053$ ). Results of the linear regression of each limb bone per population and strain are shown in Table S2.1a. Further, outbred and inbred groups were examined after the linear regression for the effects of sex and population / strain for each group separately and the results of the regression are shown in the Table S2.1b. The

hybrid group showed significant association with age only in ulna ( $r^2 = 0.025$ ,  $p = 0.027$ , slope = 0.013) and femur bone ( $r^2 = 0.032$ ,  $p = 0.011$ , slope = 0.021), however for the consistency, all bones were regressed for this factor in the hybrid group.

### Matrix repeatability

Repeatability of matrices per population pooled for sex ranged from 0.814 – 0.978 for correlation and 0.986 – 0.991 for covariance matrices (Table 2.1a). In hybrid subgroups the values were lower due to smaller sample sizes (correlation: 0.719 - 0.918; covariance: 0.980 - 0.986; Table 2.1b) as well as in each sex per population separately (0.686 - 0.967 for correlation and 0.967 - 0.984 for covariance matrices; Table S2.4). Repeatabilities of matrices per group (outbred / inbred and hybrid) were high and ranged between 0.920 – 0.987 for correlation and 0.993 – 0.997 for covariance matrices (Table 2.1). The lowest values for correlation matrices are recorded in MC males and in the H2 hybrid subgroup, which indicate that sampling impact could affect estimation of these matrices. The higher values of repeatability which were found for correlation matrices from populations / strains pooled for sex, as well as for outbred, inbred and hybrid groups confirmed a lower bias due to error in correlation matrices. Covariance matrices showed high values in all compared data sets.

**Table 2.1.** Matrix repeatability for correlation and covariance matrices per population / strain and pooled groups (a); and per hybrid subgroup as well as all data pooled to the hybrid group (b).

(a)

Group	n	Sampling impact	
		Correlation	Covariance
KH	65	0.978	0.991
CB	61	0.929	0.989
MC	55	0.903	0.987
AH	52	0.917	0.986
CD1	48	0.911	0.988
PWD	60	0.966	0.987
C57BL/6J	57	0.814	0.987
Outbred	281	0.973	0.997
Inbred	117	0.987	0.993

(b)

Subgroup	n	Sampling impact	
		Correlation	Covariance
H1	33	0.819	0.986
H2	34	0.719	0.980
H3	30	0.844	0.978
H4	23	0.817	0.978
H5	34	0.918	0.984
H6	24	0.877	0.983
H7	19	0.813	0.981
All	197	0.920	0.997

*Group* – stands for different outbred populations, two inbred strains and data pooled to respective outbred and inbred groups. *Subgroup* – refers to different hybrid subgroups. *n* – sample size per group. *Sampling impact* – reflect the value of each corresponding matrix repeatability.

### *Correlation and covariance matrices*

Correlation and covariance matrices were calculated for each sex per population from the raw data (Table S2.5). On average males showed higher correlations than females. Populations from the outbred group found stronger associations between humerus - femur which are homologous elements and between ulna – tibia (CB, MC). Another high correlation was detected between femur - tibia in AH and KH populations. Differences in this pattern are found in CD1 males which yielded better connectivity in forelimb bones (humerus - ulna) and hindlimb bones (femur – tibia) as well as the highest value in ulna – tibia. In contrast, CD1 females had slightly lower correlations with the highest values between humerus - tibia and humerus - ulna. The inbred PWD strain showed stronger correlations between humerus – tibia and femur – tibia in both males and females, C57BL/6J confirmed this pattern in females, while males had additional strong correlation in ulna – tibia. After correction for sex per population (Table 2.2a), stronger associations between humerus – femur and femur – tibia were found in KH and AH populations. Further, humerus - femur and ulna - tibia were higher in CB and MC, while CD1 found in humerus - ulna and ulna – tibia. Inbred strains revealed higher correlations in humerus - tibia and femur - tibia. Moreover, this pattern in bone correlations was confirmed when data were pooled to the inbred group, while outbred group showed more similarities between humerus - femur and femur – tibia (Table 2.2b).

In hybrid subgroups (residual data for age) higher average correlations were found in groups closer to pure *Mus musculus* (H1, H6 and H7) (Table 2.2c). A pattern with higher correlation between ulna-tibia and femur-tibia was found in H1, H3 and H4, while H2 and H5 subgroups showed humerus-femur and ulna – tibia. Further, H6 and H7 showed humerus – femur, humerus – tibia, femur – tibia. When all data were pooled to the hybrid group, humerus was highly correlated with femur, while the strongest correlations were found in ulna – tibia and femur – tibia.

**Table 2.2.** Correlation (lower triangle) and covariance (upper triangle) matrices in residual values per population / strain (a); outbred and inbred group (b); hybrid subgroups, as well as data pooled to the hybrid group (c).

(a)

n	Population					n	Population				
65	KH					61	CB				
	Bone	Humerus	Ulna	Femur	Tibia		Bone	Humerus	Ulna	Femur	Tibia
	Humerus	-	0.057	0.128	0.119		Humerus	-	0.074	0.132	0.097
	Ulna	0.582	-	0.085	0.107		Ulna	0.641	-	0.123	0.118
	Femur	0.873	0.576	-	0.181		Femur	0.825	0.760	-	0.152
	Tibia	0.820	0.727	0.822	-		Tibia	0.703	0.850	0.789	-
55	MC					52	AH				
	Bone	Humerus	Ulna	Femur	Tibia		Bone	Humerus	Ulna	Femur	Tibia
	Humerus	-	0.110	0.198	0.164		Humerus	-	0.077	0.183	0.134
	Ulna	0.758	-	0.166	0.174		Ulna	0.734	-	0.099	0.084
	Femur	0.894	0.777	-	0.265		Femur	0.841	0.607	-	0.216
	Tibia	0.771	0.849	0.845	-		Tibia	0.775	0.646	0.808	-
48	CD1										
	Bone	Humerus	Ulna	Femur	Tibia						
	Humerus	-	0.148	0.172	0.202						
	Ulna	0.901	-	0.210	0.244						
	Femur	0.779	0.798	-	0.308						
	Tibia	0.892	0.906	0.849	-						
n	Strain					n	Strain				
60	PWD					57	C57BL/6J				
	Bone	Humerus	Ulna	Femur	Tibia		Bone	Humerus	Ulna	Femur	Tibia
	Humerus	-	0.045	0.078	0.069		Humerus	-	0.009	0.018	0.016
	Ulna	0.634	-	0.051	0.065		Ulna	0.636	-	0.016	0.016
	Femur	0.867	0.544	-	0.091		Femur	0.677	0.589	-	0.029
	Tibia	0.876	0.788	0.880	-		Tibia	0.760	0.747	0.748	-

(b)

n	Outbred					n	Inbred				
281	Bone	Humerus	Ulna	Femur	Tibia	117	Bone	Humerus	Ulna	Femur	Tibia
	Humerus	-	0.108	0.196	0.179		Humerus	-	0.028	0.061	0.053
	Ulna	0.757	-	0.163	0.173		Ulna	0.607	-	0.034	0.041
	Femur	0.863	0.742	-	0.285		Femur	0.824	0.470	-	0.088
	Tibia	0.823	0.828	0.855	-		Tibia	0.854	0.679	0.887	-

Pearson correlation matrices (off-diagonal, lower triangle); all *t*-tests were significant at  $p < 0.001$ . Covariances are shown with off-diagonal, upper triangle, dark grey. *n* – sample size per respective group.

(c)

n	Hybrid subgroup					n	Hybrid subgroup				
33	H1					34	H2				
	Bone	Humerus	Ulna	Femur	Tibia		Bone	Humerus	Ulna	Femur	Tibia
	Humerus	-	0.156	0.219	0.208		Humerus	-	0.089	0.172	0.126
	Ulna	0.845	-	0.203	0.201		Ulna	0.805	-	0.129	0.117
	Femur	0.881	0.890	-	0.274		Femur	0.849	0.798	-	0.192
	Tibia	0.875	0.924	0.931	-		Tibia	0.733	0.857	0.768	-
30	H3					23	H4				
	Bone	Humerus	Ulna	Femur	Tibia		Bone	Humerus	Ulna	Femur	Tibia
	Humerus	-	0.048	0.091	0.078		Humerus	-	0.121	0.157	0.194
	Ulna	0.595	-	0.092	0.102		Ulna	0.826	-	0.119	0.156
	Femur	0.739	0.680	-	0.171		Femur	0.742	0.737	-	0.226
	Tibia	0.645	0.769	0.840	-		Tibia	0.833	0.878	0.880	-
34	H5					24	H6				
	Bone	Humerus	Ulna	Femur	Tibia		Bone	Humerus	Ulna	Femur	Tibia
	Humerus	-	0.145	0.248	0.191		Humerus	-	0.214	0.335	0.332
	Ulna	0.822	-	0.203	0.196		Ulna	0.927	-	0.311	0.307
	Femur	0.953	0.804	-	0.281		Femur	0.970	0.920	-	0.485
	Tibia	0.860	0.906	0.879	-		Tibia	0.970	0.916	0.965	-
19	H7					197	All hybrid data together				
	Bone	Humerus	Ulna	Femur	Tibia		Bone	Humerus	Ulna	Femur	Tibia
	Humerus	-	0.233	0.398	0.299		Humerus	-	0.196	0.281	0.279
	Ulna	0.902	-	0.322	0.249		Ulna	0.879	-	0.244	0.263
	Femur	0.966	0.903	-	0.413		Femur	0.897	0.854	-	0.362
	Tibia	0.957	0.922	0.954	-		Tibia	0.892	0.918	0.900	-

Pearson correlation matrices (off-diagonal, lower triangle); all *t*-tests were significant at  $p < 0.001$ . Covariances are shown with off-diagonal, upper triangle, dark grey. *n* – sample size per respective group.

### *Similarities between matrices*

Correlation and covariance matrices were examined for the similarities in their structure. Significance was detected between all comparisons in covariance matrices, although correlation matrices reported significance only among few observations across different populations pooled for sex (Table S2.6a). Inbred strains showed similar structure in their correlations, as well as in observation with the KH outbred population. Among outbred populations, KH showed significant correlation only with MC. In hybrid subgroups, H1 showed significance to H3 (Table S2.6b), however, values were inflated, probably due to a smaller sample size among subgroups. When data were pooled to groups, significant values were found between all groups in correlation and covariance matrices. The highest value is recorded between the outbred and the inbred group, whereas lower value was detected among the hybrid and the inbred group in correlations (Table 2.3). Covariance matrices showed the same result between the inbred and the hybrid group, whereas the strongest similarity was present among the outbred and the hybrid group.

**Table 2.3.** Matrix correlation in correlation (lower triangle, white) and covariance (upper triangle, dark gray) matrices per group.

Group	Hybrid	Outbred	Inbred
Hybrid	-	0.976	0.938
Outbred	0.880	-	0.973
Inbred	0.709	0.954	-

All values in the table were found significant ( $p < 0.05$ ) in Mantel's test (lower triangle, white) and random skewers method (upper triangle, dark gray).

#### *Partial correlations in populations / strains and groups*

In all outbred populations pooled for sex, a correlation pattern between homologous elements was confirmed, with humerus - femur and ulna – tibia showing higher partial correlations except for CD1, where humerus – ulna, femur - tibia and ulna - tibia were higher instead. Another strong connection was noticed in hindlimbs (femur - tibia) of the AH population and absence of significance between ulna and tibia. In contrast, inbred strains showed better connections between ulna - tibia and femur – tibia (Table 2.4a). When data were pooled to the outbred group, the pattern shown in outbred populations was repeated, while the inbred group showed similarity to inbred strains (Table 2.4b). Separate observations per sex found the following correlations: humerus – femur and ulna - tibia in KH, CB, MC in both sexes; AH in males had additional strong correlation between femur – tibia. Differences were obvious in CD1 for both, males and females, while inbred strains found better connectivity between ulna – tibia and femur – tibia, only exception is seen in C57BL/6J females with the highest value in humerus – tibia, beside femur – tibia (Table S2.7). The hybrid group revealed similarity to an outbred group in the overall pattern (HF, UT, FT). While hybrid subgroups confirmed the pattern from general correlations and found additional significant relation between humerus and femur in all subgroups except H4 (Table 2.4c).

Significant correlation in overall structure between different populations was found only between KH – PWD ( $r = 0.812$ ), KH – C57BL/6J ( $r = 0.718$ ) and CBL - PWD ( $r = 0.938$ ) with  $p < 0.05$  (Table S2.8a). Similarities between matrices in hybrid subgroups were found only between H1 and H3 ( $r = 0.887$ ,  $p < 0.05$ ) and close to marginal significance between H1 to H4 and H5, as well as between H3 to H5 subgroups ( $p = 0.083$ ) (Table S2.8b). These results were consistent with previous comparison of correlation matrices, although the values reported from partial correlation similarity were slightly lower and significance between KH and MC was not detected. Comparisons between hybrid, outbred and inbred groups were all significant ( $p < 0.05$ ) with the highest correlation between the hybrid and the outbred group (0.959) and the lowest

between the outbred and the inbred (0.850). Similarity between the hybrid and the inbred group was 0.856. This pattern was different from the one in correlation matrices.

**Table 2.4.** Partial correlations between limb bones in populations / strains pooled for sex (a) and groups (b); hybrid subgroups, as well as data pooled to the hybrid group (c).

(a)

n	Population					n	Population				
65	KH					61	CB				
		Humerus	Ulna	Femur	Tibia			Humerus	Ulna	Femur	Tibia
	Humerus	1					Humerus	1			
	Ulna	-0.005	1				Ulna	-0.074	1		
	Femur	<b>0.610</b>	-0.041	1			Femur	<b>0.615</b>	<b>0.264</b>	1	
Tibia	<b>0.318</b>	<b>0.518</b>	<b>0.345</b>	1	Tibia	0.163	<b>0.628</b>	0.224	1		
55	MC					52	AH				
		Humerus	Ulna	Femur	Tibia			Humerus	Ulna	Femur	Tibia
	Humerus	1					Humerus	1			
	Ulna	0.229	1				Ulna	<b>0.469</b>	1		
	Femur	<b>0.697</b>	-0.013	1			Femur	<b>0.565</b>	-0.127	1	
Tibia	-0.078	<b>0.574</b>	<b>0.452</b>	1	Tibia	0.156	0.216	<b>0.470</b>	1		
48	CD1										
		Humerus	Ulna	Femur	Tibia						
	Humerus	1									
	Ulna	<b>0.482</b>	1								
	Femur	0.033	0.097	1							
Tibia	<b>0.350</b>	<b>0.420</b>	<b>0.450</b>	1							
n	Strain					n	Strain				
60	PWD					57	C57BL/6J				
		Humerus	Ulna	Femur	Tibia			Humerus	Ulna	Femur	Tibia
	Humerus	1					Humerus	1			
	Ulna	0.034	1				Ulna	0.147	1		
	Femur	<b>0.385</b>	<b>-0.484</b>	1			Femur	0.245	0.028	1	
Tibia	<b>0.301</b>	<b>0.729</b>	<b>0.653</b>	1	Tibia	<b>0.385</b>	<b>0.465</b>	<b>0.420</b>	1		

(b)

n	Group					n	Group				
281	Outbred					117	Inbred				
		Humerus	Ulna	Femur	Tibia			Humerus	Ulna	Femur	Tibia
	Humerus	1					Humerus	1			
	Ulna	<b>0.208</b>	1				Ulna	<b>0.202</b>	1		
	Femur	<b>0.533</b>	-0.014	1			Femur	<b>0.332</b>	<b>-0.428</b>	1	
Tibia	<b>0.171</b>	<b>0.501</b>	<b>0.442</b>	1	Tibia	<b>0.274</b>	<b>0.551</b>	<b>0.703</b>	1		

Significant values ( $p < 0.05$ ) from Pearson partial correlations are shown in bold. Significance for the values from the matrices were confirmed with edge exclusion deviance (EED) method described in the methods.

(c)

n	Hybrid subgroup					n	Hybrid subgroup				
33	H1					34	H2				
		Humerus	Ulna	Femur	Tibia			Humerus	Ulna	Femur	Tibia
	Humerus	1					Humerus	1			
	Ulna	0.130	1				Ulna	0.331	1		
	Femur	<b>0.348</b>	0.152	1			Femur	<b>0.565</b>	0.144	1	
Tibia	0.199	<b>0.534</b>	<b>0.502</b>	1	Tibia	-0.019	<b>0.602</b>	0.234	1		
30	H3					23	H4				
		Humerus	Ulna	Femur	Tibia			Humerus	Ulna	Femur	Tibia
	Humerus	1					Humerus	1			
	Ulna	0.177	1				Ulna	0.366	1		
	Femur	<b>0.467</b>	0.002	1			Femur	0.097	-0.181	1	
Tibia	-0.030	<b>0.495</b>	<b>0.613</b>	1	Tibia	0.228	<b>0.564</b>	<b>0.675</b>	1		
34	H5					24	H6				
		Humerus	Ulna	Femur	Tibia			Humerus	Ulna	Femur	Tibia
	Humerus	1					Humerus	1			
	Ulna	0.289	1				Ulna	0.268	1		
	Femur	<b>0.820</b>	-0.218	1			Femur	<b>0.461</b>	0.171	1	
Tibia	-0.096	<b>0.699</b>	<b>0.420</b>	1	Tibia	<b>0.488</b>	0.100	0.376	1		
19	H7					197	All hybrid data together				
		Humerus	Ulna	Femur	Tibia			Humerus	Ulna	Femur	Tibia
	Humerus	1					Humerus	1			
	Ulna	0.072	1				Ulna	<b>0.298</b>	1		
	Femur	<b>0.592</b>	0.113	1			Femur	<b>0.457</b>	0.001	1	
Tibia	0.383	0.409	0.319	1	Tibia	<b>0.179</b>	<b>0.565</b>	<b>0.410</b>	1		

Significant values ( $p < 0.05$ ) from Pearson partial correlations are shown in bold. Significance for the values from the matrices were confirmed with edge exclusion deviance (EED) method described in the methods.

### *Patterns of integration*

Indices were calculated for correlation and variance-covariance matrices for each sex per population / strain (Table S2.9), for populations pooled for sex, hybrid subgroups and per each respective group (Table 2.5). The data in Tables 2.5 and S2.9 are estimated from correlation and covariance matrices before resampling, while the data with resampling used in comparison among tested observations are shown in Table S2.10. My main focus was in inter-trait relationship, i.e. magnitude of morphological integration (Olson & Milller 1958), therefore, the correlation matrix is more informative in exploration of strength and pattern of association among traits (Pavlicev et al. 2009).

The index based on the average coefficient of determination ( $r^2$ ) and relative eigenvalue variances ( $EV_{rel}$ ) confirmed results which are shown for the correlation matrices, i.e. higher indices in males than females, with the highest value found in CD1 males (0.818) and the lowest in C57BL/6J females (0.295) (Table S2.9). This was reflected when populations / strains were pooled, CD1 (0.732) and C57BL/6J (0.484). Hybrid subgroups followed as well the pattern from correlation matrices with higher indices in H6 and H7 and



the lowest in H3 subgroups. Among groups, greater values were recorded for hybrid (0.792), then outbred (0.661) and the lowest in the inbred group (0.541) (Table 2.5).

**Table 2.5.** Integration indices based on correlation ( $r^2$ ,  $EV_{rel}$ ) and variance-covariance matrices (ICV) per population / strain, outbred and inbred group (a); for hybrid subgroups and data pooled to the hybrid group (b). Values shown in the second row for each population / strain, subgroup and group stand for size corrected indices.

(a)

Pop / Strain	n	$r^2$	$EV_{rel}$	ICV
KH	65	0.552	0.552	1.564
		0.008	0.012	0.989
CB	61	0.585	0.585	1.565
		0.007	0.008	0.919
MC	55	0.668	0.668	1.672
		0.004	0.004	0.947
AH	52	0.548	0.548	1.585
		0.007	0.008	0.824
CD1	48	0.732	0.732	1.719
		0.002	0.003	1.082
PWD	60	0.602	0.602	1.575
		0.009	0.013	1.312
C57BL/6J	57	0.484	0.484	1.469
		0.007	0.008	0.915
Outbred	281	0.661	0.661	1.673
		0.003	0.004	0.829
Inbred	117	0.541	0.541	1.568
		0.010	0.016	1.173

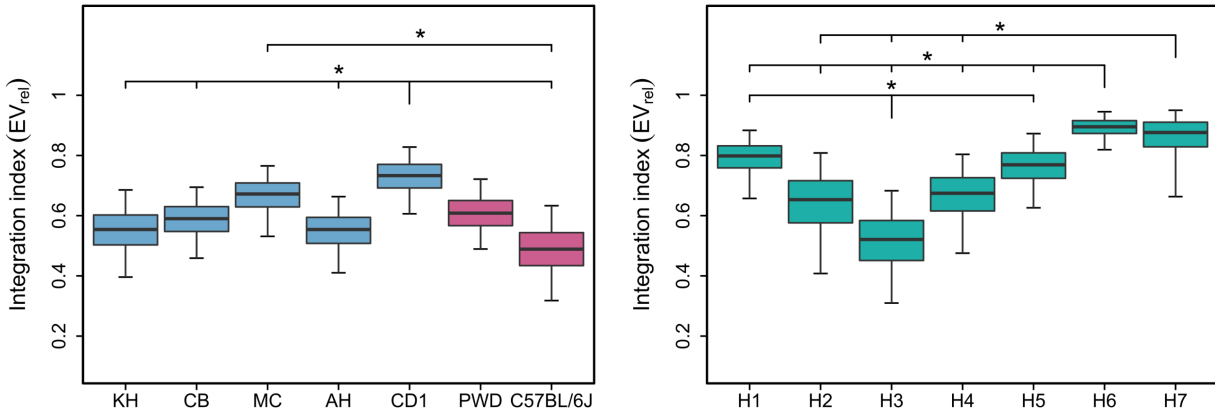
(b)

Subgroup	n	$r^2$	$EV_{rel}$	ICV
H1	33	0.795	0.795	1.798
		0.001	0.001	0.779
H2	34	0.644	0.644	1.620
		0.004	0.004	1.012
H3	30	0.513	0.513	1.546
		0.008	0.009	0.774
H4	23	0.669	0.669	1.664
		0.004	0.004	0.967
H5	34	0.761	0.761	1.763
		0.003	0.003	1.173
H6	24	0.893	0.893	1.905
		$3.26 \times 10^{-4}$	$4.75 \times 10^{-4}$	0.897
H7	19	0.873	0.873	1.887
		$4.68 \times 10^{-4}$	$6.33 \times 10^{-4}$	0.885
All	197	0.792	0.792	1.791
		0.001	0.001	0.812

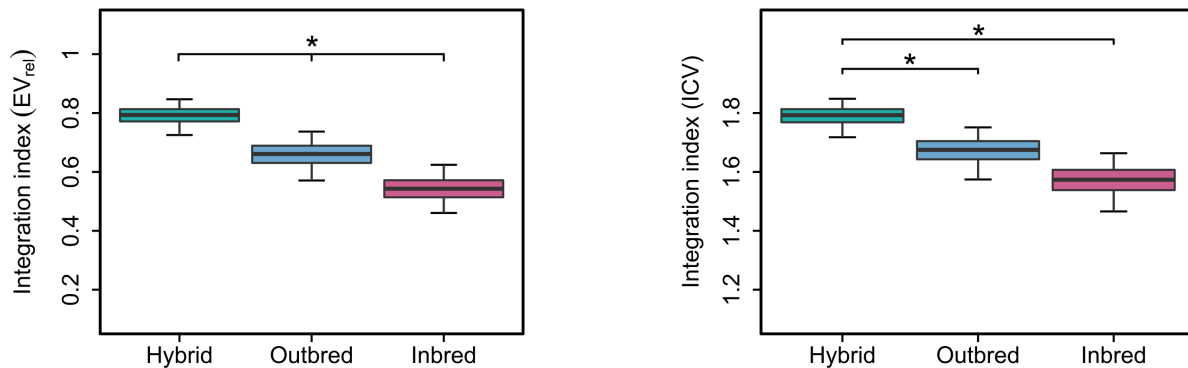
Indices for outbred and inbred group are shown for residual values after accounting for the factors population / strain and sex in each group separately. Indices in table (b) for the hybrid group (All) were corrected just for the effect of age.

Comparison of resampled eigenvalue variances from the correlation matrix ( $EV_{rel}$ ) between different populations / strains yielded significant differences among CD1 to KH, CB, AH and C57BL/6J. Significantly higher integration was found also in MC to C57BL/6J, whereas larger differences were not found between other populations and strains (Table S2.10a and Figure 2.2). Among hybrid subgroups, H6 and H7 showed the highest integration, which was significant for most of the groups, however subgroups closer to pure subspecies (H1 and H7) did not show differences. On the other hand, H3 showed the lowest integration and it was significant for most of the groups except H2 and H4 (Table S2.10b and Figure 2.2). When hybrid, outbred and inbred groups were compared, significant differences were found among all three groups (Table S2.10c and Figure 2.3, left plot). Resampled eigenvalue variances estimated from the variance-

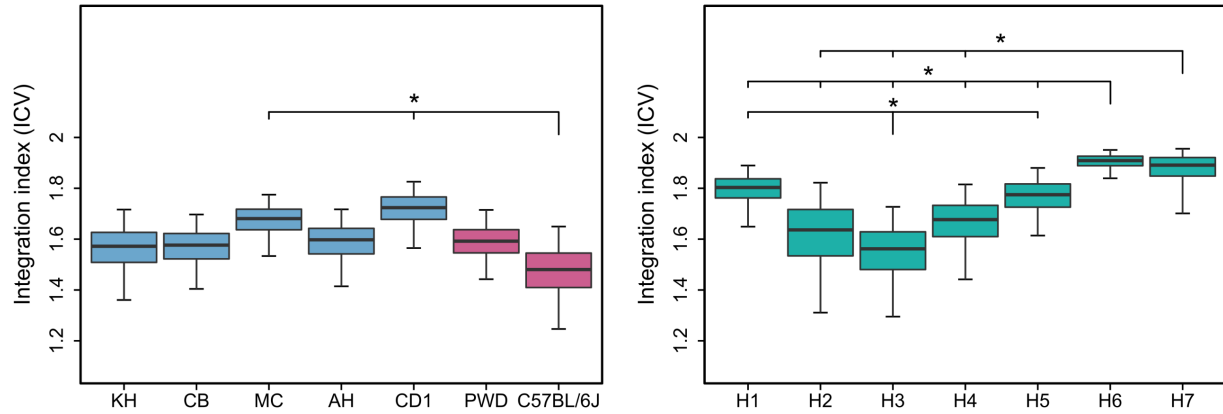
covariance matrix (ICV) showed the same trend, although significance was found only between the hybrid to outbred and inbred groups (Figure 2.3, right plot). Comparison between populations / strains, showed only C57BL/6J to be significantly lower from MC and CD1, while all other groups revealed similar integration. Hybrid subgroups showed the same result as found in the correlation matrices (Figure 2.4).



**Figure 2.2.** Distribution of resampled eigenvalue variance (10,000 replicates) from correlation matrices ( $EV_{rel}$ ) in outbred populations and inbred strains (left plot); in hybrid subgroups (right plot). Error bars indicate the 95% confidence limits of the resampled  $EV_{rel}$ . The longer line bar shows significance between corresponding population with other populations / strains.



**Figure 2.3.** Distribution of resampled eigenvalue variance from correlation (left figure,  $EV_{rel}$ ) and covariance matrices (right plot, ICV) (10,000 replicates) in the hybrid, outbred and inbred group. Error bars indicate the 95% confidence limits of the resampled data.



**Figure 2.4.** Distribution of resampled eigenvalue variance (10,000 replicates) from variance-covariance matrices (ICV) in outbred populations and inbred strains (left plot); in hybrid subgroups (right plot). Error bars indicate the 95% confidence limits of the resampled ICV. The longer line bar shows significance between corresponding population / strain with other populations.

When data were corrected for the effect of size in correlation matrices, an opposite trend was found in comparisons between the groups, with hybrids showing the lowest, while inbred group the highest integration (Tables 2.5, S2.10c and Figure S2.4). Among populations / strains PWD showed the highest value for  $EV_{rel}$  (0.013) and it was close to KH (0.012), while CD1 showed the lowest value (0.003) (Tables 2.5a and S2.10b (values correspond to median values in the table)). Nevertheless, the populations / strains did not differ, except CD1 which was significantly lower from KH, AH, PWD and C57BL/6J (Figure S2.5). In hybrid subgroups H3 showed the highest, while H6 and H7 the lowest correlation (Tables 2.5b, S2.10a and Figure S2.5). Indices based on average coefficient of determination from size corrected data showed a similar pattern, although the values were slightly lower (Table S2.10).

Correction for the first eigenvector revealed which groups are more affected by the influence of general size. This was mainly found in hybrid subgroups. In comparison between populations, CB, MC and CD1 were the most affected, while the pattern of integration between other compared populations / strains remained as with the original data. Moreover, the observation between the groups can be explained due to greater influence of size in populations of the outbred and hybrid subgroups, which resulted in a reverse pattern of integration after the correction. Support for this finding is shown in PC1 (Table S2.2) which explains most of the variation in these populations and hybrid subgroups.

Indices computed based on size corrected covariance matrices yielded a different outcome. However, the PWD strain was still the least affected and showed the highest integration in comparison to other

populations except CD1 (Figure S2.6). In addition, CD1 was different to AH, which showed the lowest value. Hybrid subgroups showed less differences, with only significance in H3 which showed the lowest integration and differed from H2 and H5 (Figure S2.6). When groups were compared after size correction, hybrid and outbred group showed similar integration which was significant only between these two groups in comparison to the inbred group (Figure S2.7).

An additional test with a larger data set, when all available data are included (regardless of larger differences between the right and the left side), showed similar patterns to the analyzed data with regular and size corrected correlation matrices ( $EV_{rel}$ ), although AH and C57BL/6J showed lower degree of integration and thus additional significance between AH to MC and PWD (Figure S2.8). In data corrected for size, results were the same as with the before mentioned data, only AH showed additional differences to MC (Figure S2.9).

As reported previously (Shirai & Marroig 2010; Haber 2011), equal results were found for indices based on relative eigenvalue variances ( $EV_{rel}$ ) and average coefficient of determination ( $r^2$ ) with original data. This was confirmed in my data set, although data corrected for size showed similar but not the same values between these two estimates. Therefore, results for indices computed from correlation matrices in raw data are discussed for  $EV_{rel}$  index, but results based on  $r^2$  are also shown in Table S2.10. Size corrected data are also reported from eigenvalue variances ( $EV_{rel(ns)}$ ), since results were similar. Moreover, differences in significance between data analyzed with  $EV_{rel(ns)}$  and  $r^2_{ns}$  were not found, except CD1 to CB and AH, and H6 to H5 which were only marginal with  $EV_{rel(ns)}$ .

## DISCUSSION

The overall results from correlation and partial correlation matrices were similar in populations / strains and hybrid subgroups, with few exceptions. Inbred strains showed stronger correlations between humerus – tibia, whereas partial correlations revealed ulna – tibia, beside femur – tibia which was common in both comparison between these two matrices. Further, this was reflected when data are pooled to the inbred group. Outbred and hybrid groups confirmed the general pattern expected from developmental factors that structure limb covariation, i.e. all comparison shared strong correlation between humerus – femur which are serially homologous bones as well as among ulna – tibia that belong to the same segment (zeugopod). Higher values were recorded also for femur – tibia which belong to the

same limb. Moreover, this connection (femur - tibia) overlapped in all three groups and confirms a better correlation in hindlimbs. Only the results for the CD1 stock differed in the pattern and showed better correlations in bones of the same limb, as well as weaker connectivity between humerus – femur. However, in CD1 animals I found relatively higher values, which might be explained by larger body size of these individuals. Greater correlation between humerus – femur is widely described across different taxa (Young & Hallgrímsson 2005; Schmidt & Fischer 2009; Young et al. 2009; Kolarov et al. 2017) and corresponds to proximo – distal patterning. These bones are proximal elements, thus better connectivity among them is directly associated to earlier developmental timing in their ossification in comparison to more distal elements (Weisbecker 2011). Similarity in the overall structure between correlation matrices was found among inbred strains, as well as between both inbred strains to KH and KH with MC. This relationship might be explained due to similar general size between these populations / strains, but might also be influenced by the fact of the same sub-species background between KH (*M. m. musculus*) and PWD (wild derived strain from *M. m. musculus*). The absence in similarity between hybrid subgroups can be due to smaller sample sizes per compared group and thus lower precision in matrix estimation. Overall, these results would suggest possible structural differences between populations and subgroups which lacked in significance in matrix correlations. In contrast, significance was found between all three compared groups (hybrid, outbred and inbred).

Comparison in the magnitude of eigenvalue variances from populations / strains did not reveal major differences between populations / strains, except for the CD1 outbred stock, which showed the highest values. In contrast, when populations / strains are pooled to groups, the outbred group showed higher degree of integration in comparison to the inbred group. This would confirm a hypothesis about stronger stabilizing selection in wild populations. However, results might be biased due to CD1 which could potentially generate larger differences between outbred and inbred groups.

Higher variability was detected among hybrid subgroups, which might be influenced by developmental causes, because individuals in this group were twice younger in comparison to other groups. Previous studies also reported that integration can change during ontogeny (Cheverud et al. 1983; Cheverud & Leamy 1985; Zelditch 1988; Zelditch & Carmichael 1989; Cane 1993; Ackermann 2005; Ivanovic et al. 2005). Further, smaller sample size per subgroup and lower repeatability of correlation matrices could also bias the result (Young et al. 2009). Nevertheless, subgroups closer to pure subspecies were found to be more integrated. When all data from subgroups were pooled to the hybrid group and compared with outbred and inbred groups, significantly higher integration was recorded in the hybrid group.

Similar results were obtained in hybrid subgroups from resampled eigenvalues from the variance-covariance matrix, which should reflect the level and pattern of variation (Pavlicev et al. 2009). While differences between outbred populations and inbred strains were seen only in the C57BL/6J strain which was less integrated, overall differences in integration between these groups were not found. This would be explained by similar variances among outbred and inbred groups.

Pattern and level of correlations between traits might be under a large influence of variation due to size (Zelditch 1988; Marroig et al. 2004). Impact of general size can obscure the biological signal of morphological integration by masking weak correlations (Riska 1986; Zelditch 1988; Magwene 2001; Marroig et al. 2004; Klingenberg 2008). Numerous genetic, developmental and ecological factors generate variation in size (Patton & Brylski 1987) which further produce higher correlations in larger individuals and thus increase overall integration between traits (Porto et al. 2013). Hallgrímsson et al. (2002) found a large proportion of variance explained by principal components in Rhesus macaques and CD1 mice which affected integration patterns in these species due to effects of the general size. Higher morphological integration due to higher variance of the character size was found in the chondrocranium of the inbred mutant brachymorph mice (Hallgrímsson et al. 2006). Overall correlation in the skull was influenced by the variance in the growth of the chondrocranium. A study in paedomorphic and metamorphic alpine newt also reported bias in the integration due to size, as the differences between the groups lacked in significance after removal of this factor (Kolarov et al. 2017).

After controlling for the effect of size in our data set and comparing the strength of morphological integration, three populations from the outbred group as well as hybrid subgroups showed to be influenced with this change and resulted in the opposite trend as the one with non-corrected data. However, higher magnitude of integration was recorded in inbred strains, although overall differences between populations and strains were less expressed. On the other hand, the pattern for hybrid subgroups differed between the correlation and covariance matrices. Further, comparison between groups found the highest level of integration in the inbred group in both compared matrices.

## **CONCLUSIONS**

The expectation to find stronger integration in more natural populations was shown in data that were not corrected for size, where the hybrid group showed the highest, while inbred group showed the lowest

level of integration. Previous studies with inbred strains detected more unstable covariance structure in these mice (Jamniczky & Hallgrímsson 2009). This was explained through genetic variance of the developmental process which underlie covariance structure, therefore, the alteration of this structure is dependent on the higher magnitude of change in variance in one of the processes when these processes are controlled by a larger amount of genetic variance (Hallgrímsson et al. 2009). Another study in inbred mice showed a lower level of integration in the strain with craniofacial malformations (Hallgrímsson, et al. 2004). However, after correction for size in our data, the inbred group showed higher integration. This might indicate on differences in functional or developmental interactions. Interestingly, similar correlation matrices were found in inbred strains (before size correction), with stronger connection between humerus and tibia. This type of inter-trait relationship is mostly described in carnivora (Martin-Serra et al. 2015) where these bones are defined as functionally equivalent (Schmidt & Fischer 2009), which was caused by reorganization in the position of certain bones. Changes in developmental processes that underlie appendicular structures can occur in order to follow biomechanical requirements which might further influence the pattern of phenotypic integration at an evolutionary level (Martin-Serra et al. 2015).

Selection pressure on fore- and hindlimbs could have different evolutionary outcomes, depending on the strength of integration between the limb bones (Sanger et al. 2011). Higher integration would constrain possible change and diversification among these structures, whereas lower levels of cohesion would allow further adaptation and independent evolution (Hallgrímsson et al. 2002; Young & Hallgrímsson 2005).

## CHAPTER III

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

Neva Škrabar<sup>1</sup>, Leslie M. Turner<sup>1,2</sup>, Luisa F. Pallares<sup>1,3</sup>, Bettina Harr<sup>1</sup>, Diethard Tautz<sup>1\*</sup>


<sup>1</sup>Max-Planck Institute for Evolutionary Biology, Plön, Germany

<sup>2</sup>Department of Biology and Biochemistry, Milner Centre for Evolution, University of Bath, Bath, UK

<sup>3</sup>Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ, USA



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

Neva Škrabar<sup>1</sup> | Leslie M. Turner<sup>1,2</sup> | Luisa F. Pallares<sup>1,3</sup> | Bettina Harr<sup>1</sup> |  
Diethard Tautz<sup>1</sup> 

<sup>1</sup>Max-Planck Institute for Evolutionary Biology, Plön, Germany

<sup>2</sup>Department of Biology and Biochemistry, Milner Centre for Evolution, University of Bath, Bath, UK

<sup>3</sup>Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ, USA

## Correspondence

Diethard Tautz, Max-Planck Institute for Evolutionary Biology, Plön, Germany.  
E-mail: tautz@evolbio.mpg.de

## Funding information

Max-Planck-Gesellschaft

## 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 genomewide 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 *Mus musculus 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.

## KEYWORDS

ecological genetics, hybrid zone, limbs, mapping, quantitative genetics

## 1 | INTRODUCTION

Vertebrate limbs are an excellent model system to study morphological integration, developmental stability and phenotypic covariance patterns (Hallgrímsson, Willmore, & Hall, 2002; Kolarov, Ivanovic, & Kalezić, 2011; Lande, 1980; Pavlicev, Wagner, Noonan, Hallgrímsson, & Cheverud, 2013; Rolian, Lieberman, & Hallgrímsson, 2010; Schmidt & Fischer, 2009; Young, 2013; Young & Hallgrímsson, 2005). The forelimb and hindlimb represent serially homologous structures, that is, repeated parts that share a developmental architecture.

During evolution, serially homologous structures are formed when the same underlying developmental programme 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 programme. 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 (Schmidt & Fischer, 2009;

Young & Hallgrímsson, 2005), 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, 2006; Hansen, Ambruster, Carlson, & Pelabon, 2003). 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 (rQTLs) (Cheverud et al., 2004; Pavlicev et al., 2008, 2013). The identification of rQTLs is of particular interest for serially homologous structures, such as the limbs, as they are thought to underlie their evolutionary specialization, such as the morphological differentiation that is evident between the forelimbs and hindlimbs in mice (Figure 1).

QTLs influencing bone length have been mapped in various studies of laboratory mouse strains derived from the subspecies *Mus musculus domesticus* (Wolf, Pomp, Eisen, Cheverud, & Leamy, 2006; Norgard et al., 2008, 2009; Pavlicev et al., 2008, 2013; Parmenter et al., 2016; —and references therein). These studies found a high degree of covariation between traits in forelimbs and hindlimbs 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, 2013) used this approach to identify loci conveying variational independence between limb bones.

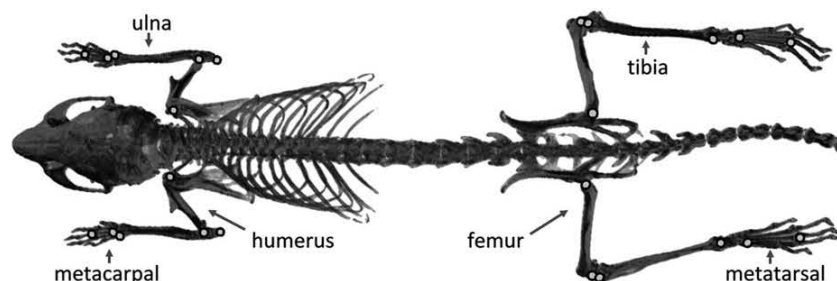
Employing a standard F2 QTL mapping approach for detecting loci connected to length variation of limb bones does not provide sufficient resolution to identify underlying causative genes (Flint, Valdar, Shifman, & Mott, 2005). Although fine-mapping approaches using advanced intercrosses provide further 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 & Eskin, 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). Genomewide 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 &

Moore, 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, Zaitlen, Reich, & Patterson, 2010).

Hybrid zones between subspecies are of particular interest for mapping, as they represent natural cases of advanced intercrosses (Rieseberg & 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 harbour 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; that is, a relatively lower marker density is required compared to standard genomewide 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 & Harr, 2014) and craniofacial traits (Pallares, Harr, Turner, & Tautz, 2014). The Western and Eastern house mouse subspecies (*Mus musculus domesticus* and *Mus musculus musculus*) form a narrow hybrid zone in the middle of Europe (Guenet & Bonhomme, 2003; Phifer-Rixey & 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 (Pallares et al., 2014; Turner & Harr, 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, Turner, & Tautz, 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 was 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 hindlimbs, and bones that belong to the same segment, that is, stylopod, zeugopod or autopod share common developmental genetic networks. On the other hand, as forelimbs and hindlimbs 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



**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



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 and 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.

## 2 | METHODS

### 2.1 | 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, Schwahn, & Harr, 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<sub>2</sub> asphyxiation between 9 and 12 weeks of age. The mapping population consists of 172 males including full-siblings, half-siblings and unrelated individuals (Table S1). Genomewide association studies of sterility traits (Turner & Harr, 2014) and craniofacial traits (Pallares et al., 2014) were previously reported for these mice.

### 2.2 | 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  $\mu$ A, voxel size: 38  $\mu$ m. 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 hindlimb (femur, tibia and metatarsal bone) using the TINA landmarking tool (Schunke, Bromiley, Tautz, & Thacker, 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 capitata-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, Gabet, Hajbi-Yonissi, & Müller, 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, Lougheed, and Handford (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 (Table S2). 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; that is, the animals with only one intact side are represented with only one measurement. A total of 172 individuals were included in the data set for association mapping.

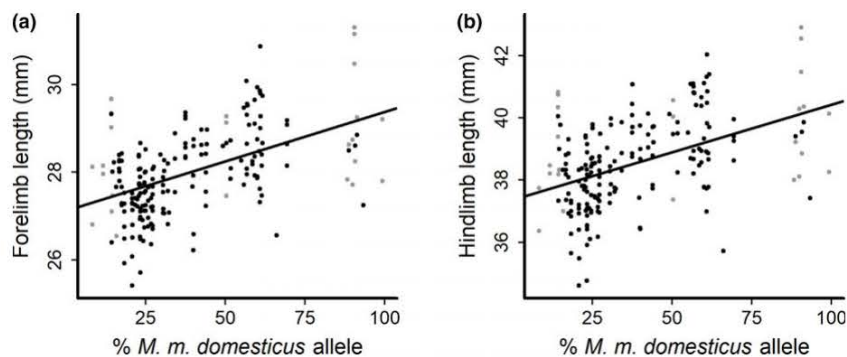
The individual bone lengths were normally distributed (Tables S3 and S4, Figures 1 and 2). Age showed small but significant correlations with metacarpal ( $r^2 = .026$ ,  $p = .036$ , slope =  $-0.003$ ) and metatarsal ( $r^2 = .045$ ,  $p = .005$ , slope =  $-0.011$ ) measurements. Hence, for consistency, 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. 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 hindlimb as the sum of femur, tibia and metatarsal bones. Following the procedure described in Pallares et al. (2016) and based on a specific SNP data set from 37 loci, we performed regressions of forelimb and hindlimb lengths versus percentage of *M. m. domesticus* alleles ("PairMid" in Table S1).

### 2.3 | 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 GRCh38/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  $1 \times 10^{-6}$  ( $VIF = 1/(1-R^2)$ ) (see Turner & Harr, 2014 for details). Essentially, this procedure removed nearby SNPs in strong LD from the data set. The SNP genotype data set contained genotype data for 185 individuals, which were filtered for SNPs  $\geq 5\%$  minor allele frequency (MAF). To use the same set of SNPs as in Turner and Harr (2014) and Pallares et al. (2014) for this 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 performed univariate linear mixed model analysis implemented in genomewide efficient mixed-model



**FIGURE 2** Correlation between limb length and genomic ancestry. Regression of forelimb (a) and hindlimb (b) over the percentage of *M. m. domesticus* alleles. Forelimb values consist of summed values of humerus, ulna and metacarpal bone ( $r^2 = .25$ ,  $p = 8.8 \times 10^{-14}$ ); hindlimb values are shown as a sum of femur, tibia and metatarsal bone ( $r^2 = .22$ ,  $p = 5.1 \times 10^{-12}$ ). Grey dots include values of 25 additional animals taken from Pallares et al. (2016) for which only partial genotype data were available; that is, these were not included in the association analyses

association—GEMMA (version 0.94.1) (Zhou & 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 centred 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, Carbonetto, & Stephens, 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 hyperparameters (PVE, PGE) were recorded for every 10th 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; that is, 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 & Stephens, 2014) of bone phenotypes because preliminary mapping analyses showed our sample size was too small to reliably reach convergence.

## 2.4 | Statistical analysis

To define genomewide 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 data set. The significance thresholds for each bone were then defined as the 5th 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 was 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 5th percentile of values for 10,000 permutations for each analysis.

We also visually inspected the Manhattan plots for obvious “peaks,” that is, 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 cut-off. A significance threshold of  $p < 10^{-5}$  is expected to yield one false-positive SNP association per trait in our data set 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 data set, we do not analyse 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 data set (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).

## 2.5 | Genetic correlation

We used bivariate restricted maximum-likelihood analysis implemented in the genomewide complex trait analysis—GCTA (version 1.26.0) (Yang, Lee, Goddard, & Visscher, 2011; Yang et al., 2010) to estimate the genetic correlation between all combinations of two traits. This analysis essentially compares the phenotypic similarity and the genetic similarity between individuals within and across two traits. The genetic correlation ( $r_g$ ) is defined as  $r_g = \frac{\text{COV}_g(t_1, t_2)}{\sqrt{\text{Var}_g(t_1) \cdot \text{Var}_g(t_2)}}$ , where ( $\text{var}_g(t_i)$ ) is the additive genetic variance of trait  $i$  and covariance ( $\text{cov}_g(t_i, t_j)$ ) 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 population structure (Yang et al., 2011). 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, especially in small samples such as ours (Visscher et al., 2014).

## 2.6 | 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 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).

## 3 | RESULTS

### 3.1 | 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 (Table S2), and hence, we averaged these measures as well as their measuring errors.

Across the hybrids, lengths of forelimbs and hindlimbs 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).

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

### 3.2 | Genomewide association mapping and genetic architecture

Genomewide association mapping included 143,592 SNPs and each bone was analysed 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 Figure S3). When using genomewide 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 Section 2), distributed across nine chromosomes and six bones, with an overlapping region on chr9 for humerus and tibia, and on chr16 for metacarpal and metatarsal bones (Table S5). The size of LD regions around significant SNPs varied from 51 bp to 601 kb (median 114 kb).

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 data set (i.e., “SNP heritability”—Wray et al., 2013) are reported in Table 2.

**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 <sup>a</sup>	0.25	0.20	0.18	0.12	5.69	0.53

<sup>a</sup>The per cent of measurement error is independent of the units of the measured objects.

**TABLE 2** Phenotypic and genetic correlations

Trait	Humerus	Ulna	Femur	Tibia	Metacarpal	Metatarsal
Humerus	<b>0.77<sup>c</sup></b>	0.86 <sup>a</sup>	0.88 <sup>a</sup>	0.88 <sup>a</sup>	0.40 <sup>a</sup>	0.55 <sup>a</sup>
Ulna	0.91 <sup>b</sup>	<b>0.82<sup>c</sup></b>	0.84 <sup>a</sup>	0.91 <sup>a</sup>	0.47 <sup>a</sup>	0.64 <sup>a</sup>
Femur	0.96 <sup>b</sup>	0.92 <sup>b</sup>	<b>0.85<sup>c</sup></b>	0.89 <sup>a</sup>	0.34 <sup>a</sup>	0.51 <sup>a</sup>
Tibia	0.88 <sup>b</sup>	0.91 <sup>b</sup>	0.96 <sup>b</sup>	<b>0.85<sup>c</sup></b>	0.43 <sup>a</sup>	0.62 <sup>a</sup>
Metacarpal	0.51 <sup>b</sup>	0.48 <sup>b</sup>	0.50 <sup>b*</sup>	0.54 <sup>b</sup>	<b>0.68<sup>c</sup></b>	0.74 <sup>a</sup>
Metatarsal	0.54 <sup>b</sup>	0.62 <sup>b</sup>	0.54 <sup>b</sup>	0.60 <sup>b</sup>	0.77 <sup>b</sup>	<b>0.67<sup>c</sup></b>

<sup>a</sup>Phenotypic correlation coefficients (off-diagonal, upper triangle, dark grey) in 172 genotyped individuals (Pearson product-moment correlation, all *t*-tests were significant at  $p < .001$ ).

<sup>b</sup>Genetic correlations from bivariate analysis, based on genomewide 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,  $p < .05$ , one-sided test). \*For femur-metacarpal, the correlation was significantly different from one, but not from zero.

<sup>c</sup>Proportion of phenotypic variation explained (PVE) by all the SNPs used in the mapping from Bayesian sparse linear mixed model (diagonal—bold, light grey).

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) (Table S6), 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 S1 and Table S7.

For highly polygenic traits, we expect that the proportion of variance in bone length explained by each chromosome should be correlated with chromosome length (Berenos et al., 2015; Yang et al., 2011). 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 Figure S4).

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

As our sample size is limited, we used simulations to assess the power of our data set to detect significant SNPs at our chosen significance threshold. As detailed in Supplementary Text S2 (plus Table S8), our power to detect at least one causal SNP as significant of 10 (or 15, respectively) simulated is close to 100%. However, in none of the 1,000 simulated data sets 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.

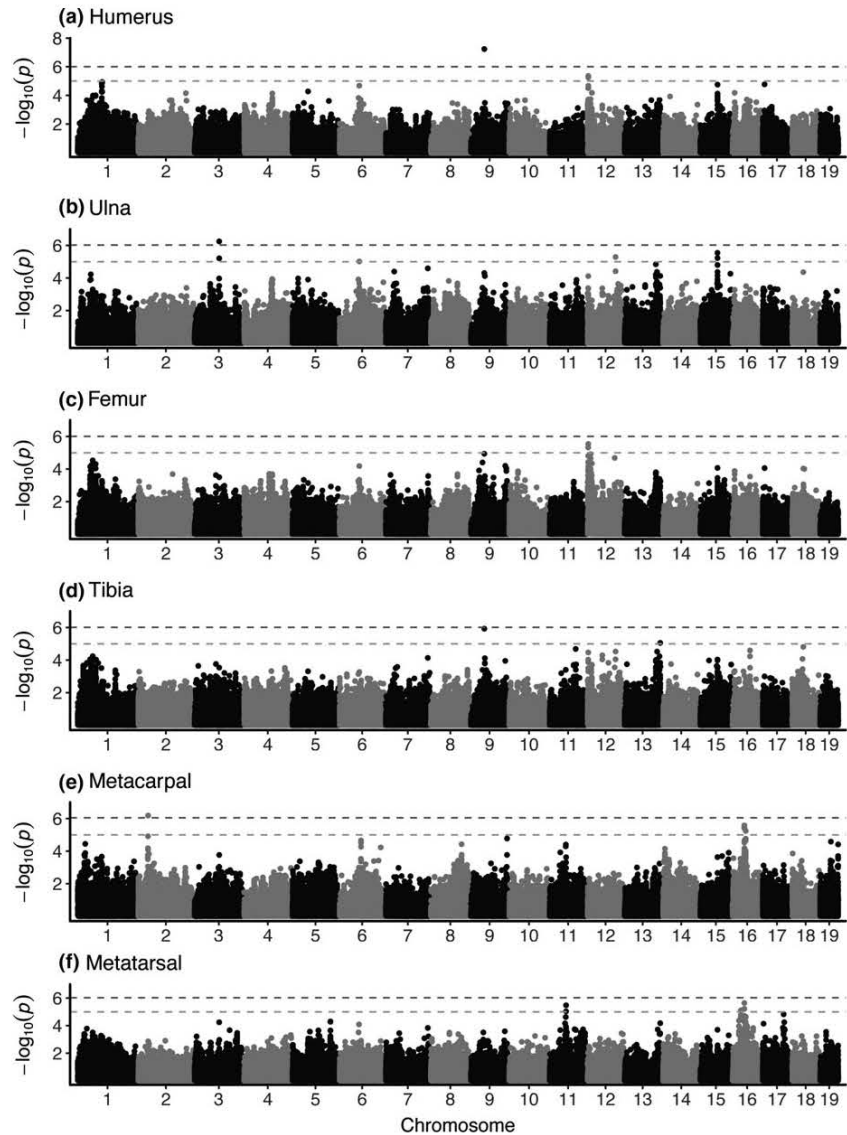
### 3.4 | Candidate genes

To identify candidate genes influencing limb bone length in mice, we evaluated all annotated genes (coding and noncoding) 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 Section 4 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.

### 3.5 | 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 Section 2 for specific combinations of bones). This revealed 17 SNPs above the genomewide permutation thresholds for the following bone combinations





**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 genomewide threshold (GWT) based on 10,000 permutations (upper line) in (a) humerus (GWT =  $9.87 \times 10^{-7}$ ), (b) ulna (GWT =  $9.45 \times 10^{-7}$ ), (c) femur (GWT =  $9.87 \times 10^{-7}$ ), (d) tibia (GWT =  $9.65 \times 10^{-7}$ ), (e) metacarpal (GWT =  $8.96 \times 10^{-7}$ ), (f) metatarsal bone (GWT =  $9.47 \times 10^{-7}$ )

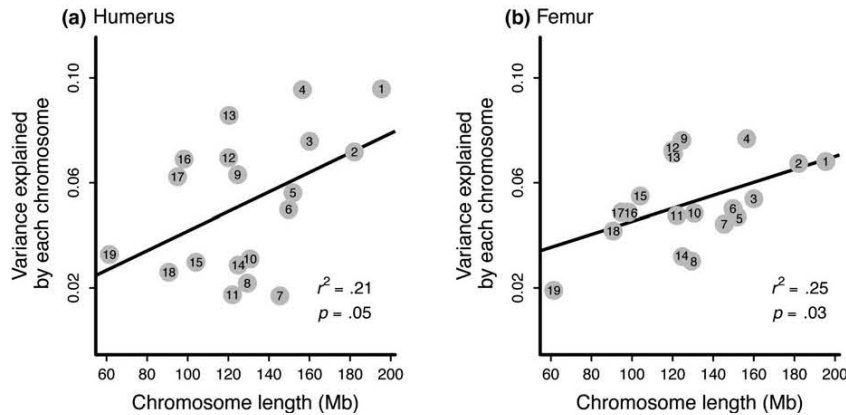
(the first capital letter refers to the bone tested, while the second capital letter for the one used as a covariate): HcovU, FcovT, HcovF, UcovT, UcovF and FcovU (Table 4 and Table S9). Manhattan plots are provided in Figure S6, and the corresponding Q-Q-plots are provided in Figure S7. None of the SNPs significant at genomewide 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 (min = 113 bp, 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 Table S9).

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 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 Section 4 for details).

## 4 | DISCUSSION

We have explored the use of animals from a natural hybrid zone for genomewide 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



**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

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 permutation tests. For highly polygenic traits, permutation tests are overly conservative, especially if sample sizes are relatively small (Yang et al., 2010, 2011). 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  $p < 10^{-4}$  (Table S5), also indicating correlated responses across traits. Interestingly, the chr9 locus shows an additional association with ulna and femur at this lower  $p$ -value cut-off.

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 (Hallgrímsson et al., 2002; Young & Hallgrímsson, 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 (Capdevila & Belmonte, 2000; Shubin, Tabin, & Carroll, 1997), 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 & 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.

#### 4.1 | 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; Sheeba, Andrade, & Palmeirim, 2016; Tickle, 2006). Interestingly, none of these genes overlap with any candidate regions identified here (>1 Mb 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/hindlimb differences in mice, while we found it in the individual bone mapping of metacarpal length, a bone that was not included in their analysis.



**TABLE 3** Genomic regions associated with limb bone length, identified by individual bone analysis

Bone	Region	N of SNPs <sup>a</sup>	<i>p</i> -value of the best SNP <sup>b</sup>	Size (Mb) <sup>c</sup>	Genes in the region <sup>d</sup>
Metacarpal	chr2: 35028134–35029436	1 (2)	<b>6.46 × 10<sup>-7</sup></b>	0.001 [0.114]	<i>Hc</i> , <i>Al182371</i>
Ulna	chr3: 82469792–82583792	2 <sup>e</sup>	<b>5.60 × 10<sup>-7</sup></b>	[0.114]	<i>Npy2r</i>
Ulna	chr6: 67069424–67183424	1	9.56 × 10 <sup>-6</sup>	[0.114]	<i>E230016M11Rik</i> , <i>AK079709</i> , <i>AK039826</i>
Humerus	chr9: 46604715–46718715	1	<b>5.74 × 10<sup>-8</sup></b>	[0.114]	<i>Gm22805</i>
Tibia		1	1.17 × 10 <sup>-6</sup>		
Metatarsal	chr11: 55936031–55964675	2 (3)	3.29 × 10 <sup>-6</sup>	0.029 [0.114]	<i>Gm12239</i>
Femur	chr12: 5161413–5762255	2 (12)	2.88 × 10 <sup>-6</sup>	0.601	<i>Klhl29</i> , <i>2810032G03Rik</i>
Humerus	chr12: 5792918–5907264	2 (12)	4.31 × 10 <sup>-6</sup>	0.114	<i>AK135963</i>
Ulna	chr12: 9377662–93851912	1 (3)	5.09 × 10 <sup>-6</sup>	0.074 [0.114]	/
Tibia	chr13: 117841280–117955280	1	8.69 × 10 <sup>-6</sup>	[0.114]	<i>Hcn1</i>
Ulna	chr15: 58538615–58676719	2 (9)	2.83 × 10 <sup>-6</sup>	0.138	<i>Fer116</i>
Metatarsal	chr16: 30263056–30263107	1 (2)	7.74 × 10 <sup>-6</sup>	5.10 × 10 <sup>-5</sup> [0.114]	<i>Cpn2</i> , <i>Lrrc15</i> , <i>Gp5</i> , <i>Atp13a3</i>
Metacarpal	chr16: 41712286–42136622	3 (33)	2.62 × 10 <sup>-6</sup>	0.424	<i>Lsamp</i>
Metatarsal		2 (33)	2.27 × 10 <sup>-6</sup>		
Metacarpal	chr16: 45984088–46202709	1 (16)	5.68 × 10 <sup>-6</sup>	0.219	<i>Cd96</i> , <i>Gm4737</i> , <i>Plcx2d<sup>f</sup></i>

<sup>a</sup>Number of SNPs in the region with  $p < 10^{-5}$ , in parentheses number of total SNPs that are in LD within the same region.

<sup>b</sup>*p*-values in bold represent the SNPs above the permutation-based threshold.

<sup>c</sup>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.

<sup>d</sup>Annotated genes overlapping with the region; genes involved in limb development, phenotype or expression are highlighted in bold.

<sup>e</sup>two SNPs within 6 kb, but not in LD.

<sup>f</sup>Overlap with previous QTL study on limb length in Pavlicev et al. (2013).

Another gene implicated in metacarpal length is the haemolytic 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 sub-cortical 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 signalling 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 (Schlötterer, 2015; Tautz & Domazet-Lošo, 2011). 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 *Mus musculus musculus* populations (can be seen in the genome sequence data provided in Harr et al. (2016)); that is, 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 coreceptor of *wnt* signalling with a direct involvement in osteoblastogenesis and cartilage development (Joeng, Schumacher, Zylstra-Diegel, Long, & Williams, 2011). Humerus length is associated with a region including *Morc3*, which is involved in calcium homeostasis and osteoblast differentiation (Jadhav, Teguh, Kenny, Tickner, & Xu, 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, Tran, & Cooper,

**TABLE 4** Genomic regions associated with limb bone length, identified by mapping with covariates

Bone <sup>a</sup>	Region	N of SNPs <sup>b</sup>	p-value of the best SNP <sup>c</sup>	Size (Mb) <sup>d</sup>	Genes in the region <sup>e</sup>
UcovF	chr1: 3625815–4718067	2 (87)	$5.63 \times 10^{-7}$	1.092	<i>Sox17</i> , <i>Rp1</i> , <i>Xkr4</i> , <i>AK149000</i>
FcovU	chr1: 9262814–9380814	1	$5.92 \times 10^{-6}$	[0.118]	<i>Sntg1</i>
HcovF	chr1: 23690398–23716894	1 (3)	$1.39 \times 10^{-6}$	0.026 [0.118]	/
FcovT	chr1: 84059655–84952620	2 (21)	$4.21 \times 10^{-8}$	0.893	<i>Fbxo36</i> , <i>Pid1</i> , <i>Dner</i> <sup>f*</sup> , <i>Trip12</i> <sup>*</sup> , <i>Slc16a14</i> , <i>Mir6353</i> , <i>AK032919</i> , <i>AK036072</i>
FcovH	chr1: 119187233–119322362	1 (9)	$8.77 \times 10^{-6}$	0.135	/
HcovF	chr1: 119320805–119438805	1	$4.10 \times 10^{-6}$	[0.118]	<i>Inhbb</i> , <i>AK039419</i>
HcovT	chr1: 128607010–128725010	1	$7.48 \times 10^{-6}$	[0.118]	/
HcovF	chr1: 152359185–152477185	1	$3.60 \times 10^{-6}$	[0.118]	<i>Tsen15</i> , <i>Colgalt2</i>
HcovF	chr1: 154018113–154036315	1 (3)	$4.87 \times 10^{-6}$	0.018 [0.118]	<i>Gm29291</i> , <i>Gm28286</i> , <i>AK043564</i> , <i>AK154552</i>
HcovF	chr1: 154053011–154077631	2 (2)	$5.20 \times 10^{-6}$	0.025 [0.118]	<i>AK043564</i> , <i>AK154552</i>
HcovT		2 (2)	$3.29 \times 10^{-6}$		
TcovH		2 (2)	$9.85 \times 10^{-6}$		
FcovH	chr2: 8781928–9002418	2 (10)	$3.87 \times 10^{-6}$	0.22	/
MCcovMT	chr2: 33909493–34027493	1	$5.15 \times 10^{-6}$	[0.118]	<i>AK162388</i> , <i>C230014012Rik</i>
MCcovMT	chr2: 54244165–54348111	1 (3)	$6.20 \times 10^{-6}$	0.104	/
TcovF	chr3: 100168735–100360505	1 (9)	$6.63 \times 10^{-6}$	0.192	<i>Gdap2</i>
HcovT	chr3: 130829175–130833777	1 (2)	$6.90 \times 10^{-6}$	0.005 [0.118]	/
HcovF	chr3: 154809225–154926757	1 (3)	$8.46 \times 10^{-6}$	0.118	<i>Tnni3k</i>
UcovF	chr5: 8672123–8674096	1 (2)	$2.48 \times 10^{-6}$	0.002 [0.118]	<i>Abcb1a</i> , <i>Rundc3b</i>
FcovT	chr6: 44025034–44505141	1 (3)	$4.17 \times 10^{-8}$	0.48	/
FcovH	chr6: 102314940–102422804	1 (3)	$7.91 \times 10^{-6}$	0.108	<i>4930587E11Rik</i> , <i>Cntn3</i>
FcovT	chr6: 134066708–135439926	6 (37)	$6.77 \times 10^{-9}$	1.373	<i>2810454H06Rik</i> , <i>Apold1</i> , <i>Gprc5d</i> , <i>Gm19434</i> , <i>Gsg1</i> , <i>Pbp2</i> , <i>Bcl2l14</i> , <i>Borcs5</i> , <i>Crebl2</i> , <i>Cdkn1b</i> , <i>Lockd</i> , <i>Ddx47</i> , <i>Gprc5a</i> , <i>Hebp1</i> , <i>Fam234b</i> , <i>Gsg1</i> , <i>Emp1</i> , <i>Lrp6</i> , <i>Mansc1</i> , <i>Dusp16</i> , <i>Gpr19</i> , <i>Etv6</i>
FcovT	chr6: 139526562–139581085	3 (8)	$2.77 \times 10^{-7}$	0.055 [0.118]	<i>Rergl</i> , <i>Pik3c2g</i>
HcovF	chr6: 144869562–144987562	1	$1.03 \times 10^{-6}$	[0.118]	/
FcovH		1	$8.08 \times 10^{-6}$		
TcovF	chr6: 146155662–146273662	1	$7.51 \times 10^{-6}$	[0.118]	<i>Itpr2</i> <sup>f*</sup>
HcovF	chr6: 146457299–146575299	1	$1.94 \times 10^{-6}$	[0.118]	<i>Ints13</i> ( <i>Asun</i> ), <i>Itpr2</i> <sup>f*</sup>
TcovH	chr7: 34695176–34813176	1	$4.19 \times 10^{-6}$	[0.118]	<i>Chst8</i>
FcovU	chr7: 43234461–43352461	1	$5.22 \times 10^{-7}$	[0.118]	<i>Zfp715</i> , <i>Siglecfl</i>
FcovT	chr8: 45416400–45534400	1	$5.67 \times 10^{-6}$	[0.118]	<i>Sorbs2</i> <sup>f**</sup>
TcovF	chr12: 67683133–68174899	1 (27)	$3.11 \times 10^{-6}$	0.492	/
FcovH	chr12: 86591551–86709551	1	$5.48 \times 10^{-6}$	[0.118]	<i>Vash1</i> , <i>Angel1</i>
TcovH	chr15: 44447489–44565489	1	$3.51 \times 10^{-6}$	[0.118]	<i>Pkhd111</i>
FcovU	chr15: 59470046–59482266	1 (3)	$7.14 \times 10^{-7}$	0.012 [0.118]	<i>Nsmce2</i> , <i>AK080559</i>
FcovU	chr15: 72720799–72733036	1 (2)	$6.61 \times 10^{-6}$	0.012 [0.118]	<i>Trappc9</i>
UcovT	chr15: 75470303–76162911	2 (37)	$3.29 \times 10^{-6}$	0.693	<i>Ly6 h</i> , <i>Gpihbp1</i> , <i>Zfp41</i> , <i>Mafa</i> , <i>Gsdmd</i> , <i>Mroh6</i> , <i>Naprt</i> , <i>Tigd5</i> , <i>Pycl</i> , <i>Tsta3</i> , <i>Zfp623</i> , <i>Ccdc166</i> , <i>Mapk15</i> , <i>Mir6952</i> , <i>Top1mt</i> , <i>Rhpn1</i> , <i>Eef1d</i> , <i>Zfp707</i> , <i>Fam83 h</i> , <i>Scrib</i> <sup>f***</sup> , <i>Puf60</i> , <i>Nrbp2</i> , <i>Eppk1</i> , <i>BC024139</i> , <i>Zc3 h3</i>
UcovT	chr15: 76148236–76266236	1	$8.84 \times 10^{-8}$	[0.118]	<i>Mir1942</i> , <i>Grina</i> , <i>Mir6953</i> , <i>Plec</i> , <i>Parp10</i>
HcovF	chr16: 92988497–93106497	1	$3.55 \times 10^{-6}$	[0.118]	/

(Continues)



TABLE 4 (Continued)

Bone <sup>a</sup>	Region	N of SNPs <sup>b</sup>	p-value of the best SNP <sup>c</sup>	Size (Mb) <sup>d</sup>	Genes in the region <sup>e</sup>
HcovU	chr16: 93764330–94205101	4 (5)	<b>8.30 × 10<sup>-7</sup></b>	0.441	<b>Chaf1b, Cldn14, Dopey2, Morc3, Sim2, Hlcs, AK009785</b>
HcovF		2 (5)	8.98 × 10 <sup>-7</sup>		
HcovT		3 (5)	1.72 × 10 <sup>-6</sup>		
HcovF	chr17: 30935073–31053073	1	<b>3.49 × 10<sup>-7</sup></b>	[0.118]	<b>Umod1, Glp1r, AK138161</b>
HcovF	chr17: 31860580–32474832	6 (17)	5.47 × 10 <sup>-6</sup>	0.614	<b>Rrp1b, Ephx3, Gm4432, A530088E08Rik, Pglyrp2, Pdxk-ps, Notch3<sup>f***</sup>, Akap8, Akap8l, Rasal3, Cyp4f39, Hsf2bp, Brd4, Wiz</b>
HcovU	chr19: 56931261–56931374	1 (2)	3.45 × 10 <sup>-6</sup>	1.13 × 10 <sup>-4</sup>	<b>Afap1l2, Vwa2</b>
UcovH		1 (2)	5.90 × 10 <sup>-6</sup>		

<sup>a</sup>Annotation: first capital letter stands for the bone tested, second capital letter for the one used as a covariate.

<sup>b</sup>Number of SNPs in the region with  $p < 10^{-5}$ , in parentheses number of total SNPs that are in LD within the same region.

<sup>c</sup>p-values in bold represent the SNPs above the permutation-based threshold.

<sup>d</sup>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.

<sup>e</sup>Annotated genes overlapping with the region; genes involved in limb development, phenotype or expression are highlighted in bold.

<sup>f</sup>Overlap with previous QTL studies in limb length: \*Norgard et al., 2011; \*\*Pavlicev et al., 2013; \*\*\*Kenney-Hunt et al., 2006.

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 programme, but subspecies may be too closely related to show such constraints, and 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., Norgard et al., 2008, 2009; Parmenter et al., 2016; Pavlicev et al., 2008, 2013; Wolf et al., 2006). 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).

## 5 | 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.

## AUTHOR CONTRIBUTIONS

The study was designed by N.S., L.F.P. and D.T.; the mice and their genotypes were provided by L.M.Z and B.H.; N.S. performed the analyses with input from B.H. and L.M.T.; N.S. and D.T. wrote the manuscript with input from B.H. and L.M.T. and L.F.P.

## DATA ACCESSIBILITY

Files for raw limb measures, age corrected measures, the kinship matrix and mapping files are deposited in Dryad under <https://doi.org/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).

## ORCID

Diethard Tautz  <http://orcid.org/0000-0002-0460-5344>

## REFERENCES

- Andrades, J. A., Nimni, M. E., Becerra, J., Eisenstein, R., Davis, M., & Sorgente, N. (1996). Complement proteins are present in developing endochondral bone and may mediate cartilage cell death and vascularization. *Experimental Cell Research*, 227, 208–213. <https://doi.org/10.1006/excr.1996.0269>
- Bab, I., Gabet, Y., Hajbi-Yonissi, C., & Müller, R. (2007). *Micro-Tomographic atlas of the mouse skeleton*. Springer Verlag. <https://doi.org/10.1007/978-0-387-39258-5>.
- Baroy, T., Kresse, S. H., Skarn, M., Stabell, M., Castro, R., Lauvrak, S., ... Meza-Zepeda, L. A. (2014). Reexpression of LSAMP inhibits tumor growth in a preclinical osteosarcoma model. *Molecular Cancer*, 13, 93. <https://doi.org/10.1186/1476-4598-13-93>
- Berenos, C., Ellis, P. A., Pilkington, J. G., Lee, S. H., Gratten, J., & Pemberton, J. M. (2015). Heterogeneity of genetic architecture of body size traits in a free-living population. *Molecular Ecology*, 24, 1810–1830. <https://doi.org/10.1111/mec.13146>
- Blake, J. A., Eppig, J. T., Kadin, J. A., Richardson, J. E., Smith, C. L., Bult, C. J., & Mouse Genome Database Group (2017). Mouse Genome Database (MGD)-2017: Community knowledge resource for the laboratory mouse. *Nucleic Acids Research*, 45, D723–D729. <https://doi.org/10.1093/nar/gkw1040>
- Bush, W. S., & Moore, J. H. (2012). Chapter 11: Genome-wide association studies. *Plos Computational Biology*, 8, e1002822. <https://doi.org/10.1371/journal.pcbi.1002822>
- Capdevila, J., & Belmonte, J. C. I. (2000). Perspectives on the evolutionary origin of tetrapod limbs. *Journal of Experimental Zoology*, 288, 287–303. [https://doi.org/10.1002/\(ISSN\)1097-010X](https://doi.org/10.1002/(ISSN)1097-010X)
- Cardelli, M., Zirngibl, R. A., Boetto, J. F., McKenzie, K. P., Troy, T. C., Turksen, K., & Aubin, J. E. (2013). Cartilage-specific overexpression of ERR gamma results in chondrodysplasia and reduced chondrocyte proliferation. *PLoS ONE*, 8, e81511. <https://doi.org/10.1371/journal.pone.0081511>
- Cheverud, J. M., Ehrich, T. H., Vaughn, T. T., Koreishi, S. F., Linsey, R. B., & Pletscher, L. S. (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. [https://doi.org/10.1002/\(ISSN\)1097-010X](https://doi.org/10.1002/(ISSN)1097-010X)
- Claude, J. (2008). *Morphometrics with R*. New York: Springer Verlag.
- Cooper, K. L., Sears, K. E., Uygur, A., Maier, J., Baczkowski, K. S., Brosnahan, M., ... Tabin, C. J. (2014). Patterning and post-patterning modes of evolutionary digit loss in mammals. *Nature*, 511, 41–U537. <https://doi.org/10.1038/nature13496>
- Deary, I. J., Yang, J., Davies, G., Harris, S. E., Tenesa, A., Liewald, D., ... Redmond, P. (2012). Genetic contributions to stability and change in intelligence from childhood to old age. *Nature*, 482, 212–215. <https://doi.org/10.1038/nature10781>
- Flint, J., & Eskin, E. (2012). Genome-wide association studies in mice. *Nature Reviews Genetics*, 13, 807–817. <https://doi.org/10.1038/nrg3335>
- 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. <https://doi.org/10.1038/nrg1576>
- Guenet, J. L., & Bonhomme, F. (2003). Wild mice: An ever-increasing contribution to a popular mammalian model. *Trends in Genetics*, 19, 24–31. [https://doi.org/10.1016/S0168-9525\(02\)00007-0](https://doi.org/10.1016/S0168-9525(02)00007-0)
- Hall, B. K. (1995). Homology and embryonic-development. *Evolutionary Biology*, 28(28), 1–37.
- Hallgrímsson, B., Willmore, K., & Hall, B. K. (2002). Canalization, developmental stability, and morphological integration in primate limbs. *Yearbook of Physical Anthropology*, 45(45), 131–158. [https://doi.org/10.1002/\(ISSN\)1096-8644](https://doi.org/10.1002/(ISSN)1096-8644)
- Hansen, T. F. (2006). The evolution of genetic architecture. *Annual Review of Ecology Evolution and Systematics*, 37, 123–157. <https://doi.org/10.1146/annurev.ecolsys.37.091305.110224>
- Hansen, T. F., Armbruster, W. S., Carlson, M. L., & 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. <https://doi.org/10.1002/jez.b.14>
- Harr, B., Karakoc, E., Neme, R., Teschke, M., Pfeifle, C., Pezer, Ž., ... Abai, M. R. (2016). Genomic resources for wild populations of the house mouse, *Mus musculus* and its close relative *Mus spretus*. *Scientific Data*, 3, 160075. <https://doi.org/10.1038/sdata.2016.75>
- Jadhav, G., Teguh, D., Kenny, J., Tickner, J., & Xu, J. K. (2016). Morc3 mutant mice exhibit reduced cortical area and thickness, accompanied by altered haematopoietic stem cells niche and bone cell differentiation. *Scientific Reports*, 6, 25964. <https://doi.org/10.1038/sre.p25964>
- Joeng, K. S., Schumacher, C. A., Zylstra-Diegel, C. R., Long, F. X., & Williams, B. O. (2011). Lrp5 and Lrp6 redundantly control skeletal development in the mouse embryo. *Developmental Biology*, 359, 222–229. <https://doi.org/10.1016/j.ydbio.2011.08.020>
- Kang, H. M., Sul, J. H., Zaitlen, N. A., Kong, S. Y., Freimer, N. B., Sabatti, C., & Eskin, E. (2010). Variance component model to account for sample structure in genome-wide association studies. *Nature Genetics*, 42, 348–U110. <https://doi.org/10.1038/ng.548>
- Kenney-Hunt, J. P., Vaughn, T. T., Pletscher, L. S., Peripato, A., Routman, E., Cothran, K., ... Cheverud, J. M. (2006). Quantitative trait loci for body size components in mice. *Mammalian Genome*, 17, 526–537. <https://doi.org/10.1007/s00335-005-0160-6>
- Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M., & Haussler, D. (2002). The human genome browser at UCSC. *Genome Research*, 12, 996–1006. <https://doi.org/10.1101/gr.229102>
- Kolarov, N. T., Ivanovic, A., & Kalezić, M. L. (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. <https://doi.org/10.1002/jez.b.21401>
- Lande, R. (1980). The genetic covariance between characters maintained by pleiotropic mutations. *Genetics*, 94, 203–215.
- Laurie, C. C., Nickerson, D. A., Anderson, A. D., Weir, B. S., Livingston, R. J., Dean, M. D., ... Nachman, M. W. (2007). Linkage disequilibrium in wild mice. *Plos Genetics*, 3, 1487–1495.
- Logan, M. (2003). Finger or toe: The molecular basis of limb identity. *Development*, 130, 6401–6410. <https://doi.org/10.1242/dev.00956>
- Martin-Serra, A., Figueirido, B., Perez-Claros, J. A., & Palmqvist, P. (2015). Patterns of morphological integration in the appendicular skeleton of mammalian carnivores. *Evolution*, 69, 321–340. <https://doi.org/10.1111/evo.12566>
- Moore, T. Y., Organ, C. L., Edwards, S. V., Biewener, A. A., Tabin, C. J., Jenkins, F. A. Jr, & Cooper, K. L. (2015). Multiple phylogenetically distinct events shaped the evolution of limb skeletal morphologies associated with bipedalism in the jerboas. *Current Biology*, 25, 2785–2794. <https://doi.org/10.1016/j.cub.2015.09.037>
- Norgard, E. A., Jarvis, J. P., Roseman, C. C., Maxwell, T. J., Kenney-Hunt, J. P., Samocha, K. E., ... Wolf, J. B. (2009). Replication of long-bone length QTL in the F-9-F-10 LG, SM advanced intercross. *Mammalian Genome*, 20, 224–235. <https://doi.org/10.1007/s00335-009-9174-9>
- Norgard, E. A., Lawson, H. A., Pletscher, L. S., Wang, B., Brooks, V. R., Wolf, J. B., & Cheverud, J. M. (2011). Genetic factors and diet affect long-bone length in the F-34 LG, SM advanced intercross. *Mammalian Genome*, 22, 178–196. <https://doi.org/10.1007/s00335-010-9311-5>
- Norgard, E. A., Roseman, C. C., Fawcett, G. L., Pavličev, M., Morgan, C. D., Pletscher, L. S., ... Cheverud, J. M. (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. <https://doi.org/10.1359/jbmr.080210>



- Pallares, L. F., Carbonetto, P., Gopalakrishnan, S., Parker, C. C., Ackert-Bicknell, C. L., Palmer, A. A., & Tautz, D. (2015). Mapping of craniofacial traits in outbred mice identifies major developmental genes involved in shape determination. *Plos Genetics*, *11*, e1005607. <https://doi.org/10.1371/journal.pgen.1005607>
- Pallares, L. F., Harr, B., Turner, L. M., & 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. <https://doi.org/10.1111/mec.12968>
- Pallares, L. F., Turner, L. M., & 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. <https://doi.org/10.1007/s00427-016-0550-7>
- Parmenter, M. D., Gray, M. M., Hogan, C. A., Ford, I. N., Broman, K. W., Vinyard, C. J., & Payseur, B. A. (2016). Genetics of skeletal evolution in unusually large mice from Gough Island. *Genetics*, *204*, 1559–1572. <https://doi.org/10.1534/genetics.116.193805>
- Pavlicev, M., Kenney-Hunt, J. P., Norgard, E. A., Roseman, C. C., Wolf, J. B., & Cheverud, J. M. (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, G. P., Noonan, J. P., Hallgrímsson, B., & Cheverud, J. M. (2013). Genomic correlates of relationship QTL involved in Fore- versus hind limb divergence in mice. *Genome Biology and Evolution*, *5*, 1926–1936. <https://doi.org/10.1093/gbe/evt144>
- Phifer-Rixey, M., & Nachman, M. W. (2015). Insights into mammalian biology from the wild house mouse *Mus musculus*. *eLife*, *4*, e05959. <http://doi.org/10.7554/eLife.05959>
- Price, A. L., Zaitlen, N. A., Reich, D., & Patterson, N. (2010). New approaches to population stratification in genome-wide association studies. *Nature Reviews Genetics*, *11*, 459–463. <https://doi.org/10.1038/nrg2813>
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., ... Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, *81*, 559–575. <https://doi.org/10.1086/519795>
- R Core Team (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Rieseberg, L. H., & Buerkle, C. A. (2002). Genetic mapping in hybrid zones. *American Naturalist*, *159*, S36–S50. <https://doi.org/10.1086/338371>
- Rolian, C., Lieberman, D. E., & Hallgrímsson, B. (2010). The coevolution of human hands and feet. *Evolution*, *64*, 1558–1568. <https://doi.org/10.1111/j.1558-5646.2009.00944.x>
- Schlötterer, C. (2015). Genes from scratch - the evolutionary fate of de novo genes. *Trends in Genetics*, *31*, 215–219. <https://doi.org/10.1016/j.tig.2015.02.007>
- Schmidt, M., & Fischer, M. S. (2009). Morphological integration in mammalian limb proportions: dissociation between function and development. *Evolution*, *63*, 749–766. <https://doi.org/10.1111/j.1558-5646.2008.00583.x>
- Schoengraf, P., Lambris, J. D., Recknagel, S., Kreja, L., Liedert, A., Brenner, R. E., ... Ignatius, A. (2013). Does complement play a role in bone development and regeneration? *Immunobiology*, *218*, 1–9. <https://doi.org/10.1016/j.imbio.2012.01.020>
- Schunke, A. C., Bromiley, P. A., Tautz, D., & Thacker, N. A. (2012). TINA manual landmarking tool: Software for the precise digitization of 3D landmarks. *Frontiers in Zoology*, *9*, 6. <https://doi.org/10.1186/1742-9994-9-6>
- Sheeba, C. J., Andrade, R. P., & 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. <https://doi.org/10.1016/j.semdb.2015.01.007>
- Shi, Y. C., Lin, S., Wong, I. P. L., Baldock, P. A., Aljanova, A., Enriquez, R. F., ... Macia, L. (2010). NPY neuron-specific Y2 receptors regulate adipose tissue and trabecular bone but not cortical bone homeostasis in mice. *PLoS ONE*, *5*, e11361. <https://doi.org/10.1371/journal.pone.0011361>
- Shubin, N., Tabin, C., & Carroll, S. (1997). Fossils, genes and the evolution of animal limbs. *Nature*, *388*, 639–648. <https://doi.org/10.1038/41710>
- Sul, J. H., Bilow, M., Yang, W. Y., Kostem, E., Furlotte, N., He, D., & Eskin, E. (2016). Accounting for population structure in gene-by-environment interactions in genome-wide association studies using mixed models. *Plos Genetics*, *12*, e1005849. <https://doi.org/10.1371/journal.pgen.1005849>
- Tautz, D., & Domazet-Loso, T. (2011). The evolutionary origin of orphan genes. *Nature Reviews Genetics*, *12*, 692–702. <https://doi.org/10.1038/nrg3053>
- Tickle, C. (2006). Making digit patterns in the vertebrate limb. *Nature Reviews Molecular Cell Biology*, *7*, 45–53. <https://doi.org/10.1038/nrm1830>
- Tsutsumi, R., Tran, M. P., & Cooper, K. L. (2017). Changing while staying the same: preservation of structural continuity during limb evolution by developmental integration. *Integrative and Comparative Biology*, *57*, 1269–1280. <https://doi.org/10.1093/icb/ix092>
- Turner, L. M., & Harr, B. (2014). Genome-wide mapping in a house mouse hybrid zone reveals hybrid sterility loci and Dobzhansky-Muller interactions. *Elife*, *3*, e02504. <http://doi.org/10.7554/eLife.02504>
- Turner, L. M., Schwahn, D. J., & 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. <https://doi.org/10.1111/j.1558-5646.2011.01445.x>
- Visscher, P. M., Hemani, G., Vinkhuyzen, A. A. E., Chen, G. B., Lee, S. H., Wray, N. R., ... Yang, J. (2014). Statistical power to detect genetic (co)variance of complex traits using SNP data in unrelated samples. *Plos Genetics*, *10*, e1004269. <https://doi.org/10.1371/journal.pgen.1004269>
- Wheeler, H. E., Shah, K. P., Brenner, J., Garcia, T., Aquino-Michaels, K., Cox, N. J., ... GTEx Consortium (2016). Survey of the heritability and sparse architecture of gene expression traits across human tissues. *Plos Genetics*, *12*, e1006423. <https://doi.org/10.1371/journal.pgen.1006423>
- Wolf, J. B., Pomp, D., Eisen, E. J., Cheverud, J. M., & Leamy, L. J. (2006). The contribution of epistatic pleiotropy to the genetic architecture of covariation among polygenic traits in mice. *Evolution & Development*, *8*, 468–476. <https://doi.org/10.1111/j.1525-142X.2006.00120.x>
- Wray, N. R., Yang, J., Hayes, B. J., Price, A. L., Goddard, M. E., & Visscher, P. M. (2013). Pitfalls of predicting complex traits from SNPs. *Nature Reviews Genetics*, *14*, 507–515. <https://doi.org/10.1038/nrg3457>
- Yang, J. A., Benyamin, B., McEvoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R., ... Goddard, M. E. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics*, *42*, 565–U131. <https://doi.org/10.1038/ng.608>
- Yang, H., Ding, Y. M., Hutchins, L. N., Szatkiewicz, J., Bell, T. A., Paigen, B. J., ... Churchill, G. A. (2009). A customized and versatile high-density genotyping array for the mouse. *Nature Methods*, *6*, 663–U655. <https://doi.org/10.1038/nmeth.1359>
- Yang, J. A., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2011). GCTA: A tool for genome-wide complex trait analysis. *American Journal of Human Genetics*, *88*, 76–82. <https://doi.org/10.1016/j.ajhg.2010.11.011>
- Yang, J., Manolio, T. A., Pasquale, L. R., Boerwinkle, E., Caporaso, N., Cunningham, J. M., ... Hill, W. G. (2011). Genome partitioning of genetic variation for complex traits using common SNPs. *Nature Genetics*, *43*, 519–U544. <https://doi.org/10.1038/ng.823>
- Yezerinac, S. M., Lougheed, S. C., & Handford, P. (1992). Measurement error and morphometric studies – statistical power and observer

- experience. *Systematic Biology*, 41, 471–482. <https://doi.org/10.1093/sysbio/41.4.471>
- Young, N. M. (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, N. M., & Hallgrímsson, B. (2005). Serial homology and the evolution of mammalian limb covariation structure. *Evolution*, 59, 2691–2704. <https://doi.org/10.1111/j.0014-3820.2005.tb00980.x>
- Zacco, A., Cooper, V., Chantler, P. D., Fisher-Hyland, S., Horton, H. L., & Levitt, P. (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, e1003264. <https://doi.org/10.1371/journal.pgen.1003264>
- Zhou, X., & Stephens, M. (2012). Genome-wide efficient mixed-model analysis for association studies. *Nature Genetics*, 44, 821–U136. <https://doi.org/10.1038/ng.2310>
- Zhou, X., & Stephens, M. (2014). Efficient multivariate linear mixed model algorithms for genome-wide association studies. *Nature Methods*, 11, 407–409. <https://doi.org/10.1038/nmeth.2848>

#### SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

**How to cite this article:** Škrabar N, Turner LM, Pallares LF, Harr B, Tautz D. Using the *Mus musculus* hybrid zone to assess covariation and genetic architecture of limb bone lengths. *Mol Ecol Resour*. 2018;00:1–14. <https://doi.org/10.1111/1755-0998.12776>

## Contributions to the thesis

### Chapter three

This chapter was published in the internationally peer reviewed journal *Molecular Ecology Resources*.

Citation: Škrabar N, Turner LM, Pallares LF, Harr B, Tautz D. 2018. Using the *Mus musculus* hybrid zone to assess covariation and genetic architecture of limb bone lengths. *Mol Ecol Resour.* Mar 9. doi: 10.1111/1755-0998.12776. [Epub ahead of print]

The study was designed by me together with D. Tautz and L. F. Pallares. I collected the phenotypic data, performed the analyses, and wrote the first version of the manuscript. B. Harr and L. M. Turner provided the samples (mice from the hybrid zone) and the genotypes. B. Harr and L. M. Turner assisted with the analyses. B. Harr and L. M. Turner, L. F. Pallares and D. Tautz assisted with writing the final version.

## Perspectives

Anatomical structures in vertebrates impose numerous challenges in understanding the complexity that underlie their proper development, function and evolution (Esteve-Altava et al. 2018). Rise of new technologies and approaches enabled a better understanding of different patterns of integration and modularity between body parts. A variety of traits, from fitness related to locomotory are prone to constant changes and responses to different selective pressures. However, the main patterning processes are not disturbed and allow coordinated alteration across characters. Morphometry found wide application in different fields and provided quantitative evaluation of bone remodeling, mainly related to medical research. Further, evolutionary importance was noticed in exploration of quantitative traits. Analyses of different phenotypes helped in understanding complex networks, mostly between development and function, but also continued a step further implementing information from high-throughput sequencing approaches and possibility to identify genetic regions which might be involved in trait variation. Previous genome-wide association studies were mostly oriented toward single phenotypes, while growing interest in connectivity between different structures incorporated multi-trait genetic analyses (Porter & O'Reily 2016). Association mapping with multiple traits increased the power to detect variants affecting correlated traits (Furlotte & Eskin 2015). The inclusion of outbred lab mice, as well as close to natural hybrid zone populations provide numerous advantages in resolving these interesting interactions between morphological characters, and also open new questions.



## Acknowledgments

Most of all, I would like to thank to my supervisor Prof. Dr. Diethard Tautz, who made all of this work possible. Thank you for guiding me through many difficulties, for great discussions, comments, constructive critics, motivation, freedom and independence, as well as for a lot of understanding, patience, time and support.

I would like to thank Prof. Dr. Thomas Bosch and Prof Dr. Arne Nolte for their input, suggestions and feedback I have received in the beginning of my doctoral studies.

I am deeply grateful to my coauthors Bettina Harr, Leslie M. Turner and Luisa F. Pallares for all the feedback, generous help, support, time and wonderful collaboration. With special thanks to Bettina Harr and Leslie M. Turner for providing the mice and for answering numerous questions I had.

Many thanks to Elke Blohm-Sievers and Bastian Poerschke for their kind help with scanning the mice, as well as for the great friendship and time outside of work.

I would like to thank to the entire mouse team, with special thanks to Christine Pfeifle, Heike Harre, Maik Görtz-Sonnwald, Anika Jonas, Susanne Reinsch, Janine Wolf, Camilo Medina and Till Sckerl, who generously helped me with the mouse care, otherwise it would not be possible to finish everything in time. I also appreciate the great communication we had during the entire project.

I would like to thank to my colleagues Anja Schunke, Kira Delmore, Maryam Keshavarz, Derek Caetano-Anolles, Kristian Ullrich, Guy Reeves, Wen Yu Zhang, Lutz Becks, Cemalettin Bekpen, Chen Xie, Sven Künzel, Peter Refki, Sophie von Merten, Predrag Kalajdžić for many valuable discussions, input and feedback they gave during my doctoral studies.

Thank to Prof. Dr. Gabriel Marroig for helpful comments and kind explanations considering some questions in the second chapter of the thesis.

Many thanks to my office mates and wonderful friends Rebecca Krebs-Wheaton and Stefan Dennenmoser for all the support, help, advice, discussions and many insightful comments.

I am thankful to Arne Nolte and Stefan Dennenmoser for generously providing medaka fish which I first used for experimental purpose and a lot of help from Ralf Schmuck with their care.

I would like to say thank you to all the colleagues from the Max Planck Institute for Evolutionary Biology in Plön for a wonderful atmosphere and time, their willingness to help and their kindness. My time at the Institute was filled with interesting seminars, workshops, talks as well as many great events.

I am deeply thankful to my friends, Marija, Milica, Goga, Bojana, Ana and Milan for all the love and invaluable friendship, for being with me in many moments despite the distance. Federica, Ana, Malavi, Ela, Mayra, Alina, Loukas, Luka, Michael, Guénolé, Juan, Jatin, Wiola, Johana, Devika, Gillian, Filipa, Dušica, thank you for all the wonderful and unforgettable moments, travels, parties, talks, advices, solving many difficulties and always being supportive.

Special thanks to my partner Robert Bakarić for love, constant encouragement, being with me in stressful moments, for incredible understanding, many scientific discussions and sharing knowledge with me.

I am especially grateful to my parents Jasenka and Stane and to my brother Bojan for unconditional love, strength, concern, motivation, supporting my endeavor towards science and the pursuit of my dreams.

## References

- Ackermann, R. R. (2005). Ontogenetic integration of the hominoid face. *J Hum Evol*, *48*, 175-197.
- Ackermann, R. R., & Cheverud, J. M. (2000). Phenotypic covariance structure in tamarins (genus *Saguinus*): a comparison of variation patterns using matrix correlation and common principal component analysis. *Am J Phys Anthropol*, *111*, 489-501.
- Albertson, R. C., & Kocher, T. D. (2005). Genetic architecture sets limits on transgressive segregation in hybrid cichlid fishes. *Evolution*, *59*, 686-690.
- Albrechtova, J., Albrecht, T., Baird, S. J., Macholan, M., Rudolfsen, G., Munclinger, P., et al. (2012). Sperm-related phenotypes implicated in both maintenance and breakdown of a natural species barrier in the house mouse. *Proc Biol Sci*, *279*, 4803-4810.
- Aldinger, K. A., Sokoloff, G., Rosenberg, D. M., Palmer, A. A., & Millen, K. J. (2009). Genetic variation and population substructure in outbred CD-1 mice: implications for genome-wide association studies. *PLoS One*, *4*, e4729.
- Alibert, P., Auffray, J-C. (2003). Genomic coadaptation, outbreeding depression, and developmental instability. In: *Developmental instability: causes and consequences* (ed. Polak, M.), pp. 116-134. Oxford University Press, Oxford
- Alibert, P., Fel-Clair, F., Manolakou, K., Britton-Davidian, J., Auffray, J-C. (1997). Developmental stability, fitness, and trait size in laboratory hybrids between European subspecies of the house mouse. *Evolution*, *51*, 1284-1295.
- Alibert, P., Renaud, S., Dod, B., Bonhomme, F., Auffray, J-C. (1994). Fluctuating asymmetry in the *Mus musculus* hybrid zone: a heterotic effect in disrupted co-adapted genomes. *Proceedings of the Royal Society B – Biological Sciences*, *258*, 53-59.
- Arnold, B.C., Groeneveld, R.A. (1995). Measuring skewness with respect to the mode. *The American Statistician*, *49*, 34-38.
- Atchley, W. R., Hall, B. K. (1991). A model for development and evolution of complex morphological structures. *Biological Reviews*, *66*, 101-157.
- Atchley, W. R., Rutledge, J. J., & Cowley, D. E. (1982). A multivariate statistical analysis of direct and correlated response to selection in the rat. *Evolution*, *36*, 677-698.
- Auerbach, B. M., Ruff, C. B. (2006). Limb bone bilateral asymmetry: variability and commonality among modern humans. *J Hum Evol*, *50*, 203-218.
- Bab, I., Gabet, Y., Hajbi-Yonissi, C., & Müller, R. (2007). *Micro-Tomographic atlas of the mouse skeleton*. Springer Verlag.
- Bader, R. S. (1965). Fluctuating asymmetry in the dentition of the house mouse. *Growth*, *29*, 291-300.
- Baird, S.J.E., & Macholán, M. (2012). What can the *Mus musculus musculus/M. m. domesticus* hybrid zone tell us about speciation? In: *Evolution of the house mouse*, (eds. Macholán, M., Baird, S.J.E., Munclinger, P., & Piálek J.), pp. 334-372. Cambridge studies in morphology and molecules: new paradigms in evolutionary biology. Cambridge University Press, Cambridge.
- Balmford, A., Jones, I.L., & Thomas, A.L.R. (1993). On avian asymmetry: evidence of natural selection for symmetrical tails and wings in bird. *Proc Roy Soc Lond Ser. B*, *252*, 245-251.

- Barton, N. H., Etheridge, A. M., Véber, A. (2017). The infinitesimal model: definition, derivation and implications. *Theor Pop Biol*, 118, 50-73.
- Barton, N. H., & Hewitt, G. M. (1985). Analysis of hybrid zones. *Annual Review of Ecology and Systematics*, 16, 113-148.
- Bell, E., Andres, B., & Goswami, A. (2011). Integration and dissociation of limb elements in flying vertebrates: a comparison of pterosaurs, birds and bats. *J Evol Biol*, 24, 2586-2599.
- Boag, P. T. (1983). The heritability of external morphology in Darwin's ground finches (*Geospiza*) on Isla Daphne Major, Galapagos. *Evolution*, 37, 877-894.
- Boell, L. (2013). Lines of least resistance and genetic architecture of house mouse (*Mus musculus*) mandible shape. *Evolution & Development*, 15, 197-204.
- Boell, L., Pallares, L. F., Brodski, C., et al. (2013). Exploring the effects of gene dosage on mandible shape in mice as a model for studying the genetic basis of natural variation. *Development Genes and Evolution*, 223, 279-287.
- Boell, L., Tautz, D. (2011). Micro-evolutionary divergence patterns of mandible shapes in wild house mouse (*Mus musculus*) populations. *BMC Evol Biol*, 11, 306.
- Boissinot, S., & Boursot, P. (1997). Discordant phylogeographic patterns between the Y chromosome and mitochondrial DNA in the house mouse: selection on the Y chromosome? *Genetics*, 146, 1019-1034.
- Bonhomme, F., Catalan, J., Britton-Davidian, J., Chapman, V. M., Moriwaitu, D., Nevo, E., Thaler, L. (1984). Biochemical diversity and evolution in the genus *Mus*. *Biochem Genet*, 22, 275-303.
- Boyle, E. A., Li, Y. I., Pritchard, J. K. (2017). An expanded view of complex traits: from polygenic to omnigenic. *Cell*, 169, 1177-1186.
- Britton-Davidian, J., Fel-Clair, F., Lopez, J., Alibert, P., & Boursot, P. (2005). Postzygotic isolation between the two European subspecies of the house mouse: estimates from fertility patterns in wild and laboratory-bred hybrids. *Biological Journal of the Linnean Society*, 84, 379-393.
- Brommer, J.E., Pihlajamaki, O., Kolunen, H. & Pietiainen, H. (2003). Life-history consequences of partial-moult asymmetry. *J Anim Ecol*, 72, 1057-1063.
- Bush, W. S., Moore, J. H. (2012). Chapter 11: Genome-Wide Association Studies. *Plos Computational Biology*, 8, e1002822.
- Butterfield, N.C., McGlinn, E., Wicking, C. (2010). The molecular regulation of vertebrate limb patterning. *Curr Top Dev Biol*, 90, 319-341.
- Cane, W. P. (1993). The ontogeny of postcranial integration in the common tern, *Sterna Hirundo*. *Evolution*, 47, 1138-1151.
- Carter, A. J. R., Weier, T. M., Houle, D. (2009). The effect of inbreeding on fluctuating asymmetry of wing veins in two laboratory strains of *Drosophila melanogaster*. *Heredity*, 102, 563-572.
- Casellas, J. (2011). Inbred mouse strains and genetic stability: a review. *Animal*, 5, 1-7.
- Chapin, R. E., Morrissey, R. E., Gulati, D. K., Hope, E., Barnes, L. H., Russell, S. A., & Kennedy, S. R. (1993). Are mouse strains differentially susceptible to the reproductive toxicity of ethylene glycol monomethyl ether? A study of three strains. *Fundam Appl Toxicol*, 21, 8-14.

- Cheverud, J. M. (1982). Phenotypic, genetic, and environmental morphological integration in the cranium. *Evolution*, *36*, 499-516.
- Cheverud, J. M. (1984). Quantitative genetics and developmental constraints on evolution by selection. *J Theor Biol*, *110*, 155-171.
- Cheverud, J. M. (1988). A comparison of genetic and phenotypic correlations. *Evolution*, *42*, 958-968.
- Cheverud, J. M. (1989). A comparative analysis of morphological variation patterns in the Papionins. *Evolution*, *43*, 1737-1747.
- Cheverud, J. M. (1995). Morphological integration in the saddle-back tamarin (*Saguinus-Fuscicollis*) Cranium. *American Naturalist*, *145*, 63-89.
- Cheverud, J. M. (1996a). Developmental integration and the evolution of pleiotropy. *American Zoologist*, *36*, 44-50.
- Cheverud, J. M. (1996b). Quantitative genetic analysis of cranial morphology in the cotton-top (*Saguinus oedipus*) and saddle-back (*S. fuscicollis*) tamarins. *Journal of Evolutionary Biology*, *9*, 5-42.
- Cheverud, J. M., & Leamy, L. J. (1985). Quantitative genetics and the evolution of ontogeny. III. Ontogenetic changes in correlation structure among live-body traits in randombred mice. *Genet Res*, *46*, 325-335.
- Cheverud, J. M., Leamy, L. J., Atchley, W. R., & Rutledge, J. J. (1983). Quantitative genetics and the evolution of ontogeny: I. Ontogenetic changes in quantitative genetic variance-components in randombred mice. *Genetical Research*, *42*, 65-75.
- Cheverud, J. M., & Marroig, G. (2007). Comparing covariance matrices: Random skewers method compared to the common principal components model. *Genetics and Molecular Biology*, *30*, 461-469.
- Cheverud, J. M., Rutledge, J. J., & Atchley, W. R. (1983). Quantitative genetics of development: genetic correlations among age-specific trait values and the evolution of ontogeny. *Evolution*, *37*, 895-905.
- Cheverud, J. M., Wagner, G. P., & Dow, M. M. (1989). Methods for the comparative analysis of variation patterns. *Systematic Zoology*, *38*, 201-213.
- Chia, R., Achilli, F., Festing, M. F., & Fisher, E. M. (2005). The origins and uses of mouse outbred stocks. *Nat Genet*, *37*, 1181-1186.
- Clarke, G. M. (1993). The genetic basis of developmental stability. I. Relationships between stability, heterozygosity and genomic coadaptation. *Genetica*, *89*, 15-23.
- Clarke, G. M., Oldroyd, B. P., & Hunt, P. (1992). The genetic basis of developmental stability in *Apis Mellifera*: Heterozygosity versus genic balance. *Evolution*, *46*, 753-762.
- Clarke, J. H., Mithen, R., Brown, J. K., & Dean, C. (1995). QTL analysis of flowering time in *Arabidopsis thaliana*. *Mol Gen Genet*, *248*, 278-286.
- Claude, J. (2008). *Morphometrics with R*. New York: Springer Verlag.
- Cloutier, R., & P. E. Ahlberg. (1996). Morphology, characters, and the interrelationships of basal sarcopterygians. In: *Interrelationships of fishes* (eds. Stiassny, M. L. J., Parenti, L. R., & Johnson, G. D.), pp. 325-337. Academic Press, London.

- Cock, A. G. (1966). Genetical aspects of metrical growth and form in animals. *Quarterly Review of Biology*, *41*, 131-190.
- Collins, A. L., & Sullivan, P. F. (2013). Genome-wide association studies in psychiatry: what have we learned? *Br J Psychiatry*, *202*, 1-4.
- Comeault, A. A., Soria-Carrasco, V., Gompert, Z., Farkas, T. E., Buerkle, C. A., Parchman, T. L., & Nosil, P. (2014). Genome-wide association mapping of phenotypic traits subject to a range of intensities of natural selection in *Timema cristinae*. *Am Nat*, *183*, 711-727.
- Cooper, K. L. (2011). The lesser Egyptian jerboa, *Jaculus jaculus*: a unique rodent model for evolution and development. *Cold Spring Harb Protoc*, *2011*, 1451-1456.
- D'Agostino, R. B. (1986). Tests for the normal distribution. In: *Goodness-of-fit-techniques* (eds. D'Agostino, R. B., & Stephens, M. A.), pp. 367-419. Marcel Dekker Inc., New York.
- Davis, A. P., & Capecchi, M. R. (1994). Axial homeosis and appendicular skeleton defects in mice with a targeted disruption of *hoxd-11*. *Development*, *120*, 2187-2198.
- Davis, A. P., & Capecchi, M. R. (1996). A mutational analysis of the 5' *HoxD* genes: dissection of genetic interactions during limb development in the mouse. *Development*, *122*, 1175-1185.
- Debat, V., Alibert, P., David, P., Paradis, E., Auffray, J-C. (2000). Independence between developmental stability and canalization in the skull of the house mouse. *Proc Biol Sci*, *267*, 423-430.
- Dixon, W., Mood, A. (1946). The statistical sign test. *Journal of American Statistical Association*, *41*, 557-566.
- Dobzhansky, T. (1950). Genetics of natural populations. XIX. Origin of heterosis through natural selection in populations of *Drosophila pseudoobscura*. *Genetics*, *35*, 288-302.
- Dobzhansky, T., & Spassky, B. (1968). Genetics of natural populations. XI. Heterotic and deleterious effects of recessive lethals in populations of *Drosophila pseudoobscura*. *Genetics*, *59*, 411-425.
- Dod, B., Jermiin, L. S., Boursot, P., Chapman, V. H., Tonnes-Nielsen, J., & Bonhomme, F. (1993). Counterselection on sex chromosomes in the *Mus musculus* European hybrid zone. *J. Evol. Biol.* *6*, 529-546.
- Dolle, P., Dierich, A., LeMeur, M., Schimmang, T., Schuhbaur, B., Chambon, P., & Duboule, D. (1993). Disruption of the *Hoxd-13* gene induces localized heterochrony leading to mice with neotenic limbs. *Cell*, *75*, 431-441.
- Dongen, S. V. (2006). Fluctuating asymmetry and developmental instability in evolutionary biology: past, present and future. *J Evol Biol*, *19*, 1727-1743.
- Dosselman, D. J., Schaalje, G. B., Sites, J. W. (1998). An analysis of fluctuating asymmetry in a hybrid zone between two chromosome races of the *Sceloporus grammicus* complex (Squamata: Phrynosomatidae) in central Mexico. *Herpetologica* *54*, 434-447.
- Emlen, D. J., Marangelo, J., Ball, B., & Cunningham, C. W. (2005). Diversity in the weapons of sexual selection: horn evolution in the beetle genus *Onthophagus* (Coleoptera: Scarabaeidae). *Evolution*, *59*, 1060-1084.
- Esteve-Altava, B., Molnar, J. L., Johnston, P., Hutchinson, J. R., & Diogo, R. (2018). Anatomical network analysis of the musculoskeletal system reveals integration loss and parcellation boost during the fins-to-limbs transition. *Evolution*, *72*, 601-618.

- Falconer, D. S. (1981). *Introduction to Quantitative Genetics*, Ed. 2. Longmans Green, London/New York.
- Farnum, C. E. (2007). Postnatal growth of fins and limbs through endochondral ossification. In: *Fins into Limbs* (ed. Hall, B. K.). University of Chicago Press, Chicago.
- Fel-Clair, F., Lenormand, T., Catalan, J., Grobert, J., Orth, A., Boursot, P., et al. (1996). Genomic incompatibilities in the hybrid zone between house mice in Denmark: evidence from steep and non-coincident chromosomal clines for Robertsonian fusions. *Genet Res*, *67*, 123-134.
- Festing, M. F. W. (1979). *Inbred strains in biomedical research*. Oxford University Press, New York.
- Festing, M. F. W. (1993). *International index of laboratory animals*. 6th edition. Newbury, England.
- Festing, M.F.W. (1996). Origins and characteristics of inbred strains of mice. In: *Genetic variants and strains of the laboratory mouse* (eds. Lyon, M. F., Rastan, S., & Brown, S. D. M.), pp. 1537-1576. Oxford University Press, Oxford.
- Fisher, R. A. (1950). Gene frequencies in a cline determined by selection and diffusion. *Biometrics*, *6*, 353-361.
- Flint, J., & Eskin, E. (2012). Genome-wide association studies in mice. *Nat Rev Genet*, *13*, 807-817.
- Flint, J., & Mackay, T. F. (2009). Genetic architecture of quantitative traits in mice, flies, and humans. *Genome Res*, *19*, 723-733.
- Freeman, D. C., Graham, J. H., Byrd, D. W., McArthur E. D., & Turner, W. A. (1995). Narrow hybrid zone between two subspecies of big sagebrush, *Artemisia tridentata* (Asteraceae). III. Developmental instability. *American Journal of Botany*, *82*, 1144-1152.
- Furlotte, N. A., & Eskin, E. (2015). Efficient multiple-trait association and estimation of genetic correlation using the matrix-variate linear mixed model. *Genetics*, *200*, 59-68.
- Gabriel, S. I., Stevens, M. I., Mathias Mda, L., & Searle, J. B. (2011). Of mice and 'convicts': origin of the Australian house mouse, *Mus musculus*. *PLoS One*, *6*, e28622.
- Galeotti, P., Sacchi, R., Vicario, V. (2005). Fluctuating asymmetry in body traits increases predation risks: tawny owl selection against asymmetric woodmice. *Evol Ecol*, *19*, 405-418.
- Gibson-Brown, J. J., Agulnik, S. I., Chapman, D. L., Alexiou, M., Garvey, N., Silver, L. M., & Papaioannou, V. E. (1996). Evidence of a role for T-box genes in the evolution of limb morphogenesis and the specification of forelimb/hindlimb identity. *Mech Dev*, *56*, 93-101.
- Gilbert, S. F. (2000a). *Developmental Biology*. 6th edition. The Questions of Developmental Biology. Sunderland (MA): Sinauer Associates.
- Gilbert, S. F. (2000b). *Developmental Biology*. 6th edition. Osteogenesis: The Development of Bones. Sunderland (MA): Sinauer Associates.
- Gonzales, N. M., & Palmer, A. A. (2014). Fine-mapping QTLs in advanced intercross lines and other outbred populations. *Mamm Genome*, *25*, 271-292.
- Good, J. M., Handel, M. A., & Nachman, M. W. (2008). Asymmetry and polymorphism of hybrid male sterility during the early stages of speciation in house mice. *Evolution*, *62*, 50-65.
- Graham, J. H. (1992). Genomic coadaptation and developmental stability in hybrid zones. *Acta Zool Fenn*, *191*, 121-131.

- Graham, J. H., & Felley, J. D. (1985). Genomic coadaptation and developmental stability within introgressed populations of *Enneacanthus Gloriosus* and *E. Obesus* (Pisces, Centrarchidae). *Evolution*, *39*, 104-114.
- Haber, A. (2011). A comparative analysis of integration indices. *Evolutionary Biology*, *38*, 476-488.
- Hagg, A. C. (2016). Assessment of skeletal and dental fluctuating asymmetry in two historic Dutch populations, MSc Dissertation, University of Pretoria, Pretoria, viewed 180601 <<http://hdl.handle.net/2263/56936>>
- Haldane, J. B. (1948). The theory of a cline. *J Genet*, *48*, 277-284.
- Hall, B. K. (1995). *Homology and embryonic development*. *Evolutionary Biology*, *28*, 1-37.
- Hall, B. K. (2007). *Fins into limbs : evolution, development, and transformation*. University of Chicago Press, Chicago.
- Hallgrímsson, B. (1993). Fluctuating asymmetry in *Macaca fascicularis*: Implications for the etiology of developmental noise. *Am J Phys Anthropol*, Suppl *16*, 102.
- Hallgrímsson, B. (1998). Fluctuating asymmetry in the mammalian skeleton: evolutionary and developmental implications. *Evolutionary Biology*, *30*, 187-251.
- Hallgrímsson, B. (1999). Ontogenetic patterning of skeletal fluctuating asymmetry in rhesus macaques and humans: evolutionary and developmental implications. *International Journal of Primatology* *20*, 121-151.
- Hallgrímsson, B., Brown, J. J., Ford-Hutchinson, A. F., Sheets, H. D., Zelditch, M. L., & Jirik, F. R. (2006). The brachymorph mouse and the developmental-genetic basis for canalization and morphological integration. *Evol Dev*, *8*, 61-73.
- Hallgrímsson, B., Dorval, C. J., Zelditch, M. L., & German, R. Z. (2004). Craniofacial variability and morphological integration in mice susceptible to cleft lip and palate. *J Anat*, *205*, 501-517.
- Hallgrímsson, B., Jamniczky, H., Young, N. M., Rolian, C., Parsons, T. E., Boughner, J. C., & Marcucio, R. S. (2009). Deciphering the palimpsest: Studying the relationship between morphological integration and phenotypic covariation. *Evol Biol*, *36*, 355-376.
- Hallgrímsson, B., Miyake, T., Wilmore, K., Hall, B. K. (2003). Embryological origins of developmental stability: size, shape and fluctuating asymmetry in prenatal random bred mice. *J Exp Zool B Mol Dev Evol*, *296*, 40-57.
- Hallgrímsson, B., Willmore, K., & Hall, B. K. (2002). Canalization, developmental stability, and morphological integration in primate limbs. *Yearbook of Physical Anthropology*, *45*, 131-158.
- Hardouin, E. A., Chapuis, J. L., Stevens, M. I., van Vuuren, J. B., Quillfeldt, P., Scavetta, R. J., et al. (2010). House mouse colonization patterns on the sub-Antarctic Kerguelen Archipelago suggest singular primary invasions and resilience against re-invasion. *BMC Evol Biol*, *10*, 325.
- Hardouin, E. A., Orth, A., Teschke, M., Darvish, J., Tautz, D., & Bonhomme, F. (2015). Eurasian house mouse (*Mus musculus* L.) differentiation at microsatellite loci identifies the Iranian plateau as a phylogeographic hotspot. *BMC Evol Biol*, *15*, 26.
- Hiadlovská, Z., Vošlajerová Bímová, B., Mikula, O., Piálek, J., Macholán, M. (2013). Transgressive segregation in a behavioural trait? Explorative strategies in two house mouse subspecies and their hybrids. *Biological Journal of the Linnean Society*, *108*, 225-235.



- Holmes, D. J. (2003). DBA/2 mouse. *Science of Aging Knowledge Environment*, 44, 3.
- Houle, D., & Meyer, K. (2015). Estimating sampling error of evolutionary statistics based on genetic covariance matrices using maximum likelihood. *J Evol Biol*, 28, 1542-1549.
- Ivanovic, A., Kalezic, M. L., & Aleksic, I. (2005). Morphological integration of cranium and postcranial skeleton during ontogeny of facultative paedomorphic European newts (*Triturus vulgaris* and *Triturus alpestris*). *Amphibia-Reptilia*, 26, 485-495.
- Jamniczky, H. A., & Hallgrímsson, B. (2009). A comparison of covariance structure in wild and laboratory murid crania. *Evolution*, 63, 1540-1556.
- Jolicoeur, P. (1963). The multivariate generalization of the allometry equation. *Biometrics*, 19, 497-499.
- Jones, E. P., Johannsdottir, F., Gunduz, I., Richards, M. B., & Searle, J. B. (2011). The expansion of the house mouse into north-western Europe. *Journal of Zoology*, 283, 257-268.
- Kenney-Hunt, J. P., Vaughn, T. T., Pletscher, L. S., Peripato, A., Routman, E., Cothran, K., et al. (2006). Quantitative trait loci for body size components in mice. *Mamm Genome*, 17, 526-537.
- Kim, S. (2015). ppcor: An R package for a fast calculation to semi-partial correlation coefficients. *Communications for Statistical Applications and Methods*, 22, 665-674.
- Kingsolver, J. G., & Koehl, M. A. R. (1994). Selective factors in the evolution of insect wings. *Annual Review of Entomology*, 39, 425-451.
- Klingenberg, C. P. (2008). Morphological integration and developmental modularity. *Annual Review of Ecology Evolution and Systematics*, 39, 115-132.
- Klingenberg, C. P., Leamy, L. J., & Cheverud, J. M. (2004). Integration and modularity of quantitative trait locus effects on geometric shape in the mouse mandible. *Genetics*, 166, 1909-1921.
- Klingenberg, C. P., & Nijhout, H. F. (1999). Genetics of fluctuating asymmetry: a developmental model of developmental instability. *Evolution*, 53, 358-375.
- Kochi, Y. (2016). Genetics of autoimmune diseases: perspectives from genome-wide association studies. *Int Immunol*, 28, 155-161.
- Kolarov, N. T., Cvijanovic, M., Denoel, M., & Ivanovic, A. (2017). Morphological integration and alternative life history strategies: A case study in a facultatively paedomorphic newt. *Journal of Experimental Zoology Part B-Molecular and Developmental Evolution*, 328, 737-748.
- Komsta, L., & Novomestky, F. (2015). moments: Moments, cumulants, skewness, kurtosis and related tests. R package version 0.14. 2015.
- Kronenberg, H. (2003). Developmental regulation of the growth plate. *Nature*, 423, 332-336.
- Lande, R. (1979). Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry. *Evolution*, 33, 402-416.
- Lande, R. (1980). The genetic covariance between characters maintained by pleiotropic mutations. *Genetics*, 94, 203-215.
- Leamy, L. (1974). Heritability of osteometric traits in a randombred population of mice. *J Hered*, 65, 109-120.
- Leamy, L. (1975). Component analysis of osteometric traits in randombred house mice. *Systematic Zoology*, 24, 176-190.

- Leamy, L. (1977). Genetic and environmental correlations of morphometric traits in randombred house mice. *Evolution*, *31*, 357-369.
- Leamy, L. (1984). Morphometric studies in inbred and hybrid mouse. V. Directional and fluctuating asymmetry. *American Naturalist*, *123*, 579-593.
- Leamy, L. (1992). Morphometric studies in inbred and hybrid mouse. VII. Heterosis in fluctuating asymmetry at different ages. *Acta Zoologica Fennica*, *191*, 111-120.
- Leamy, L. (1997). Genetic analysis of fluctuating asymmetry for skeletal characters in mice. *J Hered*, *88*, 85-92.
- Leamy, L. (1999). Heritability of directional and fluctuating asymmetry for mandible characters in randombred mice. *Journal of Evolutionary Biology*, *12*, 146-155.
- Leamy, L., & Bradley, D. (1982). Static and growth allometry of morphometric traits in randombred house mice. *Evolution*, *36*, 1200-1212.
- Leamy, L. J., Klingenberg, C. P., Sherratt, E., Wolf, J. B., Cheverud, J. M. (2015). The genetic architecture of fluctuating asymmetry of mandible size and shape in a population of mice: another look. *Symmetry*, *7*, 146-163.
- Leamy, L. J., Pomp, D., Eisen, E. J., Cheverud, J. M. (2000). Quantitative trait loci for directional but not fluctuating asymmetry of mandible characters in mice. *Genet Res Camb*, *76*, 27-40.
- Leamy, L. J., Pomp, D., Eisen, E. J., & Cheverud, J. M. (2002). Pleiotropy of quantitative trait loci for organ weights and limb bone lengths in mice. *Physiol Genomics*, *10*, 21-29.
- Leamy, L. J., Routman, E. J., & Cheverud, J. M. (1997). A search for quantitative trait loci affecting asymmetry of mandibular characters in mice. *Evolution*, *51*, 957-969.
- Leamy, L. J., Routman, E. J., & Cheverud, J. M. (1998). Quantitative trait loci for fluctuating asymmetry of discrete skeletal characters in mice. *Heredity (Edinb)*, *80*, 509-518.
- Leamy, L., & Sustarsic, S. S. (1978). A morphometric discriminant analysis of agouti genotypes in C57B/6 house mice. *Systematic Zoology* *27*, 49-60.
- Leary, R. F., Allendorf, F. W., & Knudsen, K. L. (1983). Developmental stability and enzyme heterozygosity in rainbow trout. *Nature*, *301*, 71-72.
- Leary, R. F., Allendorf, F. W., & Knudsen, K. L. (1984). Superior developmental stability of heterozygotes at enzyme loci in salmonid fishes. *American Naturalist*, *124*, 540-551.
- Leary, R. F., Allendorf, F. W., & Knudsen, K. L. (1985). Developmental instability and high meristic counts in interspecific hybrids of salmonid fishes. *Evolution*, *39*, 1318-1326.
- Leary, R. F., Allendorf, F. W., Knudsen, K. L., & Thorgaard, G. H. (1985). Heterozygosity and developmental stability in gynogenetic diploid and triploid rainbow trout. *Heredity (Edinb)*, *54*, 219-225.
- Lens, L., Van Dongen, S., Galbusera, P., Schenck, T., Matthysen, E., Van De Castele, T. (2000). Developmental instability and inbreeding in natural bird populations exposed to different levels of habitat disturbance. *J Evol Biol*, *13*, 889-896.
- Lerner, I. M. (1954). *Genetic homeostasis*. London: Oliver and Boyd, London.
- Leslie, J. F., & Vrijenhoek, R. C. (1980). Consideration of Muller's ratchet mechanism through studies of genetic linkage and genomic compatibilities in clonally reproducing *Poeciliopsis*. *Evolution*, *34*, 1105-1115.

- Lewontin, R. C. (1966). On the measurement of relative variability. *Systematic Zoology*, *15*, 141-142.
- Lewontin, R. C. (1983). Gene, organism and environment. In: *Evolution from molecules to men* (ed. Bendall, D. S.), pp. 273-285. Cambridge University Press, New York.
- Lexer, C., Welch, M. E., Durphy, J. L., & Rieseberg, L. H. (2003). Natural selection for salt tolerance quantitative trait loci (QTLs) in wild sunflower hybrids: implications for the origin of *Helianthus paradoxus*, a diploid hybrid species. *Mol Ecol*, *12*, 1225-1235.
- Liang, Z., Wei, J., Zhao, J., Liu, H., Li, B., Shen, J., & Zheng, C. (2008). The statistical meaning of kurtosis and its new application to identification of persons based on seismic signals. *Sensors (Basel, Switzerland)*, *8*, 5106-5119.
- Locke, A. E., Kahali, B., Berndt, S. I., Justice, A. E., Pers, T. H., Day, F. R., et al. (2015). Genetic studies of body mass index yield new insights for obesity biology. *Nature*, *518*, 197-206.
- Ludwig, W. (1932). Das Rechts-Links Problem im Teirreich und beim Menschen. Springer, Berlin. 496 pp.
- Lynch, C. J. (1969). The so-called Swiss mouse. *Lab Anim Care*, *19*, 214-220.
- Macholan, M., Munclinger, P., Sugerkova, M., Dufkova, P., Bimova, B., Bozikova, E., et al. (2007). Genetic analysis of autosomal and X-linked markers across a mouse hybrid zone. *Evolution*, *61*, 746-771.
- Magwene, P. M. (2001). New tools for studying integration and modularity. *Evolution*, *55*, 1734-1745.
- Manley, S. A. M., & Ledig, F. T. (1979). Photosynthesis in black and red spruce and their hybrid derivatives: ecological isolation and hybrid adaptive inferiority. *Can J Bot*, *57*, 305-314.
- Mariani, F. V., & Martin, G. R. (2003). Deciphering skeletal patterning: clues from the limb. *Nature*, *423*, 319-325.
- Markow, T. A. (1995). Evolutionary Ecology and Developmental Instability. *Annual Review of Entomology*, *40*, 105-120.
- Marroig, G., & Cheverud, J. M. (2001). A comparison of phenotypic variation and covariation patterns and the role of phylogeny, ecology, and ontogeny during cranial evolution of new world monkeys. *Evolution*, *55*, 2576-2600.
- Marroig, G., De Vivo, M., & Cheverud, J. M. (2004). Cranial evolution in sakis (Pithecia, Platyrrhini). II: Evolutionary processes and morphological integration. *J Evol Biol*, *17*, 144-155.
- Marroig, G., Shirai, L. T., Porto, A., de Oliveira, F. B., & De Conto, V. (2009). The evolution of modularity in the mammalian skull II: Evolutionary consequences. *Evolutionary Biology*, *36*, 136-148.
- Martin-Serra, A., Figueirido, B., Perez-Claros, J. A., & Palmqvist, P. (2015). Patterns of morphological integration in the appendicular skeleton of mammalian carnivores. *Evolution*, *69*, 321-340.
- Mather, K. (1953). Genetical control of stability in development. *Heredity*, *7*, 297-336.
- Mayr, E. (1970). *Populations, Species, and Evolution: An Abridgment of Animal species and Evolution*. Harvard University Press, Cambridge.
- Melo, D., Garcia, G., Hubbe, A., Assis, A. P., & Marroig, G. (2015). EvolQG - An R package for evolutionary quantitative genetics. *F1000Res*, *4*, 925.
- Mikula, O., & Macholán, M. (2008). There is no heterotic effect upon developmental stability in the ventral side of the skull within the house mouse hybrid zone. *J Evol Biol*, *21*, 1055-106710.

- Mitton, J. B., & Grant, M. C. (1984). Associations among protein heterozygosity, growth-rate, and developmental homeostasis. *Annual Review of Ecology and Systematics*, *15*, 479-499.
- Moller, A. P., & Thornhill, R. (1997). A meta-analysis of the heritability of developmental stability. *J evol biol*, *10*, 1-16.
- Moore, T. Y., Organ, C. L., Edwards, S. V., Biewener, A. A., Tabin, C. J., Jenkins, F. A., Jr., & Cooper, K. L. (2015). Multiple phylogenetically distinct events shaped the evolution of limb skeletal morphologies associated with bipedalism in the jerboas. *Curr Biol*, *25*, 2785-2794.
- Morse, H.C. (1978). *Origins of Inbred Mice*. Academic, New York.
- Mott, R., Talbot, C. J., Turri, M. G., Collins, A. C., & Flint, J. (2000). A method for fine mapping quantitative trait loci in outbred animal stocks. *Proc Natl Acad Sci U S A*, *97*, 12649-12654.
- Mukai, T., & Nagano, S. (1983). The genetic structure of natural populations of *Drosophila melanogaster*. XVI. Excess of additive genetic variance of viability. *Genetics*, *105*, 115-134.
- Nolte, A. W., & Sheets, H. D. (2005). Shape based assignment tests suggest transgressive phenotypes in natural sculpin hybrids (Teleostei, Scorpaeniformes, Cottidae). *Front Zool*, *2*, 11.
- Norgard, E. A., Jarvis, J. P., Roseman, C. C., Maxwell, T. J., Kenney-Hunt, J. P., Samocha, K. E., et al. (2009). Replication of long-bone length QTL in the F9-F10 LG,SM advanced intercross. *Mamm Genome*, *20*, 224-235.
- Norgard, E. A., Roseman, C. C., Fawcett, G. L., Pavlicev, M., Morgan, C. D., Pletscher, L. S., 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. *J Bone Miner Res*, *23*, 887-895.
- Ng, L. J., Wheatley, S., Muscat, G. E., et al. (1997). SOX9 binds DNA, activates transcription, and coexpresses with type II collagen during chondrogenesis in the mouse. *Developmental Biology*, *183*, 108-121.
- Okbay, A., Beauchamp, J. P., Fontana, M. A., Lee, J. J., Pers, T. H., Rietveld, C. A., et al. (2016). Genome-wide association study identifies 74 loci associated with educational attainment. *Nature*, *533*, 539-542.
- Olson, E. C., & Miller, R. L. (1958). *Morphological integration*. University of Chicago Press, Chicago.
- Otowa, T., Hek, K., Lee, M., Byrne, E. M., Mirza, S. S., Nivard, M. G., et al. (2016). Meta-analysis of genome-wide association studies of anxiety disorders. *Mol Psychiatry*, *21*, 1485.
- Pallares, L. F., Carbonetto, P., Gopalakrishnan, S., Parker, C. C., Ackert-Bicknell, C. L., Palmer, A. A., & Tautz, D. (2015). Mapping of craniofacial traits in outbred mice identifies major developmental genes involved in shape determination. *Plos Genetics*, *11*, e1005607.
- Pallares, L. F., Harr, B., Turner, L. M., & Tautz, D. (2014). Use of a natural hybrid zone for genomewide association mapping of craniofacial traits in the house mouse. *Mol Ecol*, *23*, 5756-5770.
- Pallares, L. F., Turner, L. M., & Tautz, D. (2016). Craniofacial shape transition across the house mouse hybrid zone: implications for the genetic architecture and evolution of between-species differences. *Dev Genes Evol*, *226*, 173-186.
- Palmer, A. R. (1994). Fluctuating asymmetry analyses: A primer. In: *Developmental instability: Its origins and evolutionary implications* (ed. Markow, T. A.), pp. 335-364. Kluwer, Dordrecht, Netherlands.
- Palmer, A. R. (1996). Waltzing with asymmetry. *BioScience*, *46*, 518-532.

- Palmer, A. R., & Strobeck, C. (1986). Fluctuating asymmetry: measurement, analysis, patterns. *Annual Review of Ecology and Systematics*, 17, 391-421.
- Palmer, A. R., & Strobeck, C. (1992). Fluctuating asymmetry as a measure of developmental stability: Implications of non-normal distributions and power of statistical tests. *Acta Zoologica Fennica*, 191, 57-72.
- Palmer, R., & Strobeck, C. (2003). Fluctuating asymmetry analysis unplugged. In: *Developmental instability: Causes and consequences* (ed. Polak, M.), pp. 279-319. Oxford University Press, Oxford.
- Parker, C. C., Gopalakrishnan, S., Carbonetto, P., Gonzales, N. M., Leung, E., Park, Y. J., et al. (2016). Genome-wide association study of behavioral, physiological and gene expression traits in outbred CFW mice. *Nat Genet*, 48, 919-926.
- Parker, C. C., & Palmer, A. A. (2011). Dark matter: Are mice the solution to missing heritability? *Front. Genet*, 2, 32.
- Pavlicev, M., Cheverud, J. M., & Wagner, G. P. (2009). Measuring morphological integration using eigenvalue variance. *Evolutionary Biology*, 36, 157-170.
- Pavlicev, M., & Wagner, G. P. (2012). A model of developmental evolution: selection, pleiotropy and compensation. *Trends Ecol Evol*, 27, 316-322.
- Pavlicev, M., Wagner, G. P., Noonan, J. P., Hallgrímsson, B., & Cheverud, J. M. (2013). Genomic correlates of relationship QTL involved in fore- versus hind limb divergence in mice. *Genome Biol Evol*, 5, 1926-1936.
- Patterson, B. D., & Patton, J. L. (1990). Fluctuating asymmetry and allozymic heterozygosity among natural populations of pocket gophers (*Thomomys bottae*). *Biol J Linn Soc*, 40, 21-36.
- Patton, J. L., & Brylski, P. V. (1987). Pocket gophers in alfalfa fields: causes and consequences of habitat-related body size variation. *Am Nat*, 130, 493-506.
- Payseur, B. A., & Hoekstra, H. E. (2005). Signatures of reproductive isolation in patterns of single nucleotide diversity across inbred strains of mice. *Genetics*, 171, 1905-1916.
- Payseur, B. A., Krenz, J. G., & Nachman, M. W. (2004). Differential patterns of introgression across the X chromosome in a hybrid zone between two species of house mice. *Evolution*, 58, 2064-2078.
- Pearson, E. S., & Hartley, H. O. eds. (1966). *Biometrika Tables for Statisticians*. Cambridge University Press, Cambridge.
- Pooley, S. M. (1972). Growth tables for 66 strains and stocks of laboratory animals. *Lab. Anim. Sci.* 22, 758-779.
- Polak, M., Møller, A. P., Gangestad, S. W., Kroeger D. E., Manning, J. T., Thornhill, R. (2003). Does an individual asymmetry parameter exist? A meta-analysis. In: *Developmental instability: causes and consequences* (ed. Polak, M.), pp. 81-96. Oxford University Press, Oxford.
- Porter, H. F., & O'Reilly, P. F. (2017). Multivariate simulation framework reveals performance of multi-trait GWAS methods. *Sci Rep*, 7, 38837.
- Porto, A., Shirai, L. T., de Oliveira, F. B., & Marroig, G. (2013). Size variation, growth strategies, and the evolution of modularity in the mammalian skull. *Evolution*, 67, 3305-3322.
- R Core Team (2016). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

- Ramos, P. S., Shedlock, A. M., & Langefeld, C. D. (2015). Genetics of autoimmune diseases: insights from population genetics. *J Hum Genet*, *60*, 657-664.
- Ranganath, H. A., & Aruna, S. (2003). Hybridization, transgressive segregation and evolution of new genetic systems in *Drosophila*. *J Genet*, *82*, 163-177.
- Raufaste, N., Orth, A., Belkhir, K., Senet, D., Smadja, C., Baird, S. J. E., et al. (2005). Inferences of selection and migration in the Danish house mouse hybrid zone. *Biological Journal of the Linnean Society*, *84*, 593-616.
- Reeves, N. M., Auerbach, B. M., Sylvester, A. D. (2016). Fluctuating and directional asymmetry in the long bones of captive cotton-top Tamarins (*Saguinus oedipus*). *Am J Phys Anthropol*, *160*, 41-51.
- Reno, P. L., McCollum, M. A., Cohn, M. J., Meindl, R. S., Hamrick, M., & Lovejoy, C. O. (2008). Patterns of correlation and covariation of anthropoid distal forelimb segments correspond to Hoxd expression territories. *J Exp Zool B Mol Dev Evol*, *310*, 240-258.
- Rice, M. C., & O'Brien, S. J. (1980). Genetic variance of laboratory outbred Swiss mice. *Nature*, *283*, 157-161.
- Rieseberg, L. H., Archer, M. A., & Wayne, R. K. (1999). Transgressive segregation, adaptation and speciation. *Heredity (Edinb)*, *83*, 363-372.
- Rieseberg, L. H., & Buerkle, C. A. (2002). Genetic mapping in hybrid zones. *Am Nat*, *159 Suppl 3*, S36-50.
- Riska, B. (1986). Some models for development, growth, and morphometric correlation. *Evolution*, *40*, 1303-1311.
- Robertson, F. W., & Reeve, E. C. (1952). Heterozygosity, environmental variation and heterosis. *Nature*, *170*, 286.
- Rockman, M. V. (2012). The QTN program and the alleles that matter for evolution: all that's gold does not glitter. *Evolution*, *66*, 1-17.
- Ross, K. G., & Robertson, J. L. (1990). Developmental stability, heterozygosity, and fitness in two introduced fire ants (*Solenopsis invicta* and *S. richteri*) and their hybrid. *Heredity*, *64*, 93-103.
- Sage, R. D. (1981). Wild mice. In: *The mouse in biochemical research* (eds. Foster, H. L., Small, J. D., Fox, J. G. (1982)), pp. 39-90. Academic Press, New York.
- Sanchez-Roige, S., Gray, J. C., MacKillop, J., Chen, C. H., & Palmer, A. A. (2018). The genetics of human personality. *Genes Brain and Behavior*, *17*, e12439.
- Sanger, T. J., Norgard, E. A., Pletscher, L. S., Bevilacqua, M., Brooks, V. R., Sandell, L. J., & Cheverud, J. M. (2011). Developmental and genetic origins of murine long bone length variation. *J Exp Zool B Mol Dev Evol*, *316B*, 146-161.
- Schaefer, K., Lauc, T., Mitterocher, P., Gunz, P., Bookstein, F. (2006). Dental arch asymmetry in an isolated Adriatic community. *Am J Phys Anthropol*, *129*, 132-142.
- Schluter, D. (1996). Adaptive radiation along genetic lines of least resistance. *Evolution*, *50*, 1766-1774.
- Schmidt, M., & Fischer, M. S. (2009). Morphological integration in mammalian limb proportions: dissociation between function and development. *Evolution*, *63*, 749-766.
- Schunke, A. C., Bromiley, P. A., Tautz, D., & Thacker, N. A. (2012). TINA manual landmarking tool: software for the precise digitization of 3D landmarks. *Front Zool*, *9*, 6.

- Searle, J. B., Jamieson, P. M., Gunduz, I., Stevens, M. I., Jones, E. P., Gemmill, C. E., & King, C. M. (2009). The diverse origins of New Zealand house mice. *Proc Biol Sci*, 276, 209-217.
- Shirai, L. T., & Marroig, G. (2010). Skull modularity in Neotropical marsupials and monkeys: size variation and evolutionary constraint and flexibility. *Journal of Experimental Zoology Part B-Molecular and Developmental Evolution*, 314b, 663-683.
- Small, K. M., & Potter, S. S. (1993). Homeotic transformations and limb defects in Hox A11 mutant mice. *Genes Dev*, 7, 2318-2328.
- Soule, M., & J. Cuzin-Roudy J. (1982). Allomeric variation. 2. Developmental instability of extreme phenotypes. *Amer Natur*, 129, 765-786.
- Stelkens, R., & Seehausen, O. (2009). Genetic distance between species predicts novel trait expression in their hybrids. *Evolution*, 63, 884-897.
- Steppan, S. J. (1997). Phylogenetic analysis of phenotypic covariance structure. I. Contrasting results from matrix correlation and common principal component analyses. *Evolution*, 51, 571-586.
- Svenson, K. L., Gatti, D. M., Valdar, W., Welsh, C. E., Cheng, R., Chesler, E. J., et al. (2012). High-resolution genetic mapping using the Mouse Diversity outbred population. *Genetics*, 190, 437-447.
- Swaddle, J.P. (1997). Within-individual changes in developmental stability affect flight performance. *Behavioral Ecology*, 8, 601-604.
- Szymura, J. M., & Barton, N. H. (1986). Genetic analysis of a hybrid zone between the fire-bellied toads, *Bombina bombina* and *Bombina variegata*, near Cracow in Southern Poland. *Evolution*, 40, 1141-1159.
- Tamura, K., Yonei-Tamura, S., & Izpisua Belmonte, J. C. (1999). Differential expression of Tbx4 and Tbx5 in Zebrafish fin buds. *Mech Dev*, 87, 181-184.
- Thoday, J. M. (1958). Homeostasis in a selection experiment. *Heredity* 12, 401-415.
- Thorpe, R. S. (1981). The morphometrics of the mouse: A review. In *Biology of the house mouse*, ed. R. J. Berry, 85-125. *Symposium of the Zoological Society of London* 47. London: Zoological Society of London.
- Tomkins, J. L., & Kotiaho, J. S. (2002). Fluctuating Asymmetry. In: *Encyclopedia of Life Sciences* (MacMillan Publishers Ltd.), pp. 1-5. Nature Publishing Group, London.
- Tucker, P. K., Sage, R. D., Warner, J., Wilson, A. C., & Eicher, E. M. (1992). Abrupt cline for sex-chromosomes in a hybrid zone between two species of mice. *Evolution*, 46, 1146-1163.
- Turner, L. M., & Harr, B. (2014). Genome-wide mapping in a house mouse hybrid zone reveals hybrid sterility loci and Dobzhansky-Muller interactions. *Elife*, 3, e02504.
- Turner, L. M., Schwahn, D. J., & 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.
- Turner, T. L., Stewart, A. D., Fields, A. T., Rice, W. R., & Tarone, A. M. (2011). Population-based resequencing of experimentally evolved populations reveals the genetic basis of body size variation in *Drosophila melanogaster*. *PLoS Genet*, 7, e1001336.
- Valdar, W., Solberg, L. C., Gauguier, D., Burnett, S., Klenerman, P., Cookson, W. O., et al. (2006). Genome-wide genetic association of complex traits in heterogeneous stock mice. *Nat Genet*, 38, 879-887.

- Vanlerberghe, F., Boursot, P., Catalan, J., Gerasimov, S., Bonhomme, F., Botev, B. A., & Thaler, L. (1988). Genetic analysis of the hybrid zone between two murine subspecies *Mus musculus domesticus* and *Mus musculus musculus* in Bulgaria. *Genome*, *30*, 427-437.
- Vanlerberghe, F., Dod, B., Boursot, P., Bellis, M., Bonhomme, F. (1986). Absence of Y-chromosome introgression across the hybrid zone between *Mus musculus domesticus* and *Mus musculus musculus*. *Genet Res*, *48*, 191-197.
- Van Valen, L. (1962). A study of fluctuating asymmetry. *Evolution*, *16*, 125-142.
- Van Valen, L. (1965). The study of morphological integration. *Evolution*, *19*, 347-349.
- Villmoare, B., Fish, J., & Jungers, W. (2011). Selection, morphological integration, and strepsirrhine locomotor adaptations. *Evolutionary Biology*, *38*, 88-99.
- Visscher, P. M., Brown, M. A., McCarthy, M. I., & Yang, J. (2012). Five years of GWAS discovery. *Am J Hum Genet*, *90*, 7-24.
- Waddington, C. H. (1957). *The strategy of the genes*. Allen & Unwin, London
- Wagner, G. P. (1984). On the eigenvalue distribution of genetic and phenotypic dispersion matrices: evidence for a nonrandom organization of quantitative character variation. *Journal of Mathematical Biology*, *21*, 77-95.
- Wagner, G. P., & Altenberg, L. (1996). Perspective: Complex adaptations and the evolution of evolvability. *Evolution*, *50*, 967-976.
- Wagner, G. P., Booth, G., & Bagheri-Chaichian, H. (1997). A population genetic theory of canalization. *Evolution*, *51*, 329-347.
- Wagner, G. P., & Chiu, C. H. (2001). The tetrapod limb: a hypothesis on its origin. *J Exp Zool*, *291*, 226-240.
- Waldmann, P. (1999). The effect of inbreeding and population hybridization of developmental instability in petals and leaves of the rare plant *Silene diclinis* (Caryophyllaceae). *Heredity*, *83*, 138-144.
- Weisbecker, V. (2011). Monotreme ossification sequences and the riddle of mammalian skeletal development. *Evolution*, *65*, 1323-1335.
- Wellik, D. M., & Capecchi, M. R. (2003). Hox10 and Hox11 genes are required to globally pattern the mammalian skeleton. *Science*, *301*, 363-367.
- Whitlock, M. C., & Fowler, K. (1997). The instability of studies of instability. *Journal of Evolutionary Biology*, *10*, 63-67.
- Willmore, K. E., Young, N. M., & Richtsmeier, J. T. (2007). Phenotypic variability: Its components, measurement and underlying developmental processes. *Evolutionary Biology*, *34*, 99-120.
- Wood, A. R., Esko, T., Yang, J., Vedantam, S., Pers, T. H., Gustafsson, S., et al. (2014). Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat Genet*, *46*, 1173-1186.
- Woods, L. C. S. (2014). QTL mapping in outbred populations: successes and challenges. *Physiological Genomics*, *46*, 81-90.
- Wooten, M. C., & Smith, M. H. (1986). Fluctuating asymmetry and genetic variability in a natural population of *Mus musculus*. *Journal of Mammalogy*, *67*, 725-732.



- Wray, N. R., Yang, J., Hayes, B. J., Price, A. L., Goddard, M. E., & Visscher, P. M. (2013). Pitfalls of predicting complex traits from SNPs. *Nat Rev Genet*, *14*, 507-515.
- Wright, E., Hargrave, M. R., Christiansen, J., Cooper, L., Kun, J., Evans, T., et al. (1995). The Sry-related gene Sox9 is expressed during chondrogenesis in mouse embryos. *Nature Genetics*, *9*, 15-20.
- Wright, S. (1968). *Evolution and genetics of populations*. Chicago University Press, Chicago.
- Yalcin, B., Nicod, J., Bhomra, A., Davidson, S., Cleak, J., Farinelli, L., et al. (2010). Commercially available outbred mice for genome-wide association studies. *PLoS Genet*, *6*, e1001085.
- Yampolsky, L. Y., & Scheiner, S. M. (1994). Developmental noise, phenotypic plasticity, and allozyme heterozygosity in *Daphnia*. *Evolution*, *48*, 1715-1722.
- Yang, J., Benyamin, B., McEvoy, B. P., Gordon, S., Henders, A. K., Nyholt, et al. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nat Genet*, *42*, 565-569.
- Yezerinac, S. M., Loughheed, S. C., & Handford, P. (1992). Measurement error and morphometric studies: statistical power and observer experience. *Systematic Biology*, *41*, 471-482.
- Young, N. M., & Hallgrímsson, B. (2005). Serial homology and the evolution of mammalian limb covariation structure. *Evolution*, *59*, 2691-2704.
- Young, N. M., Hallgrímsson, B., & Garland, T. (2009). Epigenetic effects on integration of limb lengths in a mouse model: Selective breeding for high voluntary locomotor activity. *Evolutionary Biology*, *36*, 88-99.
- Zelditch, M. L. (1988). Ontogenetic variation in patterns of phenotypic integration in the laboratory rat. *Evolution*, *42*, 28-41.
- Zelditch, M. L., & Carmichael, C. (1989). Ontogenetic variation in patterns of developmental and functional integration in skulls of *Sigmodon Fulviventer*. *Evolution*, *43*, 814-824.
- Zhao, Q., Eberspaecher, H., Lefebvre V., De Crombrughe B. (1997). Parallel expression of Sox9 and Col2a1 in cells undergoing chondrogenesis. *Development Dynamics*, *209*, 377-386.

# Digital Supplement

Supporting information with supplementary tables and figures of this dissertation will be available in a digital archive.

Description of tables and figures in a supplementary material:

## List of Tables

### CHAPTER I

Table <b>S1.1.</b> List of samples for the mice from the hybrid zone.	1
Table <b>S1.2.</b> Two-way, mixed model ANOVA.	5
Table <b>S1.3.</b> Tests for departure from normality, skew, kurtosis and directional asymmetry.	9
Table <b>S1.4.</b> Tests for directional asymmetry per population / strain and hybrid group.	14
Table <b>S1.5.</b> Tests for departure from normality, skew and kurtosis per population / strain and hybrid group.	17
Table <b>S1.6.</b> Tests of associations between trait asymmetry and a character size.	22
Table <b>S1.7.</b> Two-way ANOVA test for differences in FA between sexes and among traits.	26
Table <b>S1.8.</b> Differences in FA between sexes per bone in each population / strain and group.	27
Table <b>S1.9.</b> Differences in FA due to age in hybrid group.	28
Table <b>S1.10.</b> Descriptive statistics of each bone per population / strain and group.	33
Table <b>S1.11.</b> Two-way ANOVA test for differences in FA between groups and among traits.	37
Table <b>S1.12.</b> Comparison of FA levels between bones.	39
Table <b>S1.13.</b> Comparison of FA levels between populations / strains and groups for the same bone.	40
Table <b>S1.14.</b> Comparison of scaled FA differences between bones with all bones included.	41
Table <b>S1.15.</b> Comparison of scaled FA differences between bones in non-filtered data.	43
Table <b>S1.16.</b> FA1 and FA8 indices for each bone and composite measure of FA (FA17) per population / strain and hybrid group.	45
List <b>S1.1.</b> Outlier list.	11

### CHAPTER II

Table <b>S2.1.</b> Linear regression for each bone length.	48
Table <b>S2.2.</b> Proportion of variance explained by PC1 from correlation and covariance matrix.	48
Table <b>S2.3.</b> Descriptive statistics of each bone per population / strain, per group, and per hybrid subgroup.	52
Table <b>S2.4.</b> Matrix repeatability for correlation and covariance matrices per sex in each population / strain.	53
Table <b>S2.5.</b> Correlation and covariance matrices in raw data for each sex per population / strain.	54
Table <b>S2.6.</b> Similarities between correlation and covariance matrices per population and hybrid subgroups.	55
Table <b>S2.7.</b> Partial correlations per sex per population / strain.	56
Table <b>S2.8.</b> Similarities between matrices from partial correlations between populations / strains and hybrid subgroups.	57

Table <b>S2.9</b> . Integration indices based on correlation ( $r^2$ ) and variance / covariance matrices (ICV) per sex in each population / strain.	58
Table <b>S2.10</b> . Results for the indices from correlation ( $r^2$ , EV) and covariance (ICV) matrices.	58

### CHAPTER III

**Note: Table identifiers are re-numbered for consistency (Table S1 in Skrabar et al. (2018) is listed as Table S3.1, etc.)**

Table <b>S3.1</b> . List of mouse samples used for mapping. The table is shown in an extended version of the Chapter I, Table S1.1.	
Table <b>S3.2</b> . Results of two-sample $t$ -test with equal variances between the right and the left side of the limb bones in raw measurement data.	66
Table <b>S3.3</b> . Normality test of limb bone length in raw data of 172 individuals.	67
Table <b>S3.4</b> . Normality test of limb bone length in residual values for age in 172 individuals.	68
Table <b>S3.5</b> . Associated SNPs from the univariate mapping of individual limb bones.	70
Table <b>S3.6</b> . PVE versus PGE comparisons.	73
Table <b>S3.7</b> . Posterior mean estimates for the effect size parameters from BSLMM for the best SNP from every selected region in the univariate mapping of individual limb bones.	73
Table <b>S3.8</b> . Power of GWAS in the house mouse hybrid zone under an additive model and various genetic architectures.	75
Table <b>S3.9</b> . Associated SNPs from the univariate mapping with including another bone as a covariate.	77

## List of Figures

### CHAPTER I

Figure <b>S1.1</b> . Histograms for possible outliers based on larger differences between the right and the left side.	12
Figure <b>S1.2</b> . Histograms for possible outliers based on larger differences between the right and the left side.	13
Figure <b>S1.3</b> . Frequency distributions for each bone per population and sex in the outbred group.	19
Figure <b>S1.4</b> . Frequency distributions for each bone per population and sex in the outbred group.	20
Figure <b>S1.5</b> . Frequency distributions for each bone per strain and sex in the inbred group.	21
Figure <b>S1.6</b> . Frequency distributions for each bone per population in the outbred group.	29
Figure <b>S1.7</b> . Frequency distributions for each bone per strain in the inbred group.	30
Figure <b>S1.8</b> . Frequency distributions for each bone in six hybrid subgroups.	31
Figure <b>S1.9</b> . Frequency distributions for each bone per group.	32
Figure <b>S1.10</b> . Frequency distributions per bone with pooled data for each bone.	32
Figure <b>S1.11</b> . Bone length in six hybrid subgroups.	36
Figure <b>S1.12</b> . Difference in FA levels between four bones in populations from the outbred group.	37
Figure <b>S1.13</b> . Difference in FA levels between four bones in strains from the inbred group.	38
Figure <b>S1.14</b> . Difference in FA levels between four bones in subgroups of the hybrid group.	38
Figure <b>S1.15</b> . Difference in FA levels between four bones with pooled data per bone.	39

### CHAPTER II

Figure <b>S2.1</b> . Quantile-quantile plots for residual limb length measures per population / strain.	49
Figure <b>S2.2</b> . Quantile-quantile plots for residual limb length measures in hybrid subgroups.	50
Figure <b>S2.3</b> . Quantile-quantile plots for residual limb length measures per group.	51

Figure <b>S2.4.</b> Distribution of resampled eigenvalue variance from correlation matrices in data corrected for size in in the hybrid, outbred and inbred group.	62
Figure <b>S2.5.</b> Distribution of resampled eigenvalue variance from correlation matrices in data corrected for size in populations / strains and in hybrid subgroups.	62
Figure <b>S2.6.</b> Distribution of resampled eigenvalues from variance-covariance matrices in data corrected for size in populations / strains and in hybrid subgroups.	63
Figure <b>S2.7.</b> Distribution of resampled eigenvalue variance from variance-covariance matrices in data corrected for size in in the hybrid, outbred and inbred group.	63
Figure <b>S2.8.</b> Distribution of resampled eigenvalue variance from correlation matrices in non-filtered data in populations / strains and in hybrid, outbred and inbred groups.	64
Figure <b>S2.9.</b> Distribution of resampled eigenvalue variance from correlation matrices in non-filtered data, corrected for size, in populations / strains and in hybrid, outbred and inbred groups.	64

### CHAPTER III

**Note: Figure identifiers are re-numbered for consistency (Figure S1 in Skrabar et al. (2018) is listed as Figure S3.1, etc.)**

Figure <b>S3.1.</b> Quantile-quantile plot of each limb bone length in raw data.	67
Figure <b>S3.2.</b> Quantile-quantile plot of each limb bone length in residual values for age.	68
Figure <b>S3.3.</b> Q-Q plot showing SNPs associations for each limb bone length based on univariate analysis adjusted for population stratification using a relatedness matrix.	69
Figure <b>S3.4.</b> Relationship between phenotypic variance explained by each chromosome and chromosome length.	74
Figure <b>S3.5.</b> Distribution of the number of causal regions within 10Mb of the causal SNP in 1,000 simulated replicates.	76
Figure <b>S3.6.</b> Manhattan plot showing SNPs associations for each limb bone length based on univariate analysis with including another bone as a covariate.	83
Figure <b>S3.7.</b> Q-Q plot showing SNPs associations for each limb bone length based on univariate analysis with including another bone as a covariate, adjusted for population stratification using a relatedness matrix.	84
Text <b>S3.1.</b> Effect size for significant regions.	73
Text <b>S3.2.</b> Power simulations of GWAS in hybrid zone population.	75

## Curriculum Vitae

<b>Name</b>	Neva Škrabar
<b>Date and place of birth</b>	11.02.1986 in Belgrade, Serbia
<b>Nationality</b>	Serbian
<b>Place of residence</b>	Rodomstorstraße 37, 24306, Plön, Germany
<b>Education</b>	
1993 – 2001	Primary School, Belgrade, Serbia “Vojvoda Radomir Putnik”
2001 – 2005	4 <sup>th</sup> Belgrade Grammar School Department of Natural Sciences
2005 – 2012	Diploma / MSc in Biology Faculty of Biology, University of Belgrade Study Group: Biology, Department: Genetics
Since February 2014	Doctoral Studies Max Planck Institute for Evolutionary Biology Department of Evolutionary Genetics Supervisor: Diethard Tautz Plön, Germany Christian-Albrecht-University of Kiel

## Affidavit

Hiermit erkläre ich, dass die vorliegende Arbeit

- nach Inhalt und Form meine eigene ist, abgesehen von der Beratung durch meinen Betreuer Prof. Dr.

Diethard Tautz und die Beiträge der Ko-Autoren für die im Kapitel 3 wieder gegebene Publikation.

- an keiner anderen Stelle im Rahmen eines Prüfungsverfahrens vorgelegen hat, noch nicht veröffentlicht ist und auch nicht zur Veröffentlichung eingereicht wurde.

- unter Einhaltung der Regeln guter wissenschaftlicher Praxis der Deutschen Forschungsgemeinschaft entstanden ist.

Plön, den 14.06.2018

Neva Škrabar