QUANTITATIVE GENETICS OF SKULL SHAPE IN SOAY SHEEP (Ovis aries) USING GEOMETRIC MORPHOMETRICS

A thesis submitted to the University of Manchester for the degree of DOCTOR OF PHILOSOPHY in the Faculty of Life Sciences

2015

ELIS MARINA DAMASCENO SILVA
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. General Introduction</td>
<td>12</td>
</tr>
<tr>
<td>1.1 Geometric morphometrics</td>
<td>12</td>
</tr>
<tr>
<td>1.1.1 The Study of Shape</td>
<td>12</td>
</tr>
<tr>
<td>1.1.2 Integration and Modularity</td>
<td>14</td>
</tr>
<tr>
<td>1.1.3 Fluctuating asymmetry</td>
<td>17</td>
</tr>
<tr>
<td>1.2 Quantitative Genetics</td>
<td>19</td>
</tr>
<tr>
<td>1.2.1 Estimation of genetic parameters</td>
<td>19</td>
</tr>
<tr>
<td>1.2.2 Breeder’s Equation</td>
<td>21</td>
</tr>
<tr>
<td>1.2.3 Estimation of Natural Selection</td>
<td>23</td>
</tr>
<tr>
<td>1.2.4 Hypothetical Selection</td>
<td>24</td>
</tr>
<tr>
<td>1.3 Soay sheep</td>
<td>27</td>
</tr>
<tr>
<td>1.3.1 Historical background</td>
<td>27</td>
</tr>
<tr>
<td>1.3.2 Anatomical description</td>
<td>28</td>
</tr>
<tr>
<td>1.3.3 Population monitoring</td>
<td>29</td>
</tr>
<tr>
<td>1.4 Authors’ Contributions and Thesis Outline</td>
<td>31</td>
</tr>
<tr>
<td>1.5 Literature Cited</td>
<td>32</td>
</tr>
<tr>
<td>2. Chapter 2:</td>
<td>39</td>
</tr>
<tr>
<td>Multi-level integration and modularity in the skull of Soay sheep</td>
<td>39</td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>39</td>
</tr>
<tr>
<td>2.2 Methods</td>
<td>43</td>
</tr>
<tr>
<td>2.2.1 Shape data</td>
<td>43</td>
</tr>
<tr>
<td>2.2.2 Ontogenetic stages</td>
<td>47</td>
</tr>
<tr>
<td>2.2.3 G matrix</td>
<td>48</td>
</tr>
<tr>
<td>2.2.4 Matrix correlation</td>
<td>49</td>
</tr>
<tr>
<td>2.2.5 Overall integration</td>
<td>50</td>
</tr>
<tr>
<td>2.2.6 Modularity hypotheses</td>
<td>50</td>
</tr>
<tr>
<td>2.3 Results</td>
<td>52</td>
</tr>
<tr>
<td>2.3.1 Overall Integration</td>
<td>53</td>
</tr>
<tr>
<td>2.3.2 Matrix correlations</td>
<td>56</td>
</tr>
<tr>
<td>2.3.3 Modularity hypotheses</td>
<td>57</td>
</tr>
<tr>
<td>2.4 Discussion</td>
<td>61</td>
</tr>
<tr>
<td>2.4.1 Overall Integration</td>
<td>61</td>
</tr>
<tr>
<td>2.4.2 Matrix correlations</td>
<td>62</td>
</tr>
</tbody>
</table>
2.4.3 Modularity hypotheses .............................................................................63
2.4.4 Integration, modularity & evolvability .....................................................65
2.5 Literature cited .........................................................................................65

3. Chapter 3: .................................................................................................72

The role of genetic integration on the response to selection in the skull of Soay
sheep ...........................................................................................................72

3.1 Introduction ..............................................................................................72
3.2 Method .......................................................................................................74
  3.2.1 Data collection and morphometric analyses ......................................75
  3.2.2 Quantitative genetic analyses .............................................................76
  3.2.3 Hypothetical selection .........................................................................77
3.3 Results .......................................................................................................81
3.4 Discussion ................................................................................................86
3.5 Literature Cited .........................................................................................89

4. Chapter 4: .................................................................................................94

Estimating selection of skull shape using geometric morphometrics ..........94

4.1 Introduction ..............................................................................................94
4.2 Methods .....................................................................................................97
  4.2.1 Study population ................................................................................97
  4.2.2 Fitness measure .................................................................................98
  4.2.3 Shape data ........................................................................................99
  4.2.4 Estimation of Selection Gradients ....................................................100
  4.2.5 Linear Selection ...............................................................................102
  4.2.6 Nonlinear Selection .........................................................................102
4.3 Results .......................................................................................................103
4.4 Discussion ................................................................................................105

5. Chapter 5: .................................................................................................113

Evidence of climate-change-induced heterochrony in the skull of Soay sheep 113

5.1 Introduction ..............................................................................................113
5.2 Methods .....................................................................................................116
5.3 Results .......................................................................................................120
5.4 Discussion ................................................................................................125
5.5 Literature cited ........................................................................................128

6. Chapter 6: .................................................................................................133
Testing the influence of environment, life-history, and genetics on fluctuating asymmetry levels................................................................. 133

6.1 Introduction........................................................................................................... 133

6.2 Methods .................................................................................................................. 136
   6.2.1 Fluctuating asymmetry scores ................................................................. 136
   6.2.2 Environmental & Morphological Variables.............................................. 138
   6.2.3 Statistical analyses......................................................................................... 140
   6.2.4 Quantitative genetics...................................................................................... 141

6.3 Results.................................................................................................................... 142

6.4 Discussion............................................................................................................... 145
   6.4.1 Fluctuating asymmetry and fitness........................................................... 145
   6.4.2 Fluctuating asymmetry and environment.................................................. 146
   6.4.3 Fluctuating asymmetry and size.................................................................. 147
   6.4.4 Inheritance of fluctuating asymmetry......................................................... 147

6.5 Literature Cited..................................................................................................... 149

7. General Conclusions............................................................................................... 155

7.1 Literature Cited..................................................................................................... 159

Final word count: 42,985
LIST OF FIGURES

Chapter 1- Introduction
Figure 1. Representation of genetic integration in the shape space ................... 26
Figure 2. Scheme photographs of Soay sheep individuals, each displaying one type of horn morph ................................................................. 29

Chapter 2
Figure 1. Scheme showing location of landmarks in the skull ......................... 44
Figure 2. PCA scatterplot of individual variation (P matrix) corrected for sex, showing the different ontogenetic stages ........................................... 48
Figure 3. Adjacency graphs and set of landmarks involved in each modularity hypothesis. ................................................................................... 51
Figure 4. Bar charts showing the percentage of variance explain by principal components .................................................................................. 54
Figure 5. Shape changes associated with PC1 - individual variation ............... 55
Figure 6. Shape changes associated with PC1 - G matrix ................................ 55
Figure 7. Shape changes associated with PC1 - FA matrix ............................... 56
Figure 8. Bar charts showing RV coefficient frequencies for FN and OC hypothesis, using individual variation, fluctuating asymmetry, and genetic matrix. ............................................................... 58
Figure 9. Bar charts showing RV coefficient frequencies for FN hypothesis, using individual variation and fluctuating asymmetry, in different ontogenetic stages: yearlings, young adults, adults and old animals. ........................................ 60

Chapter 3
Figure 1. Scheme showing location of landmarks in the skull ....................... 76
Figure 2. Decomposition of selection response into direct, total and correlated selection - frontality of the orbits ......................................................... 81
Figure 3. Decomposition of selection response into direct, total and correlated selection - retraction of the nose ......................................................... 83
Figure 4. Decomposition of selection response into direct, total and correlated selection - reduced molars ................................................................. 84
Figure 5. Decomposition of selection response into direct, total and correlated selection - reduced braincase ............................................................. 85
Chapter 4

Figure 1. Scheme showing location of landmarks in the skull .......................... 100

Figure 2. Scatterplots from regressions between relative fitness and skull shape for males and females. ................................................................. 104

Figure 3. Diagram showing the skull shape selected in males ....................... 104

Figure 4. Diagram showing the skull shape selected in females ...................... 105

Chapter 5

Figure 1. Scheme showing location of landmarks in the skull ......................... 118

Figure 2. Graph from ANOVA results, showing relationship between centroid size and year of birth ................................................................. 121

Figure 3. Scatterplot from regression between shape and year of birth, and diagram with the associated shape changes. ................................. 122

Figure 4. Scatterplot from regression between shape and log centroid size, and diagram with the associated shape changes.................................. 124

Chapter 6

Figure 1. Scheme showing location of landmarks in the skull ....................... 137

Figure 2. ANOVA results showing the respective mean (black circles) and standard deviations (vertical bars) of FA scores (Procrustes and Mahalanobis) for animals born in pre-crash, between-crashes, and post-crash years. ............................. 144
LIST OF TABLES:

Chapter 2
Table 1. List of landmark numbers and their respective anatomical description, and which subset it belongs for each hypothesis tested. ................................................. 46
Table 2. Results from the Procrustes ANOVA on centroid size ..................... 53
Table 3. Results from the Procrustes ANOVA on shape ............................... 53

Chapter 4
Table 1. Average values of Life-time Breeding Success (LBS) for each sex and each age group ................................................................. 99

Chapter 6
Table 1. Results from the partial correlations between FA scores and the different variables ...................................................................................... 142
Table 2. Inheritance components for Procrustes and Mahalanobis scores .......... 145
ABSTRACT

Institution: University of Manchester
Name: Elis Marina Damasceno Silva
Degree Title: PhD Evolutionary Biology
Thesis Title: Quantitative genetics of skull shape in Soay sheep (Ovis aries) using geometric morphometrics
Date: 27/09/2015

Analysing variation of skull shape in wild populations can give us an insight into evolutionary processes. By looking into patterns of morphological variation within populations, we can extrapolate and make assumptions on the patterns of variation on higher taxonomic levels. In this thesis, I collected data on skull shape of a wild population of Soay sheep using geometric morphometrics methods. I applied a multitude of quantitative genetics and geometric morphometrics methods to explore the factors behind the evolution of skull shape. I analysed several aspects of skull shape, including integration and modularity, predicted response to hypothetical selection, estimates of natural selection, presence of heterochrony, and fluctuating asymmetry (FA) in an ecological and genetic point of view. I found low levels of integration in the Soay sheep skull, along with the presence of two modules: face and neurocranium. Both the levels of integration and the pattern of modularity were found across different origins of variation, individual, genetic and developmental (FA), indicating a strong correspondence between levels. Although genetic integration is not strong, I found that most of hypothetical selection regimes result in as much correlated selection as direct selection responses. But the nose region has shown to be quite independent from the rest of the skull, with a retraction of the nose causing almost no correlated changes. Selection on skull shape is comparatively strong in the Soay sheep. The skull shape selected for in males differs from the shape selected in females. Males are selected to have wider nasal bones along with longer, flatter braincase, whereas females are selected towards narrower nasal bones, along with smaller braincases. But selection is not the only factor influencing changes in skull shape. Changes in temperature along the past 30 years are causing a change in developmental timing in Soay sheep, causing a decrease in overall body size in the population. And I found that skull shape is also being affected, thus indicating that this population might be undergoing heterochronic processes. Finally, I found that factors such as vegetation quality, parasite load and breeding success are related to FA levels, whereas population density and climate are not. I also found no significant heritability nor dominance in FA of skull shape. Overall, this thesis certainly contributes to the understanding of evolution of skull shape in natural populations. By successfully combining quantitative genetics and geometric morphometrics methods, I was able to tackle questions about the evolution of complex shapes, and give valuable insights into this still underexplored field.
DECLARATION

I hereby declare that no portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

COPYRIGHT STATEMENT

The following four notes on copyright and the ownership of intellectual property rights must be included as written below:

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trademarks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://documents.manchester.ac.uk/Doculnfo.aspx?DocID=487), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see http://www.manchester.ac.uk/library/aboutus/regulations) and in The University’s policy on Presentation of Theses.
DEDICATION AND ACKNOWLEDGMENTS

I would like to dedicate this thesis to two of the most beloved women in my life. One of them, whom I have known all my life, is my mother, Olga. The other, whom I am yet to meet, is my two-weeks-old niece, Letícia.

Eu gostaria de dedicar esta tese a duas das mulheres que eu mais amo no mundo. Uma delas eu conheço a minha vida toda, minha mãe, Olga. A outra, que eu ainda não conheci, é minha sobrinha, Letícia.

First of all, I must thank my supervisor, Chris Klingenberg, for giving me this opportunity, and being extremely supportive during the whole duration of my PhD, and the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for funding my project. I thank my lab-mates Hugo Benítez and Eli Mora for all the discussions in the lab, and the moments of leisure. I am extremely grateful to Josephine Pemberton, Jill Pilkington and Charlotte Regan, for giving me the opportunity to see the Soay sheep in their habitat in St. Kilda. I am especially thankful to Mariana Régo and Thomas Püschel for all the support they have given me this year, Vera Stankovic for the movies, chats, pints, which I will miss dearly, and to Ceferino Varón González, for all of our discussions, conversations, banter, and for his invaluable friendship.

I am also extremely grateful to my family, who has supported me unconditionally throughout my whole academic career, and have been enduring, for the past 10 years, my talking about skulls, evolution and "that geometry thing I do". Last, but certainly not least, I would like to thank Edvar Vilar, who has been the
most supportive person throughout this whole experience, for coping with the
distance, and lately also my stress, and eventual tantrums.

The past three years, and especially this year, has been extremely emotionally
and mentally challenging, and I could not have finished my PhD without the support
of everyone I mentioned, plus many more people that have, one way or another,
given me the strength to continue.
1. General Introduction

The aim of this thesis is to explore the different factors behind morphological variation in skull shape of a population of Soay sheep. To do that, I combined two methods that have been extremely relevant in evolutionary biology studies. One method, geometric morphometrics, is used to quantify, analyse and visualize information on the shape of organisms or traits. The other method, quantitative genetics, uses quantitative data, such as measurements, and combines them with information on the kinship of individuals, to estimate the genetic basis behind the variation of the quantitative variable. In this thesis, I take advantage of the vast amount of ecological and environmental information available for this particular Soay sheep population, and use it to test the effects of different factors on the variation of skull morphology. The amount of data available for each individual is also useful when correcting for variables we already know to affect skull morphology, such as sex and age, thus giving more accuracy and credibility to my results. Also contributing to the accuracy and credibility of my results, is the large sample size (more than 1,000 specimens), which is almost unprecedented in morphological studies of large mammals.

1.1 Geometric morphometrics

1.1.1 The Study of Shape

The shape of organisms has always been of great interest in many kinds of biological studies (Darwin 1859; Thompson 1942; Zelditch et al. 2012). Although
shape is infinitely more informative than size, analyzing shape requires more effort, due to its inherent multivariate composition. Geometric morphometrics is a method that combines geometry with multivariate statistics to capture information on shape (Rohlf and Marcus 1993; Astúa 2003; Zelditch et al. 2012). By quantifying shape as a multivariate trait, we manage to analyze shape as a whole, instead of as a group of linear measurements (as is done in traditional morphometrics), which not only gives us more realistic results, but also allows us to visualize the results of analyses as actual changes in shape. In geometric morphometrics, shape is defined as all the geometric information that remains after location, scale and rotational effects are removed (Kendall 1977). The method is based on placing landmarks in homologous points of the measured structures (Rohlf and Bookstein 1990; Bookstein 1991; Rohlf and Marcus 1993; Bookstein 1996; Astúa 2003; Zelditch et al. 2012). Two or three coordinates are recorded for each landmark, depending whether the data is 2-dimensional or 3-dimensional. After recording the coordinates, the landmark configurations are rotated, scaled and aligned, in order to correct for size, position, and orientation of the structure under study, thus leaving only information on shape. The effect of size is removed by scaling the centroid sizes to one unit. Centroid size is a measure of the amount of dispersion of landmarks around the centroid (centre of gravity). This process of rotating, scaling and aligning the landmarks is known as Procrustes superimposition (Dryden and Mardia 1998), and it is one of the essential preliminary steps prior to any geometric morphometrics analysis.

In symmetrical structures (where parts of the organism are repeated as mirrored copies of each other), such as the skull, the overall shape can be decomposed into two different components: symmetrical and asymmetrical. The symmetrical component is
characterized by all the variation that occurs equally in both sides of the skull, whereas the asymmetrical component pertains the variation in the differences between sides (Klingenberg et al. 2002; Klingenberg 2015). Most studies are interested in the symmetrical component of shape, since most processes behind phenotypic variation affect both sides of the skull. But variation in asymmetry can also be useful in studies that make inferences on developmental processes, since asymmetries in symmetrical structures have long been used as a measure of stress during development (Klingenberg 2003a, 2015).

The aligned landmark coordinates can be used for several sorts of statistical tests, like discriminant functions, regressions, partial least squares analyses, among others. But some analysis, such as principal component analysis and modularity analysis require covariance matrices, instead of the raw coordinates data. From the configuration of landmarks, we can also determine the variability of the data, by analyzing how the configurations relate with each other in the phenotypic space (also known as morphospace and Kendall shape space). In the phenotypic space, each set of landmarks is represented by one point, and the distances between points represent the extent of morphological disparity, whereas the direction in which the points vary represents the nature of the morphological change (Kendall 1984; Rohlf and Bookstein 1990; Bookstein 1991; Rohlf and Marcus 1993; Bookstein 1996; Dryden and Mardia 1998; Astúa 2003; Zelditch et al. 2012).

1.1.2 Integration and Modularity

One of the aspects of shape variation that has been receiving some attention recently, is how multiple traits within the same organism covary with one another
(Klingenberg 2014). This covariation, or integration, among traits is essential to the coordination of the shape and size of organs and organisms (Olson and Miller 1958; Cheverud 1982; Wagner and Altenberg 1996). For example, if the integration was absent in the human body, some people could be born with completely disproportionate bodies, with one leg being twice as long as the other, or with arms longer than the length of their bodies. But, with few exceptions, that is not what happens. Even though there is a variety of sizes and shapes within human bodies, our anatomy retains certain proportions within the length of our limbs, and between the lengths of the limbs and the torso, which means that there is some integration between the length of our limbs and torso. So, it becomes clear that the integration between traits is essential in order to maintain the functionality of organisms, and crucial to the understanding of the evolution of the different morphologies we see in nature (Cheverud 1996; Wagner 1996; Wagner and Altenberg 1996; Arthur 2001; Klingenberg 2008; Futuyma 2010).

Studying how the different traits covary, or patterns of integration, is important because only then we can understand how variation is distributed within a structure, which also determines how variation is distributed within a population and even species or higher taxa (Olson and Miller 1958). For example, Drake and Klingenberg (2010) found patterns of integration within the skull of domestic dogs in which the traits covary more within the face and the neurocranium than between these parts. In this study, they found the same pattern of integration in the skull of wolves and among several genera of the Carnivora order (Drake and Klingenberg 2010). Analyzing the level and patterns of integration between traits are, therefore, an
important step towards understanding the evolutionary processes behind phenotypic change (Wagner and Altenberg 1996).

Covariation within traits can be studied in two aspects: how much the traits covary with each other; and what is the pattern of this covariation (Klingenberg 2008). The strength of covariation within traits is known as integration (Olson and Miller 1958; Wagner 1984). Structures or organisms are considered to be highly integrated when different traits are strongly correlated with each other, and weakly integrated when traits vary independently, without affecting other traits (Olson and Miller 1958; Klingenberg 2008, 2009). The concept of modularity is closely linked with integration, since both encompass the covariation of traits. Modules are units belonging to the same organism that are more tightly integrated within themselves than with other units (Klingenberg 2008, 2009). In other words, the integration within modules is higher than between modules (Cheverud 1982; Klingenberg 2008). But it should be made clear that integration and modularity are not mutually exclusive. A structure can be integrated overall, but still be divided into modules, as long as the traits within a module are more strongly integrated (Klingenberg 2008).

There are different ways of measuring integration in a dataset. One way is by calculating the degree of covariation between the traits within a structure. Another way is by assessing the variability of the data. Integration affects the variability of structures, because when there is strong association between traits, the overall variation will be restricted to a single direction (or a few), since changes in one trait will lead to the same amount of change in another trait. Therefore, phenotypic variation in integrated structures is usually concentrated in one dimension of the shape space, the more tightly grouped the points are, in the shape space, the more
integrated the structure is (Wagner 1984; Klingenberg 2008). The patterns of integration are the specific shape changes that occur together, which can be observed as the shape changes associated with that one dimension of the shape space in which the variation is concentrated (Klingenberg 2008).

In modularity analysis, one needs to assess whether covariation within a set of traits (modules) is stronger than the covariation between different sets. Unlike integration, estimating modularity in a structure requires previously defined hypothesis of modularity (Klingenberg 2008). Many different methods have been developed to analyze modularity hypothesis, including: cluster analyses (Goswami 2006); matrix correlations (Monteiro et al. 2005; Cardini and Elton 2008); partial least squares (Mitteroecker and Bookstein 2008); 'alternative partitions' (Klingenberg 2009); and 'modularity ratios' (Porto et al. 2009).

1.1.3 Fluctuating asymmetry

As mentioned previously, one of the aspects of shape variation describes the differences between the two sides (in bilateral symmetry) of a symmetrical structure. Degrees of dissimilarities between symmetric parts of the same structure, also known as fluctuating asymmetry (FA), are often related to stress suffered during pre and post-natal development (developmental instability) (Palmer and Strobeck 1986; Markow 1995). Developmental instability represents disturbances during the development of a phenotype under given environmental conditions (Møller and Swaddle 1997). These disturbances in development, and consequently, FA levels are usually caused by poor environmental conditions (Markow 1995; Badyaev et al. 2000; Marchand et al. 2003; Oleksyk et al. 2004; Sánchez-Chardi et al. 2013; Lazić
et al. 2015; Lezcano et al. 2015; Maestri et al. 2015), and low genetic quality (Debat et al. 2000; Rego et al. 2006; Carreira et al. 2008; Albarrán-Lara et al. 2010; Mikula et al. 2010; Nouvellet et al. 2011; Quinto-Sánchez et al. 2015). Fluctuating asymmetry data can also be used in modularity analysis, to understand the developmental origins of covariation between traits (Klingenberg 2003a). Levels of covariation between FA are used in developmental contexts, because genetics and the environment only affect both sides of a structure, equally (Klingenberg 2008, 2014, 2015). Since FA is a product of developmental stress, this type of data is a great representation of development-based variation (Klingenberg 2008, 2014, 2015). Modularity of FA, or developmental modularity, can stem from shared developmental origins within modules, e.g. the traits in the face comprise one module, because they were all formed from the neural crest, or from parallel variation through the interaction of developmental pathways (Klingenberg 2003a, 2008).

Fluctuating asymmetry can be quantified in complex structures such as the skull, through geometric morphometrics methods. First, by placing landmarks in the corresponding location on right and left sides of the skull. Then, the original configuration of landmarks, along with a copy that has been reflected to its mirror image are combined in a Procrustes fit (Klingenberg 2015). The differences between the original configuration and the mirrored copy, represent asymmetry of shape. The magnitude of asymmetry is given by the distances between the original configuration and mirrored copy in the shape space, in units of procrustes distances (Klingenberg and McIntyre 1998).
1.2 Quantitative Genetics

Quantitative genetics studies the inheritance of quantitative traits, meaning continuous traits that form a continuously graded series from one extreme to the other. It was first described by Fisher (1918), who managed to combine biometrics and Mendelian inheritance theory, which are the basis for modern quantitative genetics (Falconer and Mackay 1996; Tabery 2004; Roff 2007). Quantitative genetics started as an extension of the Mendelian theory, in an effort to understand the polygenic inheritance (many loci with small effects) of measurable traits, instead of Mendelian monogenic inheritance. As population genetic studies advanced, along with statistical analysis tools and the field of genetics, quantitative genetic analyses progressed into a wide range of analyses with different applications (Falconer and Mackay 1996; Lynch and Walsh 1998). Even though quantitative genetics can and have been used to address a variety of issues, in this thesis, I apply the method to obtain the genetic basis of skull shape, to estimate the response to hypothetical selection, and to estimate selection on skull shape.

1.2.1 Estimation of genetic parameters

One way the genetic basis of quantitative traits can be estimated is through data on the degree of relatedness (kinship) between the studied individuals, or pedigree. These methods work by relating the degree of genealogical relatedness with phenotypic similarity, thus estimating the amount of phenotypic variation that is explained by the genotype. And even though obtaining pedigree information for wild populations can be considerably harder than for laboratory-raised animals, extrapolation from laboratory studies to a more general evolutionary context may be limited. Studies of the quantitative genetics of wild populations have revealed trade-
offs that are not apparent at the phenotypic level, such as maternal effects, and genotype by environmental interaction (Wilson et al. 2006; Roff 2007). There are different models available for the estimation of genetic parameters, but one model is particularly fit for multivariate pedigree analysis in wild populations. The mixed model, or animal model, uses a mix of 'fixed' and 'random' effects as explanatory terms in general linear models and analysis of variance (Lynch and Walsh 1998). In our analysis, fixed effects are factors that somehow affect the shape of the skull, such as sex and horn morph. Random effects are used to describe factors with multiple levels sampled from a population of possible values, (e.g. one individual within a population, or one population within several populations) for which the analysis provides an estimate of the variance of the effects. In the case of an animal model, the random effects of interest are the additive genetic value (or its multivariate equivalent, the G matrix) of individual animals (Kruuk 2004). The additive genetic values, or G matrix, give us an estimate of the amount of morphological variation that is derived from genetic variation (Lynch and Walsh 1998). In the context of geometric morphometrics, the G matrix is a covariance matrix, from which we are able to visualize the shape changes that derive from genetic variation.

The animal model has been used in several natural population studies, among which are the collared flycatcher (Ficedula albicollis) (Gustafsson 1986; Merila et al. 2001; Sheldon et al. 2003); the great tit (Parus major) (McCleery et al. 2004); the long-tailed tit (Aegithalos caudatus) (MacColl and Hatchwell 2003); the blue tit (Parus caeruleus) (Charmantier et al. 2004; Charmantier et al. 2004a); house sparrows (Passer domesticus) (Jensen et al. 2003). Other studies include Soay sheep (Milner et al. 2000; Coltman et al. 2001b; Coltman et al. 2003; Wilson et al. 2004), bighorn
sheep (Coltman et al. 2001a; Coltman et al. 2005; Wilson et al. 2005; Poissant et al. 2008; Réale et al. 2009), macaques, (Gagliardi et al. 2010), red squirrels (McAdam and Boutin 2003; Réale et al. 2003a; Réale et al. 2003b), and salmon (Garant et al. 2003).

To estimate the variance components of the animal model, we used a restricted maximum-likelihood analysis (REML). As any other maximum-likelihood analysis, REML identifies the set of parameters that maximizes the likelihood of observing the actual data (Lynch and Walsh 1998). But the usual ML approach assumes that fixed effects are known without error (Lynch and Walsh 1998), and so fail to account for the degrees of freedom lost in estimating fixed effects (Kruuk 2004). As a consequence, ML analyses tend to estimate downwardly biased residual variances (Lynch and Walsh 1998). To correct this bias, REML estimators maximize only the portion of the likelihood that does not depend on the fixed effects (Lynch and Walsh 1998; Kruuk 2004).

1.2.2 Breeder's Equation

The breeder’s equation was developed by Lush (1937), so that breeders could predict the phenotype resulting from artificial selection. The response to selection as given by the breeder's equation is

$$\Delta \bar{z} = \left(\frac{V_A}{V_p}\right)s = h^2s$$

where $\bar{z}$ is the mean of a trait in the population, $\Delta \bar{z}$ is the response in the mean of $z$, $h^2$ is the heritability of $z$, that is, the ratio of its additive genetic variance $V_A$ to its total phenotypic variance, $V_p$, and $s$ is the selection differential, the difference
between the means of selected and unselected parents (Lande 1979; Lynch and Walsh 1998).

Heritability is the most popular way of measuring inheritance for scalar traits. But heritability of multivariate traits such as shape, yields a more complex interpretation, that differs from the definition of heritability in an univariate context (Klingenberg 2003b; Klingenberg and Monteiro 2005). Unlike univariate estimates, multivariate heritability is a vector, composed by magnitude and direction. When estimating heritability of shape, the magnitude and direction of shape change are not really separable, because the direction in which an individual differs from the others has a genetic basis, therefore, it has to be taken into account. Since there are no effective ways to estimate shape heritability in the same sense of univariate heritability, I do not estimate heritability of skull shape per se in my analyses. But I do estimate heritability for fluctuating asymmetry of shape. Because fluctuating asymmetry is randomly distributed in the skull, only its magnitude is relevant.

Although there is no multivariate version of heritability, we can still use the genetic and phenotypic matrices to estimate the response to selection in multivariate traits, such as the skull. The multivariate version of the breeder’s equation is

$$\Delta \mu = GP^{-1} s = G \beta$$

where $\Delta \mu$ is the change in mean shape in response to selection, $G$ is the genetic covariance matrix, $P$ is the phenotypic covariance matrix, $s$ is the selection differential, and $\beta$ is the selection gradient, which is the change in relative fitness per unit of shape change (Lande 1979). Selection gradients ($\beta$), can be interpreted as the force of directional selection acting on the phenotype, whereas selection differentials
(s) do not distinguish between direct selection from indirect selection due to covariation between traits (Lande 1979; Lande and Arnold 1983).

1.2.3 Estimation of Natural Selection

The breeder's equation was the starting point for the development of several quantitative genetics methods. One of the methods derived from the breeder’s equation is the estimation of natural selection by Lande and Arnold (1983). Natural selection is at the core of the evolutionary biology field, which explains the sheer amount of studies that estimate selection, in various species, using a multitude of methods (for a review see (Kingsolver et al. 2001)). Unlike the evolutionary response to selection, which depends on genetic variation, natural selection acts on phenotypes, regardless of their genetic basis (Lande and Arnold 1983). The strength and type of natural selection can be determined from plotting fitness against phenotypic value. Fitness is how much an individual contributes, in term of genes, to the next generation, and it can be measured as total number of offspring, mating success, litter size, disease resistance, ovulation rate, etc (Endler 1986; Falconer and Mackay 1996).

Methods for estimating selection on multivariate characters were first developed by Lande and Arnold (1983), where they state that "forces of directional selection on the characters are given by the average gradient of the surface of individual relative fitness, weighted by the phenotype distribution" as shown in the equation

\[ \beta = P^{-1} \text{Cov}[w, z] \]
where $\beta$ is the linear selection gradient, $P$ is the phenotypic matrix, $w$ is relative fitness, and $z$ is a character vector (Lande and Arnold 1983). The directional selection gradient ($\beta$) is then defined as the partial regression coefficients of individual relative fitness on the characters. Meanwhile, disruptive and stabilizing selection are estimated from a quadratic multiple regression

$$\gamma = P^{-1}Cov[w,zz^T]P^{-1}$$


1.2.4 Hypothetical Selection

Using the multivariate breeder’s equation and the genetic covariance matrix, we can apply an arbitrary $s$ or $\beta$, with any chosen magnitude and direction of selection, and obtain the response to such selective forces. Predicting the outcome of selection can be very useful, especially among breeders trying to improve their crop (or cattle, etc.). In fact, the method was developed with the focus on agricultural purposes, hence the name 'breeder’s equation' (Lush 1937; Kelly 2011). But the breeder’s equation has been used for a diversity of purposes in evolutionary biology (Falconer and Mackay 1996; Kelly 2011). One particular purpose, is to assess how the structure of the genetic covariance matrix (G matrix), or more specifically, its integration, affects the outcome of selection.
Integration on the \( G \) matrix, or genetic integration, is present when few genes are responsible for morphological variation on multiple traits. Genetic integration is usually caused by pleiotropy and linkage disequilibrium (Lande 1980). Pleiotropy is when one gene influences two or more seemingly unrelated phenotypic traits, and linkage disequilibrium is the non-random association between alleles at different loci. The effects of genetic integration on the outcome of selection has been extensively debated, particularly on studies concerning phenotypic evolution (Wagner 1984, 1988, 1989; Schluter 1996; Eroukhmanoff 2009; Walsh and Blows 2009). There is a consensus among such studies, that genetic integration constrains adaptation by deflecting the direction of selection towards the direction of greatest genetic variance, or genetic line of least resistance (Lande 1979; Falconer and Mackay 1996; Schluter 1996; Klingenberg and Leamy 2001; Steppan et al. 2002; Gerhart and Kirschner 2007; Klingenberg et al. 2010).

As mentioned before, a highly integrated structure presents strong covariation among its characters, which constrains phenotypic variation. The representation of these integrated structures in the shape space (trait \( x \) versus trait \( y \)) is shown as tightly correlated points, with most of the variation represented by only one of the axes (Fig. 1). Because most of the phenotypic variation is concentrated in one axis (the genetic line of least resistance), if the optimum phenotype (where its fitness is highest) is located in a different direction than the axis of most genetic variation, then the phenotype is restrained from achieving its optimum, despite positive selection towards the optimum phenotype (Schluter 1996). The implications of these constraints are selection responses that differs from the actual selective pressures imposed in the phenotype (Eroukhmanoff 2009).
Figure 1. Representation of genetic integration in the shape space. Each data point represents the relationship between traits x and y for each specimen, the concentration of the variation in one direction is characteristic of integrated structures and it is highlighted here by the grey ellipse and lines. The three arrows represent the direction of selection (on the left), the direction of the response to selection (in the middle), and the direction of the genetic line of least resistance (on the right).

Besides measuring the overall integration of the genetic covariance matrix, we can also assess exactly how (in which direction) the integration of the G matrix deflects selection. We do that by comparing the direction of selection applied, with the direction of the response to selection. This is where hypothetical selection analyses and the breeder’s equation are extremely useful. Except by manipulated breeding experiments, or artificial selection (Hoffman and Parsons 1989; Atchley et al. 1990), it is impossible to distinguish the outcome of selection, from the selection forces acting on the population, especially when it comes to wild populations. But by applying an arbitrary s or β into the breeder’s equation and obtaining the response to these hypothetical selections, we can actually measure the how much and in which way selection is being deflected by the genetic covariance matrix (G matrix) (Klingenberg and Leamy 2001; Martínez-Abadías et al. 2009; Klingenberg et al. 2010; Martínez-Abadías et al. 2012). Unlike the estimates of natural selection,
arbitrary selection gradients/differentials allows us to know the strength and direction of the selection being applied in the population, and compare it with the strength and direction of the response to selection (Klingenberg and Leamy 2001; Martínez-Abadías et al. 2009; Klingenberg et al. 2010; Martínez-Abadías et al. 2012). The difference between the applied selection, and the response to selection can be interpreted as the influence of the genetic covariance matrix (G matrix) in the response to selection (Klingenberg and Leamy 2001; Martínez-Abadías et al. 2009; Klingenberg et al. 2010; Martínez-Abadías et al. 2012).

1.3 Soay sheep

1.3.1 Historical background

The Soay sheep (Fig. 2) carry this name due to the Scottish island they were limited to, in the archipelago of St Kilda. The origin of Soay sheep is unknown, but they are thought to be the most primitive form of domestic sheep in Europe (Boyd et al. 1964), resembling both the original wild species, the mouflon, and the domesticated Neolithic sheep which were first brought to Britain in about 5000 BC, reaching Britain’s islands (such as St Kilda) between three to four thousand years ago (Campbell 1974). Eventually, this primitive breed was replaced elsewhere by larger, more productive breeds, but a population of such sheep was left on Soay, and their existence was protected by the difficulty of access since the island is completely cliff-bounded (Boyd et al. 1964; Clutton-Brock and Pemberton 2004).

In 1930, the island of Hirta, a populated island also in the archipelago of St Kilda, was evacuated. The entire community and their animals left the island. The St. Kilda islands were sold then, to the Marquis of Bute, who transferred 107 Soay sheep
from Soay to Hirta in 1932. These sheep remained there, breeding and grazing, unmanaged, and now their descendants inhabit the island (Jewell et al. 1974; Clutton-Brock et al. 2004b).

1.3.2 Anatomical description

The Soay sheep have the appearance of an unimproved and primitive domestic sheep, they are small, narrow-bodied, with relatively long legs, short tails and narrow faces (Boyd et al. 1964; Campbell 1974; Clutton-Brock and Pemberton 2004). Adult males can reach 36 kg or more in weight in good condition, and adult females about 25 kg (Doney et al. 1974). Their main difference from wild species is that the Soay sheep coat is composed mainly of wool (Doney et al. 1974). They resemble mouflons in size and in the light markings on the face and pale-coloured rumps and stomachs that are present in most ‘wild-type’ Soay sheep (Doney et al. 1974; Clutton-Brock et al. 2004b).

Most of the rams have normal, spiral horns (85%)(Fig. 2) but a few individuals (15%) have small, deformed (scurred) horns (Fig. 2). The females have different proportions of horn types, a minority of them have normal horns, which are smaller than in males, and the majority is either scurred (35%) or hornless (40%)(polled)(Clutton-Brock et al. 2004b). Soay sheep are unusual in having an inherited polymorphism for horn development (Doney et al. 1974).
1.3.3 Population monitoring

Over time, individual naturalists visited the islands mainly to observe seabirds, wrens and mice; none were interested enough in the sheep to organize a census. In 1939, the owner estimated about 500 sheep on Hirta; in 1947, James Fisher estimated 400 to 450; in 1948, Fraser Darling estimated between 650 and 700 (Campbell 1974). In 1952 census work began, and in 1959 the sampling and marking programme started (Boyd et al. 1964). The census data provide a geographical distribution of the population through the hill, valley and cliff-terrace systems; they also provide in each area the proportions of ewes, lambs and rams (Boyd 1974). Even though the censuses started from 1952 onwards, the data obtained between 1968 and 1985 are unreliable (Clutton-Brock et al. 2004a).
Research on the sheep was restarted in 1985 by members of the Departments of Zoology and Physiology at Cambridge (T.H. Clutton-Brock, S.D. Albon and P.A. Jewell), but the project rapidly came to involve several other groups. Over the last twenty years the project has involved members of the Institute of Cell, Animal and Population Biology at Edinburgh, the Institute of Biological Science at Stirling, the Institute of Zoology, Department of Biological Sciences at Imperial College (London), the Centre for Ecology and Hydrology (Banchory) and the Macaulay Institute (Aberdeen) in addition to staff from Cambridge. Though the work initially focused on population demography and dynamics, specialists in related disciplines have extended the study to involve the genetics of the sheep, their impact on the plant population and the effects of parasite on their survival and breeding success (Clutton-Brock and Pemberton 2004).

Since 1985, members of the Soay sheep project have monitored the breeding success, growth, habitat use and survival of virtually all the sheep living in the village bay area of Hirta. For this whole period, over 95% of sheep that live on the area have been marked with colour-coded ear tags. Once they are deceased, at least 85% of the bodies of the all animals are found. For anatomical characters investigation, the skull and foreleg are collected, cleaned and stored (Clutton-Brock and Pemberton 2004).

Monitoring studies showed a persistent fluctuation in this island population, the numbers of animals are unstable due to periodic crashes, when a large number of animals die (Jewell et al. 1974). Soays rut in late autumn and bear their lambs in spring, but unlike most wild sheep, females normally conceive in their first autumn and twinning rates are relatively high (Coulson et al. 1999; Clutton-Brock et al. 2004a). Sheep populations on St Kilda are unmanaged and their survival and
reproductive rates are strongly influenced by variation in food availability caused by changes in population density and climate (Coulson et al. 1999; Coulson et al. 2001). Changes in population density affect early development, generating variation in juvenile mortality and breeding success between cohorts (Clutton-Brock et al. 1992; Clutton-Brock et al. 1996; Coulson et al. 1999; Coulson et al. 2001; Clutton-Brock et al. 2004a).

1.4 Authors’ Contributions and Thesis Outline

The entirety of this thesis was written by the author (Elis Damasceno), with suggestions from the supervisor (Chris Klingenberg). So far, no one else has participated in the construction of this thesis, but co-authors will be included prior to the submission of the manuscripts to journals. The following people will be included in all of the manuscripts: Tim Clutton-Brock has started the research in St. Kilda and was responsible for it between 1985 and 1995; Josephine Pemberton (University of Edinburgh) has taken charge of the project since 1995, they are responsible for all the research on St. Kilda, including the collection of skulls; Jill Pilkington carries out the field work in St. Kilda since 1993, tagging the animals and taking blood samples, making population censuses, and collecting data on parasites, etc. Camilo Berenos will be included in the all manuscripts except Chapter 4, for carrying out the genetic analysis using SNPs and building the pedigree.

The structure of each chapter derived from discussions between student and supervisor. All the analysis and data collection were executed by the student. The thesis is presented in the alternative format. We chose the alternative format because the many results I obtained can be divided in self-contained chapters, each focusing
in a sub-area of evolutionary biology. By submitting the thesis in the alternative format, we seek to speed the publication process, by having manuscripts ready for submission by the time I finish the PhD. In the second chapter, I analyse integration and modularity in the skull of Soay sheep. I use phenotypic, genetic, and fluctuating asymmetry data to build covariance matrices, and then measure overall integration, and test modularity hypothesis in all of them. I compare the results of different matrices to make assumptions on the origin of modularity and its evolutionary consequences. In the third chapter, I assess the effects of genetic integration in the outcome of selection by using the breeder's equation to estimate response to selection from hypothetical selection regimes. In the fourth chapter, I estimate natural selection on skull shape, using a new method that combines the analyses from Lande and Arnold (1983) and geometric morphometrics methods. In the fifth chapter, I analyse the possibility of a heterochronic process affecting the Soay sheep population. I then link this process to changes in climate throughout the past 30 years. Finally, in the sixth chapter, I calculate fluctuating asymmetry levels in the skull shape and test different environmental, and morphological variables that might have an influence on it. I also estimate the heritability and dominance of FA.

1.5 Literature Cited


effects. Genetics 122:223-234.
Zoologist 36:36-43.
multivariate genetic constraints: a geometric view of adaptation. Annual Review of
Kruuk. 2004. Maternal genetic effects set the potential for evolution in a free-living
Variation for Body Size: Using Random Regression Models in a Wild Ungulate
L. E. B. Kruuk. 2006. Quantitative genetics of growth and cryptic evolution of body
biologists. Academic Press, Waltham, MA.
2. Chapter 2:
Multi-level integration and modularity in the skull of Soay sheep

Abstract

Integration describes the intensity in which traits covary with each other, whereas modules are groups of traits that are more integrated within themselves then with other traits. Integration among traits can have several origins, such as development, genes or environment. But little is known about the interaction of the different effects as multi-level studies on integration and modularity are scarce. Studies comparing levels of integration across different ontogenetic stages are also scarce. I tested two modularity hypothesis in different levels of variation (phenotypic, genetic and developmental) of the Soay sheep skull, and then I analysed the best hypothesis (face and neurocranium) across different ontogenetic levels. I found low levels of integration in all levels. I found modularity, in all levels of variation which implies that modularity in one level influences the variation in other levels. Development and bone remodeling seem to play a major role in dividing the skull into face and neurocranium. I also found that yearlings are the most modular, and that adults are the least modular, thus confirming the palimpsest hypothesis. The presence of modules in the sheep might explain the diversity of morphological variation found among sheep breeds. Learning about shape variation in domestic animals can be a great asset to understanding processes of morphological differentiation in natural species.

2.1 Introduction

An organism is composed of several different parts, such as limbs, organs, bones, and each part is, at some degree, independent from each other. But some coordination between the parts, especially during development, is necessary to maintain the proportionality, and thus functionality of the organism as a whole (Klingenberg 2008). This coordination, or covariation between different traits is
known as integration (Olson and Miller 1958; Klingenberg 2008). Morphological
traits are considered to be highly integrated when they covary strongly with each
other, meaning that variation in one trait will be associated with a corresponding
variation in another trait (Olson and Miller 1958).

Even though the concept of morphological integration has been around for at
least half a century, the past decade has seen an increased effort to research the
subject (Klingenberg 2013). Studies on integration and modularity assess how much
the traits in one structure covary with each other, and if this covariance is organized
into modules, where traits are more strongly correlated within the module than with
other module(s) (Cheverud 1982; Klingenberg 2008, 2009). Understanding trait
variation within a structure is key to understanding evolutionary processes, since
variability of traits is an indication of its capacity to evolve, or evolvability (Cheverud

Modularity and integration can be found in different levels of variation,
stemming from differences in environment, genetics, developmental processes,
function, among others (Cheverud et al. 1997; Klingenberg et al. 2001; Klingenberg
et al. 2003; Armbruster et al. 2004; Badyaev and Foresman 2004; Klingenberg et al.
2004). Morphometric data, particularly, have been used to characterize
developmental, genetic, functional, and evolutionary modules (Klingenberg 2008).
And recent advances in geometric morphometric methods have given us means to
measure integration and modularity in each level of variation separately (Klingenberg
2009, 2010).

Integration in phenotypic variation can be analysed through the variation
among individuals, and it is the most commonly studied level of integration.
Developmental integration, on the other hand, is inferred by random differences between left and right sides, or asymmetries of a given trait. During development, these random asymmetries, or fluctuating asymmetries (FA) can be generated through disturbances in the developmental processes (developmental instability). These disturbances are usually caused by poor environmental conditions (Badyaev et al. 2000; Marchand et al. 2003; Oleksyk et al. 2004; Sánchez-Chardi et al. 2013; Lazić et al. 2015; Lezcano et al. 2015; Maestri et al. 2015), and low genetic quality (Debat et al. 2000; Rego et al. 2006; Carreira et al. 2008; Albarrán-Lara et al. 2010; Mikula et al. 2010; Nouvellet et al. 2011; Quinto-Sánchez et al. 2015). When the asymmetries of different traits covary, it could mean that these traits shared a common developmental process or pathway, or that these pathways interacted at some point during development (Klingenberg 2008). The different sides of a symmetrical structure share the same genes, and experience the same environment (unless the organism is sessile). Therefore, studying FA is an effective way of eliminating genetic and environmental variation in structures (Klingenberg 2003, 2008, 2010, 2013, 2014). Finally, genetic modularity can be assessed by estimating additive genetic covariances through quantitative genetics methods, such as the animal model (Lynch and Walsh 1998; Klingenberg and Leamy 2001; Klingenberg 2003; Kruuk 2004; Klingenberg 2008; Martínez-Abadiás et al. 2009; Martínez-Abadiás et al. 2012; Klingenberg 2013, 2014)(See Kruuk [2004] for a review on the animal model). The estimation of genetic covariances requires that the relationship between individuals is known, as well as other factors that might influence morphology, such as sex and age (Falconer and Mackay 1996; Lynch and Walsh 1998; Wilson et al. 2005; Wilson et al. 2010).
Comparing the different levels of variation can help us understand both the origins of modularity and how the different levels of variation interact (Breuker et al. 2006). Strong similarities between those levels may indicate a strong interaction between the different factors, whereas stronger modularity in one of the levels is an indication of its origins (Breuker et al. 2006; Klingenberg 2008). Comparing the patterns of covariance for fluctuating asymmetry and individual variation, specifically, can be useful to infer whether the same processes are contributing to developmental stability and canalization, both of which are measures of phenotypic constancy (Debat et al. 2000; Debat and David 2001; Hallgrimsson et al. 2002). Developmental stability is the ability to reduce developmental noise, whereas canalization is the ability to reduce the effects of environmental and genetic variability on the phenotype (Palmer 1994; Debat et al. 2000; Klingenberg 2003). Further evidence on the nature of developmental stability and canalization can be achieved by comparing FA and individual variation in different ages (Hallgrimsson 1993; Swaddle and Witter 1997; Aparicio 1998; Hallgrimsson 1999; Kellner and Alford 2003; Pélabon et al. 2006).

When analyzing complex structures, such as the skull, the use of multivariate analyses is imperative, and for modularity analyses, it is recommended that the sample size be substantially larger than the dimensionality of the data, especially when contemplating quantitative genetics methods as well (Klingenberg 2013). Large samples, with known pedigree for each specimen, are very difficult to obtain, hence the scarcity of studies with genetic modularity (Klingenberg and Leamy 2001; Klingenberg et al. 2010; Martínez-Abadías et al. 2012; Klingenberg 2013, 2014). In this study, we were able to obtain an excellent sample size, with known pedigree. We
tested two modularity hypotheses in the skull of Soay sheep, the face and
neurocranium hypothesis (FN), and the face, neurocranium and oral cavity
hypothesis (OC) (Table 1, Fig. 3). We tested the hypotheses across three levels:
individual, developmental and genetic. Then, we used the best hypothesis to compare
the modularity levels between four ontogenetic stages: yearlings, young adults, adults
and old animals (Clutton-Brock and Pemberton 2004).

2.2 Methods

The Soay sheep population under study here, come from St. Kilda
archipelago, in Scotland (Clutton-Brock and Pemberton 2004). Every sheep is tagged
individually within a few days after birth, and a skin sample is taken for paternity
analysis (Boyd et al. 1964). The population has been monitored since 1964 and the
skulls of deceased animals were collected from 1976 until 2006. All the skulls
collected in St. Kilda were deposited at the National Museum of Scotland, in
Edinburgh.

2.2.1 Shape data

My sample consists of 1,508 skulls, of which 743 are males and 765 are
females, 703 are yearlings, 271 are young adults, 362 are adults, and 161 are old
animals. A single observer (EMD) digitized 50 landmarks in three dimensions (Table
1, Fig. 1) using a MicroScribe G2X. Each skull was digitized twice to correct for
measurement errors. The raw landmark data was then imported into MorphoJ
(Klingenberg 2011), where I ran a Procrustes fit to remove effects of size, position
and orientation, followed by a Procrustes ANOVA (Klingenberg and McIntyre 1998;
Klingenberg et al. 2002). Because sheep skulls have object symmetry (a structure with
two mirrored sides, with the plane of symmetry passing through the structure), the
Procrustes fit not only corrects the data for measurement errors but also divides
variation into two components: symmetric (individual) and asymmetric (individual
by side) (Klingenberg et al. 2002). The symmetric component of shape is the average
between left and right sides and comprises variation between individuals, while the
asymmetric component holds the differences between sides (fluctuating asymmetry -
FA) (Klingenberg and McIntyre 1998; Klingenberg et al. 2002).

Figure 1. Scheme showing landmark locations on A) dorsal view; B) lateral view; and C) ventral view.
Table 1. List of landmark numbers and their respective anatomical description, and which subset it belongs for each hypothesis tested.

<table>
<thead>
<tr>
<th>Landmark numbers</th>
<th>Description</th>
<th>Subset in Hypothesis FN</th>
<th>Subset in Hypothesis OC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/9</td>
<td>Premaxillary-maxillary suture</td>
<td>Face</td>
<td>Face</td>
</tr>
<tr>
<td>2/10</td>
<td>Lacrimal-maxillary suture</td>
<td>Face</td>
<td>Face</td>
</tr>
<tr>
<td>3/11</td>
<td>Lacrimal-maxillary-malar suture</td>
<td>Face</td>
<td>Face</td>
</tr>
<tr>
<td>4/12</td>
<td>Lacrimal-frontal suture, posterior</td>
<td>Face</td>
<td>Face</td>
</tr>
<tr>
<td>5/13</td>
<td>Frontal-malar suture, supra-orbital process</td>
<td>Face</td>
<td>Neurocranium</td>
</tr>
<tr>
<td>6/14</td>
<td>Frontal-parietal-temporal suture</td>
<td>Neurocranium</td>
<td>Neurocranium</td>
</tr>
<tr>
<td>7/15</td>
<td>Asterion, posterior at occipital-parietal-temporal suture</td>
<td>Neurocranium</td>
<td>Neurocranium</td>
</tr>
<tr>
<td>8/16</td>
<td>External auditory meatus</td>
<td>Neurocranium</td>
<td>Neurocranium</td>
</tr>
<tr>
<td>17</td>
<td>Nasion, nasal-frontal suture</td>
<td>Face</td>
<td>Face</td>
</tr>
<tr>
<td>18/19</td>
<td>Supra-orbital foramen</td>
<td>Face</td>
<td>Face</td>
</tr>
<tr>
<td>20</td>
<td>Bregma, frontal-parietal suture</td>
<td>Neurocranium</td>
<td>Neurocranium</td>
</tr>
<tr>
<td>21</td>
<td>External Occipital Protuberance</td>
<td>Neurocranium</td>
<td>Neurocranium</td>
</tr>
<tr>
<td>22</td>
<td>Opisthion: dorsal lip of foramen magnum</td>
<td>Neurocranium</td>
<td>Neurocranium</td>
</tr>
<tr>
<td>No.</td>
<td>Description</td>
<td>Region 1</td>
<td>Region 2</td>
</tr>
<tr>
<td>------</td>
<td>--------------------------------------------------</td>
<td>---------------------------</td>
<td>---------------------------</td>
</tr>
<tr>
<td>23/24</td>
<td>Occipital Condyle</td>
<td>Neurocranium</td>
<td>Neurocranium</td>
</tr>
<tr>
<td>25</td>
<td>Basion, ventral lip of foramen magnum</td>
<td>Neurocranium</td>
<td>Neurocranium</td>
</tr>
<tr>
<td>26/27</td>
<td>Basilar tubercule</td>
<td>Neurocranium</td>
<td>Neurocranium</td>
</tr>
<tr>
<td>28</td>
<td>Posterior-basioccipital suture</td>
<td>Neurocranium</td>
<td>Neurocranium</td>
</tr>
<tr>
<td>29</td>
<td>Presphenoid-ethmoid spine</td>
<td>Neurocranium</td>
<td>Oral region</td>
</tr>
<tr>
<td>30/34</td>
<td>Zygo-maxillary inferior</td>
<td>Face</td>
<td>Oral region</td>
</tr>
<tr>
<td>31/35</td>
<td>Ethmoidal foramen, ventral lip</td>
<td>Neurocranium</td>
<td>Oral region</td>
</tr>
<tr>
<td>32/36</td>
<td>Temporal-malar suture, zygomatic process, anterior projection</td>
<td>Face</td>
<td>Neurocranium</td>
</tr>
<tr>
<td>33/37</td>
<td>Temporal-malar suture, zygomatic process, posterior projection</td>
<td>Face</td>
<td>Neurocranium</td>
</tr>
<tr>
<td>38/39</td>
<td>Medial end of premaxilla, anterior</td>
<td>Face</td>
<td>Face</td>
</tr>
<tr>
<td>40</td>
<td>Premaxillary-maxillary suture, posterior</td>
<td>Face</td>
<td>Oral region</td>
</tr>
<tr>
<td>41</td>
<td>Maxillary-palatine suture, anterior at midline</td>
<td>Face</td>
<td>Oral region</td>
</tr>
<tr>
<td>42</td>
<td>Palatine, posterior at midline</td>
<td>Face</td>
<td>Oral region</td>
</tr>
<tr>
<td>43/44</td>
<td>Sphenoid-palatine suture on the pterygoid process</td>
<td>Neurocranium</td>
<td>Oral region</td>
</tr>
<tr>
<td>45/48</td>
<td>Anterior 1st Premolar</td>
<td>Face</td>
<td>Oral region</td>
</tr>
</tbody>
</table>
2.2.2 Ontogenetic stages

The different ontogenetic stages of the Soay sheep are usually described as: lambs (younger than six months); juveniles (between six months and one year old); yearlings (between one and two years old); and adults (older than two years old) (Clutton-Brock et al. 2004). But these classification are based on milestones in the sheep’s lifetime, such as weaning and rutting, and not on skull morphology. In order to group the animals into different ontogenetic stages according to skull morphology, I ran a Principal Component Analysis (PCA) on the matrix of individual variation (P matrix). The analysis showed three major groupings, based on morphological similarity: yearlings (1-1.9 years), young adults (2-2.9 years), and adults older than three years (Fig 2). Even though there is not much variation in skull shape within adults, I wanted to assess whether morphological changes continue in adulthood, thus designating a group for older animals, with eight years of age or more. Animals older than eight years has been classified as belonging to a different age group in previous analysis in Soay sheep (Clutton-Brock et al. 2004). Because of the distinct morphology found among different age groups, the covariance matrices I used on subsequent PCAs and modularity analyses were pooled-within age groups.
2.2.3 $G$ matrix

I used pedigree information to obtain the genetic basis (genetic covariance matrix - G matrix) of skull shape. Maternity was inferred by observation of maternal care behaviour (Robinson et al. 2006), and paternity was inferred by analyses on single nucleotide polymorphisms (SNPs), using maximum-likelihood implemented on CERVUS (Marshall et al. 1998; Robinson et al. 2006). For each lamb, a log-likelihood ratio score was calculated for each possible father, and criteria for assigning paternity to the male with the highest score with 95% and 80% confidence were obtained by simulation (Pemberton et al. 1999; Pemberton et al. 2004).

To obtain the genetic covariance matrix (G matrix), I performed a Restricted Maximum-Likelihood based on an animal model (Lynch and Walsh 1998; Kruuk
2004; Kruuk and Hill 2008; Wilson et al. 2010) using Wombat software (Meyer 2007). Due to computational limitations, I used 24 principal components (PC) scores from a pooled within-age PCA (the maximum number supported by Wombat) instead of the raw data to calculate the \( G \) matrix (Klingenberg and Leamy 2001; Klingenberg et al. 2010; Martínez-Abadíás et al. 2012). Sex and horn morph were set as fixed effects and centroid size as a covariate, while individual and year of birth were set as random effects. The \( G \) matrix obtained through Wombat was then imported into MorphoJ (Klingenberg 2011), where it was converted into landmark coordinates for further analyses (Klingenberg et al. 2010).

2.2.4 Matrix correlation

To test the similarities between the different covariance matrices (phenotypic, genetic and FA) I ran a matrix correlation test, followed by a matrix permutation test in order to test the correlation statistically. I also tested the similarities between the \( P \) matrix and the FA covariance matrix in the different ontogenetic stages. I included the diagonal blocks of the covariance matrices (Klingenberg and McIntyre 1998; Klingenberg et al. 2002) and used landmarks for the permutation test. The use of landmarks in the permutation test differs from the way matrix correlations are usually carried out (outside a geometric morphometrics context). Normally, the coordinates would be permuted to analyse the correlation between two matrices, but in geometric morphometrics, coordinates from the same landmark are not independent of each other (Klingenberg and McIntyre 1998). So, it makes sense to maintain these associations of coordinates during permutation, and to permute the landmarks,
instead of individual coordinates in the covariance matrix (Klingenberg and McIntyre 1998).

2.2.5 Overall integration

Because the strength of integration depends on how much different traits covary with each other, one way that morphological integration can be measured is through the variance of the eigenvalues of each covariance matrix. From the PCAs on the pooled-within age groups covariance matrices $P$, $G$, and FA, we obtained both the measures of the overall integration in the form of eigenvalues variances, and the shape changes associated with the first principal component (PC1) of each matrix. The first principal component is the axis with the most amount of variation in the covariance matrix. The eigenvalues variances were scaled by total variation and number of variables, and they can range from zero to one (Young 2006).

2.2.6 Modularity hypotheses

I tested two modularity hypothesis in which the skull was divided into: (1) face and neurocranium (FN); and (2) neurocranium, oral cavity and dorsal portion of the face (OC) (Table 1, Fig. 3). The face and neurocranium are well known modules in the skull since they have different developmental origins (neural crest and paraxial mesoderm respectively) (Couly et al. 1992; Jiang et al. 2002; Noden and Trainor 2005) and they have been found in many mammals (Cheverud 1982, 1995; Hallgrimsson et al. 2004; Hallgrimsson et al. 2007; Drake and Klingenberg 2010). But because of the complex and important functions performed by the oral cavity (such as mastication and suckling), some studies describe it as a different module altogether (Marroig and Cheverud 2001; Goswami 2006).
In order to measure levels of integration and modularity in the skull, I used Escoufier’s RV coefficient (Escoufier 1973) and to test whether the modularity hypothesis is biologically accurate, I used the method of comparing alternative partitions (Klingenberg 2009). The alternative partitions method works by randomly reshuffling landmarks within one of the subsets. After reshuffling the landmarks, the parts are reassembled, a new Procrustes fit is done and the RV coefficient is recalculated. This process was repeated 20,000 times, and the new RV coefficient was computed for each time. I then compare the frequency of the RVs that are smaller than the RV generated by the original hypothesis. This analysis allows me to see whether the modules given in my hypothesis is among the best possible existing divisions of the skull by giving the percentage of random RVs that are smaller than the hypothesis’ RV. (Klingenberg 2009; Drake and Klingenberg 2010; Ivanović and Kalezić 2010; Klingenberg et al. 2010; Jojić et al. 2011, 2012; Martínez-Abadías et al. 2012; Sanger et al. 2012; Klingenberg and Marugán-Lobón 2013).
analyses, landmarks were inter-changed only within contiguous partitions. A set of landmarks is contiguous if all the landmarks are connected. The contiguity of the landmarks is given by an adjacency graph (Fig 3), that indicates which landmarks are considered to be interconnected (Klingenberg 2009).

To assess which of the two hypothesis is better fitted to the Soay sheep, I analysed the whole dataset for both symmetric and asymmetric components. I generated pooled within-group covariance matrices for the symmetric and asymmetric components. They were pooled within sex and age groups (yearlings and adults). When I tested the modularity hypothesis on the covariance matrix of the symmetric and asymmetric components, (P matrix and FA matrix, respectively) and on the G matrix estimated with Wombat. I also tested the best of the two hypotheses (FN or OC) on the P and FA matrices in all ontogenetic stages, separately.

2.3 Results

The results of the Procrustes ANOVA show that the effect of measurement errors on centroid size estimates is negligible (Table 2), but it has a greater impact on shape. The mean square procrustes distances for measurement errors is 3.48 times smaller than the mean square procrustes distances for fluctuating asymmetry (Table 3).
Table 2. Results from the Procrustes ANOVA on centroid size showing the Procrustes sum of squares (SS), the Procrustes mean squares (MS), degrees of freedom (df), Goodall’s F statistic (F) and the associated parametric P-value (p).

<table>
<thead>
<tr>
<th>Effect</th>
<th>SS</th>
<th>MS</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual</td>
<td>2736638</td>
<td>1793.3</td>
<td>1526</td>
<td>4750.3</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Error</td>
<td>576</td>
<td>0.37</td>
<td>1527</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Results from the Procrustes ANOVA on shape showing the Procrustes sum of squares (SS), the Procrustes mean squares (MS), degrees of freedom (df), Goodall’s F statistic (F) and the associated parametric P-value (p).

<table>
<thead>
<tr>
<th>Effect</th>
<th>SS</th>
<th>MS</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual</td>
<td>8.89</td>
<td>0.000076</td>
<td>115976</td>
<td>15.75</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Side</td>
<td>0.041</td>
<td>0.00062</td>
<td>67</td>
<td>127.94</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Ind. by side</td>
<td>0.49</td>
<td>0.0000048</td>
<td>102242</td>
<td>3.48</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Error</td>
<td>0.30</td>
<td>0.000014</td>
<td>218361</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2.3.1 Overall Integration

The eigenvalues variances scaled by total variation and number of variables were 0.033 for the P matrix, 0.025 for fluctuating asymmetry, and 0.066 for the G matrix (See Fig. 4 for the distribution variation among PCs). The shape changes associated with the first principal component of the P matrix include changes in the length of the snout and in the size of the braincase (Fig. 5). We found similar shape changes associated with the first principal component of the G matrix (Fig. 6). Along the first principal component of the FA matrix, we have deviations of the anterior edge of the snout towards the left and the right, relative to the rest of the skull (Fig.
7). There are also changes in the relative position of the orbits and foramen magnum (Fig. 7). Note that these shape changes are not the same as the ones from the first principal component of the P matrix, before it was pooled-within age groups (Fig. 2).

Figure 4. Bar charts showing the percentage of variance explained by each principal component for the symmetric, asymmetric, and genetic component.
**Individual variation**

Figure 5. Shape changes associated with the first principal component, obtained from a Principal Component Analysis on the P matrix. The diagrams show configurations along this shape vector at values of -0.1 (on the left) and +0.1 (on the right) Procrustes units from the average shape in the dataset.

**G matrix**

Figure 6. Shape changes associated with the first principal component, obtained from a Principal Component Analysis on the G matrix. The diagrams show configurations along this shape vector at
values of -0.1 (on the left) and +0.1 (on the right) Procrustes units from the average shape in the dataset.

Figure 7. Shape changes associated with the first principal component, obtained from a Principal Component Analysis on the FA matrix. The diagrams show configurations along this shape vector at values of -0.1 (on the left) and +0.1 (on the right) Procrustes units from the average shape in the dataset.

2.3.2 Matrix correlations

The matrix correlation showed a strong resemblance between $P$ and $G$ matrices, with a correlation of 0.96 ($P<0.0001$). The second strongest correlation was between the $P$ matrix and fluctuating asymmetry (0.62, $P=0.0004$), followed by the correlation between the $G$ matrix and fluctuating asymmetry (0.58, $P<0.0001$). The correlation between $P$ matrix and FA covariance matrix seems to decrease with age,
being 0.63 \( (P =0.0002) \) in yearlings, 0.60 \( (P =0.0001) \) in young adults, 0.58 \( (P=0.0067) \) in adults, and 0.51 \( (P=0.0011) \) in senescent animals.

2.3.3 Modularity hypotheses

Of the two hypotheses tested, the FN hypothesis was most consistent with the data. For the FN hypothesis, only 0.5\% \( (RV = 0.24) \) of alternative partitions had smaller RVs than the a priori hypothesis for the symmetric component, 1\% \( (RV = 0.23) \) for the asymmetric component, and 2\% for the \( G \) matrix \( (RV = 0.31) \). The results for the OC hypothesis is for the symmetric and genetic components, show 7\% \( (RV = 0.20) \) and 10\% \( (RV=0.28) \), respectively, of alternative partitions’ RVs being smaller than the RV generated by the hypothesis, but for the asymmetric component this hypothesis is not consistent with the data, with 77\% \( (RV=0.21) \) of the alternative partitions’ RVs being smaller than the hypothesis’ RV (Fig 8).
Figure 8. Bar charts showing the frequency in which each value of RV coefficient was generated by 20,000 different alternative partitions for both hypothesis (face and neurocranium and oral region) and each component of variation (symmetric, asymmetric and genetic) using the complete dataset. The black arrows are showing the location of the RV generated by the hypothetic partitions, which are also in display, follow by the proportion of randomly generated RVs that are smaller than the hypothesis' RV.

I only use the FN hypothesis to compare modularity levels between age groups because it is better supported than the OC hypothesis. When I tested the FN hypothesis in separate age groups I found that the skull of yearlings are more in accordance with the hypothesis than any other age group, with only 0.2% (RV = 0.28) and 0.9% (RV = 0.21) of alternative partitions having smaller RVs than the FN hypothesis for symmetric and asymmetric components respectively. I found similar
results in young adults, with 0.4% (RV=0.28) and 0.8% (RV=0.27) of alternative partitions RV coefficients being smaller than the RV coefficients generated by the FN hypothesis for symmetric and asymmetric components respectively. Animals with more than 3 years of age, especially adults, do not conform to the face and neurocranium hypothesis as well as younger animals, with 15% (RV=0.24) and 3% (RV=0.25) for adults, and 2% (RV=0.28) and 3% (RV=0.39) for old animals of alternative partitions having smaller RVs than the partitions generated by the FN hypothesis for symmetric and asymmetric components respectively (Fig 9).
Figure 9. Bar charts showing the frequency in which each value of RV coefficient was generated by 20,000 different alternative partitions for the face and neurocranium hypothesis and for symmetric and asymmetric components of variation. The black arrows are showing the location of the RV generated by the hypothetic partitions, which are also in display, follow by the proportion of randomly generated RVs that are smaller than the hypothesis’ RV.
2.4 Discussion

2.4.1 Overall Integration

The comparatively low levels of integration (Lieberman et al. 2000; Drake and Klingenberg 2010) in the symmetric component can result from the characteristics of our data. One of those characteristics is that we are analysing animals within the same population, whereas most studies have, so far, analyzed integration within species (with samples from several populations) or even higher clades (e.g. (Goswami 2005; Porto et al. 2009; Drake and Klingenberg 2010)). Because the variation within specimens from the same population is usually a lot lower than the variation between populations, the levels of integration are also lower, since integration is, in a way, also a measure of variation. Low levels of integration have also been found in human skulls (Martínez-Abadias et al 2012). This study also analysed skulls from the same population, and they also had large sample sizes (Martínez-Abadías et al. 2012). Most studies on skull integration have considerably smaller samples, along with higher integration levels. Therefore, large sample sizes might also be an explanation for the low levels of integration found in both this and the humans study. Large samples are a better representation of the variation within particular populations, and might contain a larger array of different morphologies with little variation between them, which could lead to lower levels of variation (and thus, integration) in larger samples, whereas small samples can contain morphologies from the extremes of the morphological distribution, with no intermediate morphology in the sample, which could lead to higher levels of variation (and thus, integration) in small samples.
Normally, individual variation is more integrated than fluctuating asymmetry, as found in integration analysis on dogs, wolves, and carnivorans skull (Drake and Klingenberg 2010), mouse mandibles (Klingenberg and Leamy 2001; Klingenberg 2009), and vole tooth rows (Laffont et al. 2009). Stronger integration on the genetic covariance matrix, in the other hand, do concur with previous studies do similar analysis on humans (Martínez-Abadías et al. 2009; Martínez-Abadías et al. 2012) and cricket wings (Klingenberg et al. 2010). Stronger integration in G matrices could be related to the presence of pleiotropic networks among genes (Wagner 1989; Cheverud et al. 2004; Klingenberg 2008).

2.4.2 Matrix correlations

We found a strong correlation (0.96) between P and G matrices, which was expected due to the known genotype-phenotype relationship (Wagner 1996; Wagner and Altenberg 1996) but, to some extent, also because of the part-whole relationship between the two matrices (Klingenberg et al. 2010). High correlations between phenotypic and genetic matrices has also been found in mouse mandibles (Klingenberg and Leamy 2001) and humans (Martínez-Abadías et al. 2012). The P matrix is not as strongly correlated with FA matrix (0.62) as it is with the G matrix. This indicates that the phenotype is more strongly influenced by genes then by developmental processes, even though the former still exerts some influence on the phenotype (Klingenberg 2008). The same analysis on the skull of alpine newts produced similar results (Ivanović and Kalezić 2010). The lowest correlation was found between G and FA matrices. Some correlation between these two were
expected because developmental processes mediate the expression of genetic variation in phenotypic traits (Klingenberg 2008). But they are not perfectly correlated because the expression of genetic variation is not exclusively controlled by developmental interactions.

2.4.3 Modularity hypotheses

When testing different modularity hypothesis, our results show that the FN hypothesis fits the data better than the OC hypothesis. The face and neurocranium have been found to comprise different modules in many studies (Cheverud 1982, 1995; Hallgrímsson et al. 2004; Hallgrímsson et al. 2007; Drake and Klingenberg 2010; Jojić et al. 2011), this could be explained by their different developmental origins, as the neurocranium has comes from the paraxial mesoderm and the face from neural crest (Couly et al. 1992; Jiang et al. 2002; Noden and Trainor 2005). Cheverud also hypothesises that the expansion of the brain during mammalian evolutionary history, may have led to a distinct growth pattern of the cranial vault, while face traits are more prone to functional and environmental factors even after the brain (and thus, the neurocranium) has ceased growing (Cheverud 1995).

The division of the skull into face and neurocranium seems to be recurrent in all levels of variation included in this study. Because patterns of phenotypic variation are partly dictated by genes through developmental processes, it is only expected that the same modules are found within the different levels of variation. Genetic modularity is believed to be the product of pleiotropy and linkage disequilibrium, where one gene can influence multiple traits (Nadeau et al. 2003; Klingenberg 2008). Developmental modularity, on the other hand, comes about by direct interaction
between developmental pathways (Klingenberg 2003, 2005, 2008). Modularity of individual variation is obviously affected by development, linkage disequilibrium and pleiotropy, and it is also affected by morphological plasticity (Klingenberg 2003, 2005, 2008).

Unlike previous studies (Drake and Klingenberg 2010; Jojić et al. 2012), our results show that the individual variation is just as consistent as FA in both modularity hypotheses, and in both age groups. This means that both bone remodeling and development have a great impact in dictating the patterns of integration in the skull of Soay sheep, more so than genetics. Development is known to impact modularity in the skull, and the division between face and neurocranium into modules is seen across several species of mammals (Cheverud 1982, 1995; Hallgrímsson et al. 2004; Hallgrímsson et al. 2007; Lieberman et al. 2008; Drake and Klingenberg 2010; Lieberman 2011). Whereas bone remodelling can be induced by mechanical loads, such as mastication (Kiliaridis et al. 1985). But bone remodelling can also be induced by developmental processes, thus making the two processes indistinguishable in their effects (Klingenberg 2014).

Meanwhile, the reason why skulls of yearlings and young adults are more modular than skulls of adults and old animals can be explained by the palimpsest metaphor (Hallgrímsson et al. 2009). The palimpsest hypothesis suggests that patterns of integration will change during an individual’s lifetime, with developmental and mechanical processes at each stage partially overwriting the patterns laid down earlier in life (Hallgrímsson et al. 2009). In agreement with the palimpsest hypothesis, our results, such as previous studies, show that levels of modularity vary with ontogeny (Willmore et al. 2006; Zelditch et al. 2006;
Hallgrímsson et al. 2009; Mitteroecker and Bookstein 2009; Gonzalez et al. 2011a; Gonzalez et al. 2011b).

2.4.4 Integration, modularity & evolvability

Integration and modularity has often been linked to the ability of an structure to evolve, also known as evolvability (Alberch 1991; Wagner and Altenberg 1996; Kirschner and Gerhart 1998; Klingenberg 2005). Low levels of integration, especially when coupled with the presence of modules, can facilitate processes involved in morphological change, and enable localized changes to appear in response to such processes (e.g. natural selection) (Cheverud 1982, 1984, 1996; Kirschner and Gerhart 1998; Klingenberg 2005; Goswami 2006; Cardini and Elton 2008; Klingenberg 2009; Porto et al. 2009; Shirai and Marroig 2010; Klingenberg 2013). If the structure of the G matrix of Soay sheep is representative of sheep in general, we believe that the modularity found in the Soay sheep allowed for the great morphological and genetic variation we see among different sheep breeds (Rendo et al. 2004; Gizaw et al. 2007). Similarly to what Drake and Klingenberg (2010) found in dogs and wolves, it seems that artificial selection, when applied to modular structures produces a wide range of morphological diversity. Domestic species have been overlooked when it comes to understanding evolutionary processes, but we believe they are a great study model and encourage their use in order to ascertain the processes affecting natural species.

2.5 Literature cited

Albarrán-Lara, A. L., L. Mendoza-Cuenca, S. Valencia-Avalos, A. González-Rodríguez, and K. Oyama. 2010. Leaf fluctuating asymmetry increases with hybridization and introgression between Quercus magnolifolia and Quercus resinosa (Fagaceae) through


3. Chapter 3:

The role of genetic integration on the response to selection in the skull of Soay sheep

Abstract

Genetic integration is one of the main constraints in the evolution of morphological structures, because it deflects the response to selection towards the trait combinations that have the most genetic variation. This means that the outcome of selection might not reflect the traits that were being selected in the first place. I applied four different hypothetical selection regimes on the skull shape of Soay sheep, and then quantified the direct, total and correlated response to selection, and how they differ from each other. I found that most of the hypothetical selection regimes produced widespread morphological changes. Selection towards orbit frontality and reduced molars produced similar amounts of direct and correlated responses. While selection on smaller braincases produced twice as much correlated response as direct response. But selection towards a retraction of the nose produces little correlated response, indicating a certain independence, and thus, evolvability of the trait. That might be related to the appearance of novel narial anatomy in bovids. More studies on genetic integration are needed to understand the adaptive values of constraining some traits while allowing other to vary independently.

3.1 Introduction

Genetic integration occurs when genes involved in morphological variation have a joint effect on multiple traits, which can be represented as a network of pleiotropic relations among traits (Wagner 1989; Nadeau et al. 2003; Cheverud et al. 2004; Klingenberg 2008). These pleiotropic effects, or genetic integration, play a crucial role in the outcome of evolutionary processes, serving as a potential constraint on adaptive evolution (Wagner 1984, 1988; Eroukhmanoff 2009; Walsh and Blows
Patterns of genetic integration can direct selection towards the trait combinations that have the most genetic variation (Lande 1979; Falconer and Mackay 1996; Klingenberg and Leamy 2001; Gerhart and Kirschner 2007; Klingenberg et al. 2010), also known as genetic line of least resistance (Schluter 1996). This means that, in an organism with high genetic integration, selection for localized shape changes may produce widespread morphological changes (Klingenberg and Leamy 2001; Klingenberg et al. 2010; Martínez-Abadías et al. 2012). This has severe implications for the study of natural selection, since the outcome of selection seen in natural populations, might not reflect the trait that was actually being selected for (Eroukhmanoff 2009).

But in order to assess the effect of genetic integration on the outcome of selection, one must first obtain the genetic information linked to the morphological traits under study, represented by the genetic covariance matrix (G matrix) (Cheverud 1984; Roff 1996; Kirkpatrick 2009; Walsh and Blows 2009). But obtaining the G matrix can be quite difficult, since it requires either that the relationship between the individuals (pedigree) is known, or that the genotype of the individuals is known, both of which require laborious field work and/or laboratory experiments. To overcome this difficulty, some studies have used the phenotypic (P) matrix as a proxy for the G matrix, due to their similarities (Cheverud 1982, 1988). But the P matrix is not an exact representation of the genetic basis of phenotype, since development and plasticity also produce changes in the phenotype (West-Eberhard 1989; Agrawal 2001). In recent years, the easiest and most common means of obtaining G matrices has been through laboratory experiments (Klingenberg and Leamy 2001; Klingenberg et al. 2010), even though variance-covariance matrices in
lab-raised individuals might not reflect the real covariance structure found in the wild (Jamniczky and Hallgrímsson 2009).

In this study, I evaluate the effects of genetic integration on the outcome of selection acting on skull shape of a wild population of Soay sheep with known pedigree. This is accomplished by creating different hypothetical selection scenarios and applying into the breeder’s equation (Lush 1937). These hypothetical selection scenarios act on specific shape features which are then compared to the total response to selection (Klingenberg and Leamy 2001; Martínez-Abadías et al. 2009; Klingenberg et al. 2010; Martínez-Abadías et al. 2012). If these localized selection scenarios produce widespread changes throughout the skull, it indicates that genetic integration is deflecting the evolutionary response towards a phenotype that is not the one originally selected for.

3.2 Method

The Soay sheep population under study here, come from St. Kilda archipelago, in Scotland (Clutton-Brock and Pemberton 2004). Every sheep is tagged individually within a few days after birth, and a skin sample is taken for paternity analysis (Boyd et al. 1964). The population has been monitored since 1964 and the skulls of deceased animals were collected from 1976 until 2006. All the skulls collected in St. Kilda were deposited at the National Museum of Scotland, in Edinburgh.
3.2.1 Data collection and morphometric analyses

My sample consists of 1,508 skulls, all of which had information on age, sex and year of birth. A single observer (EMD) digitized 50 landmarks in three dimensions (Fig. 1) using a MicroScribe G2X. Each skull was digitized twice to correct for measurement errors. The raw landmark data was then imported into MorphoJ (Klingenberg 2011). First, I ran a generalized Procrustes fit to extract shape information (Dryden and Mardia 1998). Since the skulls have object symmetry, I only use the symmetric component of variation in my analyses (Klingenberg and McIntyre 1998; Klingenberg et al. 2002). I then ran a Principal Components Analysis (PCA) and used the resultant principal component scores in the quantitative genetic analysis.
Figure 1. Scheme showing landmark locations on A) dorsal view; B) lateral view; and C) ventral view.

3.2.2 Quantitative genetic analyses

Pedigree information is required in order to estimate the genetic basis of a given trait. For the Soay sheep, maternity was inferred by observation of maternal care behaviour, and paternity was inferred by analyses on single nucleotide polymorphisms (SNPs) (Kijas et al. 2009), using maximum-likelihood implemented
on CERVUS (Marshall et al. 1998; Robinson et al. 2006). For each lamb, a log-likelihood ratio score was calculated for each possible father, and criteria for assigning paternity to the male with the highest score with 95% and 80% confidence were obtained by simulation (Pemberton et al. 1999; Pemberton et al. 2004).

To obtain the genetic covariance matrix (G matrix), I performed a Restricted Maximum-Likelihood based on an animal model (Lynch and Walsh 1998; Kruuk 2004; Kruuk and Hill 2008; Wilson et al. 2010) using Wombat software (Meyer 2007). Due to computational limitations, I used 24 principal components (PC) scores from a pooled within-age PCA (the maximum number supported by Wombat) instead of the raw data to calculate the G matrix (Klingenberg and Leamy 2001; Klingenberg et al. 2010; Martínez-Abadías et al. 2012). Sex and horn morph were set as fixed effects and centroid size as a covariate, while individual and year of birth were set as random effects. The G matrix obtained through Wombat was then imported into MorphoJ (Klingenberg 2011), where it was converted into landmark coordinates for further analyses (Klingenberg et al. 2010).

3.2.3 Hypothetical selection

In order to evaluate how genetic integration influences the outcome of selection, I predicted the response to different hypothetical selection scenarios (Klingenberg and Leamy 2001; Martínez-Abadías et al. 2009; Klingenberg et al. 2010; Martínez-Abadías et al. 2012). To do that, I used the multivariate version of the breeder’s equation, \( \Delta \mu = G\beta = GP^{-1}s \), where \( \Delta \mu \) is the change in mean shape in response to selection, G is the genetic covariance matrix, P is the phenotypic
covariance matrix, $\beta$ is the selection gradient, and $s$ is the selection differential (Lande 1979).

I applied four different hypothetical selection regimes to assess if selection on specific traits causes localized shape changes, or an integrated response throughout the entire skull. Since I wanted to assess the effect of localized selection, I applied changes to just a few landmarks in one small region of the skull, as long as the shape changes were anatomically credible. All the traits put under hypothetical selection are involved in important physiological and ecological functions, but I am in no way claiming that the selected shape changes are related to increased fitness. The selection scenarios created here are completely hypothetical.

**Orbit frontality** - The selection gradient for this hypothesis acts towards more frontal-oriented orbits. Forward-facing orbits are often associated with stereopsis (Collins 1921; Simons 1962), which is the perception of depth and 3-dimensional sight. I applied changes to the following landmarks: 4/12; 5/13; 30/34; 32/36; 33/37. Landmarks on the exterior (posterior) part of the orbital rim and on the zygomatic arches are moving forward whereas landmarks on the anterior portion of the orbital rim are moving backwards, thus creating more forward-facing orbits (Fig. 2).

**Retraction of the nose** - The nasal cavity is not only involved in olfactory functions, but it is also involved in warming and moistening inspired air (Negus 1956; Lieberman 2011). This function can be of extreme importance in cold and dry areas (Noback et al. 2011). In this hypothesis, landmarks 9 and 10 are shifted posteriorly and dorsally while landmarks 17, 2/10, and 3/11 are shifted anteriorly (Fig. 3). The
nasal bone is not only being reduced, but also moving upwards, a shape change that is associated with larger turbinate bones.

*Reduced molars* - The evolution of teeth throughout the mammalian evolutionary history has sometimes been decoupled from morphological changes in the rest of the skull, suggesting an evolvability of the teeth (Hunter and Jernvall 1995; Jernvall et al. 1996; Jernvall 2000; Salazar-Ciudad and Jernvall 2002). The size of the molars is sometimes positively associated with grazing efficiency (Allden and Whittaker 1970; Okello et al. 2002; Sanson 2006). An anterior-posterior reduction of molars was achieved by shifting landmarks 46/49 posteriorly and landmarks 47/50 anteriorly (Fig. 4).

*Reduced braincase* - The size of the braincase is strongly correlated with endocranial volume and thus, with the size of the brain (Finarelli 2006). Brain size has been linked to cognition, sociality (Dunbar 1992; Barton 1996; Dunbar 1998; Dunbar and Bever 1998; Finarelli and Flynn 2009), and even habitat complexity among ungulate species (Shultz and Dunbar 2006). Also, reduced brain sizes have been described as one of the consequences of domestication (Kruska 1987, 2005; Wilkins et al. 2014). Selection towards smaller braincases required the displacement of more landmarks than the previous hypotheses. It involved a posterior shift in landmarks 6/14, 18/19, 31/35, 33/37; an anterior shift in landmarks 7/15, 8/16, 21, 28; and a dorsal shift in landmarks 22, 23/24, 25; and a ventral shift in landmarks 4/12 and 20 (Fig. 5).

Because the creation of hypothetical selection gradients requires a shift in the position of the landmarks, this may cause changes in centroid size, position and
orientation of the landmark configuration. And since size, position and orientation
do not reflect changes in shape, I had to ensure that they would not affect the
response to selection (Klingenberg et al. 2010; Martínez-Abadías et al. 2012). To do
that, I projected the vectors of landmarks shifts (selection gradients) into the shape
tangent space (Dryden and Mardia 1998). By doing that, the resulting selection
gradient contained small shifts in landmarks other than the ones being targeted
initially. But those shifts are much smaller than the shifts on the "target" landmarks
(Klingenberg and Leamy 2001; Klingenberg et al. 2010; Martínez-Abadías et al.
2012). The magnitude of the standardized selection gradient (change in relative
fitness per standard deviation of shape change that corresponds to the selection
gradient) (Lande and Arnold 1983) was set to 0.1 in all of the hypothetical selection
scenarios.

After I applied the hypothetical selection gradients into the breeder’s
equation, I obtained the response to selection Δμ which was then decomposed into
direct response, total response and correlated response. Direct response is the
component of shape change in the direction of the selection gradient, and the
correlated response is the component perpendicular to it, and therefore, changes in
the direction that were not selected for (Falconer and Mackay 1996). The total
response is the combination of both direct and correlated responses. The magnitudes
of the responses were computed in units of Procrustes distances (Klingenberg and
Leamy 2001; Klingenberg et al. 2010; Martínez-Abadías et al. 2012). I also
calculated the angle between the vectors of direct and total response as a measure of
the relative genetic constraints on the response to selection (Klingenberg and Leamy
3.3 Results

The selection gradient acting on orbit frontality returns a total response with subtle changes in the position of the orbits, and a considerable change in the nose area (Fig. 2). The nasal bone shifts upwards (in relation to the teeth) and posteriorly, as a result of correlated response. The magnitude of the correlated response is larger than that of the direct response (0.04 and 0.03, respectively), as a result, the direction of total response was deflected from the direction of the selection gradient by a 56° angle (Fig. 2).

Figure 2. Hypothetical selection: frontality of the orbits. Top left: lateral and dorsal views of the
starting shape, prior to the selection event. Top right: shape changes associated with the selection gradient with arrows showing the direction to which the landmarks were shifted. Lower half: decomposition of the total response (diagonal arrow and upper right skull surface) into the direct response in the direction of the selection gradient (vertical arrow and associated skull surface) and the correlated response (horizontal arrow and associated skull surface), and the angle between the vectors of direct and total response. The skull surfaces are showing the shape changes associated with each response, with arrows pointing the direction of shape change of the affected landmarks, for better visibility. The amount of shape change for the components of the response are indicated in units of Procrustes distance.

Meanwhile, the selection towards retraction of the nose resulted in higher magnitudes of direct response (0.085), with the correlated response (0.043) acting towards very subtle changes across the entire skull, which rendered it practically impossible to visualize (Fig. 3). The shape changes seen in the total response are seemingly identical to the ones seen in the direct response to selection, with an angle of only 27° between the two vectors (Fig. 3).

Selection to reduce molar size resulted in virtually the same magnitude for both direct (0.023) and correlated responses (0.027)(Fig. 4). But just like the hypothesis for retraction of the nose, the correlated response was spread throughout the whole skull, making it difficult to visualise, even when exaggerated by 2-fold. Because of that, the total response is apparently very similar to the direct response, even though there is a 50° angle separating them (Fig. 4).

Unlike all the previous hypothetical selection regimes, selection towards smaller braincases generated a correlated response almost twice as strong (0.05) as the direct response (0.027)(Fig. 5). The total response to selection shows a reduction of the braincase coupled with an elongation of the face. And because the correlated response is stronger, the elongation of the face is far more noticeable than the reduction of the braincase. The angle between direct and total response is 62° (Fig. 5).
Figure 3. Hypothetical selection: retraction of the nose. Top left: lateral and dorsal views of the starting shape, prior to the selection event. Top right: shape changes associated with the selection gradient with arrows showing the direction to which the landmarks were shifted. Lower half: decomposition of the total response into the direct response in the direction of the selection gradient and the correlated response, and the angle between the vectors of direct and total response. The skull surfaces are showing the shape changes associated with each response, with arrows pointing the direction of shape change of the affected landmarks, for better visibility. For further details, see legend to Figure 2.
Figure 4. Hypothetical selection: reduced molars. Top left: lateral and ventral views of the starting shape, prior to the selection event. Top right: shape changes associated with the selection gradient with arrows showing the direction to which the landmarks were shifted. Lower half: decomposition of the total response into the direct response in the direction of the selection gradient and the correlated response, and the angle between the vectors of direct and total response. The skull surfaces are showing the shape changes associated with each response, with arrows pointing the direction of shape change of the affected landmarks, for better visibility. For further details, see legend to Figure 2.
Figure 5. Hypothetical selection: reduced braincase. Top left: lateral and dorsal views of the starting shape, prior to the selection event. Top right: shape changes associated with the selection gradient with arrows showing the direction to which the landmarks were shifted. Lower half: decomposition of the total response into the direct response in the direction of the selection gradient and the correlated response, and the angle between the vectors of direct and total response. The skull surfaces are showing the shape changes associated with each response, with arrows pointing the direction of shape change of the affected landmarks, for better visibility. For further details, see legend to Figure 2.
3.4 Discussion

Genetic integration in the skull of Soay sheep affects the response to selection of most of the traits in our analysis. This means that, in most cases, hypothetical selection produces global responses to localized selection, with marked deflections of the evolutionary response from the direction of the selection gradient that reflects the effects of relative constraints. But the amount of constraints set by genetic integration varied across different parts of the skull. Selection in both the orbits and teeth created similar amounts of direct and correlated response. Meanwhile, selection in the braincase and in the nose generated opposite results. Selection towards smaller braincases generated a correlated response that was stronger than the direct response, whereas selection for retraction of the nose had almost no visible correlated response. We can thus conclude that genetic integration is not pervasive in the skull of Soay sheep, but it appears to constrain some morphological changes in the skull more (reduction of braincase) than others (retraction of the nose).

Because of the constraints imposed by genetic integration, the total response to selection for reduced braincases generated a strong correlated response consisting of an elongation of the face. The pattern "small braincase-elongated face/large braincase-shortened face" is recurrent in allometry and ontogeny studies (Zelditch et al. 1992; Abdala et al. 2001; Flores et al. 2003; Weston 2003; Berge and Penin 2004; Giannini et al. 2004; Cardini and O’Higgins 2005; Marroig and Cheverud 2005; Cardini and R. W. Thorington 2006; Flores et al. 2006; Giannini et al. 2010). In general, small and neonate mammals have comparatively larger braincases and shorter snouts, whereas older and larger animals have comparatively smaller braincases and longer snouts (Langenbach and van Eijden 2001). This pattern could be an
adaptation to suckling, with the face features developing post-natally (Cheverud 1982), as the feeding requirements of the sucklings change (Langenbach and van Eijden 2001; Cardini and O’Higgins 2005), while most of the neurocranium is developed in utero (Marcucio et al. 2011; Cardini and Polly 2013). Highly constrained braincases could be a mechanism to protect the area against disadvantageous shape changes (Wagner and Altenberg 1996), since even the subtlest alterations in the brain could have fatal consequences for the individual. Previous studies in primates, including humans, also found that the changes applied to the braincase lead to correlated responses in the face (Cheverud 1982; Cheverud et al. 1992; Kohn et al. 1993).

Unlike the other hypothetical selection regimes, the localized change in the nose did not produce a strong correlated response. The total response to selection towards a retraction of the nasal bone was, in fact, very similar to the direction of the selection gradient. This indicates a certain independence of the narial morphology from the rest of the skull. Being independent from other traits, increases the variability of such a trait, which is the propensity to vary (Wagner and Altenberg 1996). The propensity to vary, in turn, increases the potential to evolve or adapt (Wagner and Altenberg 1996; Kirschner and Gerhart 1998; Klingenberg 2005; Pigliucci 2008). This potential to evolve that we have seen in the nose, might explain the array of different narial anatomies found in some ungulates (Witmer et al. 1999; Clifford and Witmer 2004). One extreme example of novel narial anatomy in bovidae is the proboscis on Saiga antelopes (Saiga tatarica) (Clifford and Witmer 2004). The skulls of Saiga antelopes show an extreme retraction of the nasal bones as one of the morphological changes associated with the proboscis (Fig. 1 in Clifford
and Witmer 2004). As in many other proboscis-bearing mammals, the nasal bones are retracted in Saiga antelopes in order to accommodate an enlarged nasal vestibule and septum cavernous mass, which, among other changes in narial anatomy, are adaptations to dusty habitats (Clifford and Witmer 2004). These adaptations work by regulating air flow, so that the inspired particulates can be collected, so that the air destined for the lungs is clean (Clifford and Witmer 2004).

Similar studies in human skulls and mouse mandibles found pervasive genetic integration that consistently deflected the total response from the direction of the selection gradient (Klingenberg and Leamy 2001; Martínez-Abadías et al. 2009; Martínez-Abadías et al. 2012). Our results differ from those found in humans and mice since not all localized shape changes produce strong correlated responses. Studies analyzing the response to hypothetical selection gradients are still scarce in the literature, especially due to the aforementioned efforts imposed by the G matrix (Klingenberg and Leamy 2001; Martínez-Abadías et al. 2009; Klingenberg et al. 2010; Martínez-Abadías et al. 2012). But understanding the underlying processes of phenotypic change is crucial for selection studies. First, it allows us to point out the actual traits that are under selection, instead of assuming that all the shape changes portrayed in the response to selection were selected for. And second, if we compare the differences in evolvability between traits in different taxa, we might be able to understand the adaptive value of certain traits being more constrained, or more evolvable than others.
3.5 Literature Cited


Collins, E. T. 1921. Changes in the visual organs associated with the adoption of arboreal life and with the assumption of erect posture. Transactions of the Ophthalmological Society of the United Kingdom 41:10-90.


4. Chapter 4:

Estimating selection of skull shape using geometric morphometrics

Abstract

Natural selection plays a central role in evolutionary theory, and therefore it has drawn the interest of many researchers since its discovery. But most studies focus on univariate measures such as size, but rarely on the shape of traits. Combining multivariate methods for quantifying shape and selection estimation methods allows us to not only quantity, but also visualize the selected shape. In this study, I combined analysis of selection on multivariate traits with geometric morphometrics in order to estimate selection on skull shape in a Soay sheep population. I found significant linear selection in males ($\beta_1 = 0.83, p = 0.031$) (Fig. 2) and females ($\beta_1 = 0.67, p = 0.0077$). I found that linear selection is favouring broader nasal bones in males, but in females, selection is acting towards narrower nasal bones and smaller braincases. Non-linear selection is absent on males and females. Selection acts on different aspects of shape in males and females, possibly due to sexual selection, since males compete for female access. My method is more thorough in the estimation of selection gradients when compared to previous methods using geometric morphometrics.

4.1 Introduction

Natural selection is one of the concepts in which the whole evolutionary theory is based upon. And since the emergence of evolutionary theory, there have been many efforts towards quantifying natural selection in a multitude of organisms (Endler 1986; Kingsolver et al. 2001; Pagel 2009). But estimating selection is hard, since it requires a measure of fitness for each individual, along with data on the trait of interest (Lande and Arnold 1983; Phillips and Arnold 1989; MacColl 2011). And this is why long-term studies using marked individuals provide great opportunities
for monitoring selection events and the evolutionary process as a whole (Milner et al. 1999a). Long-term studies of populations are not only coupled with large sample sizes and variety of different information regarding the population, but it is also more realistic, since it takes into account all environmental/ecological/genetic variability that are experienced by populations in natural circumstances.

There are several examples of natural selection events found in long-term studies of populations (for a review see (Hendry and Kinnison 1999)), some of the most iconic studies of this kind are on Darwin’s finches from Galapagos (Grant et al. 1976; Grant and Grant 1989, 1993; Grant and Grant 2002; Grant et al. 2004; Grant and Grant 2006); various populations of Guppies (Endler 1980; Magurran et al. 1992; Seghers and Magurran 1995; Reznick et al. 1997); and American house sparrows (Johnston and Selander 1964). Another excellent example of long-term studies of populations uncovering natural selection events are the Soay sheep of St. Kilda (Clutton-Brock and Pemberton 2004b; Milner et al. 2004). This particular population has been monitored for over 50 years now, with an abundance of morphological and ecological data being sampled every year (Clutton-Brock and Pemberton 2004b), thus making it ideal for selection estimation studies. So far, selection studies on the Soay sheep include traits like horn phenotype (Moorcroft et al. 1996; Clutton-Brock et al. 1997; Robinson et al. 2006, 2008), coat colour (Moorcroft et al. 1996; Gratten et al. 2008), and a few morphometric traits such as hindleg length, weight, and incisor breadth arcade (Illius et al. 1995; Milner et al. 1999a).

But as in most selection studies, the Soay sheep research focused on estimating selection on univariate morphological variables, such as size and weight
(Kingsolver et al. 2001; Pagel 2009). Selection estimates of a trait’s shape, on the other hand, are comparatively scarce in the literature. And even though the methods for estimating selection on shape, or multivariate selection, have been around for a while (Lande and Arnold 1983; Phillips and Arnold 1989), it is mostly used to correct the analyses for selection in correlated characters (Kingsolver et al. 2001).

The recent increase in accessible programs and techniques for analysing variation in complex shapes, has popularized methods like geometric morphometrics (Rohlf and Marcus 1993; Zelditch et al. 2012). The past two decades has seen a few successful attempts to combine multivariate selection methods with geometric morphometrics (Herrera 1993; Gómez et al. 2006; Gómez et al. 2008; Benitez-Vieyra et al. 2009; Nattero et al. 2010) but the possibilities within combining geometric morphometrics and selection estimates are still utterly underexplored. Using geometric morphometrics to estimate selection is extremely useful since the shape of traits is in many ways, more informative than just the size. Geometric morphometrics methods also allow us to visualize the selected shape in addition to quantifying the strength of selection. Although so far, the methods used to estimate selection using geometric morphometrics use only a small fraction of the variation captured by geometric morphometrics methods, and they do not allow for the visualization of the selected shape changes, thus not exploring the full potential of the method.

In this paper, I will introduce a new method for estimating selection using geometric morphometrics methods. I will test this method by estimating selection of skull shape in a wild population of Soay sheep (Boyd et al. 1964). This Soay sheep population is ideal for this sort of analysis because they have been studied for over 50
years, thus providing a large amount of data for each individual, such as breeding success, sex, and age (Boyd et al. 1964; Clutton-Brock and Pemberton 2004a). With this new method, I was able to utilize the maximum amount of the available variation as permitted by the analyses. And I was also able to visualize and display the aspects of shape change being selected, both of which were not possible with the previously developed methods. This study is innovative in both the method employed, and in the quality of the data, since large samples and fitness measures of wild populations of mammals are quite rare.

4.2 Methods

4.2.1 Study population

The Soay sheep population of St. Kilda has been monitored since 1964 (Clutton-Brock and Pemberton 2004b). Every individual was tagged a few days after birth, and a skin sample was taken for paternity analysis. The dates of birth and death were recorded thus allowing us to know the age of every animal in our sample (Boyd et al. 1964).

The skulls of deceased animals were collected from 1976 until 2006 and all the skulls collected in St. Kilda were then brought to and deposited at the National Museums of Scotland, in Edinburgh. The population of Soay sheep from St. Kilda undergo severe population crashes, where 50-60% of the animals die (Grubb 1974a; Clutton-Brock et al. 1991; Grenfell et al. 1992; Clutton-Brock et al. 2004a). And although the population crashes are the most intense source of selective pressure, there are also other more subtle sources of selection, such as parasites, food availability, and climate (Milner et al. 1999b; Milner et al. 2004; Wilson et al. 2004).
So even though the sheep are not predated in the island, they are under selective pressure (Illius et al. 1995; Moorcroft et al. 1996; Clutton-Brock et al. 1997; Milner et al. 1999a).

4.2.2 Fitness measure

The fitness of an individual is the contribution of genes that it makes to the next generation (Falconer and Mackay 1996). Fitness components can be described as mating ability, fertilizing ability, fertility, fecundity, and, or survivorship (Endler 1986; Falconer and Mackay 1996). In our analyses, I used Lifetime Breeding Success (LBS) as a fitness measure in the selection analyses, because it accounts for all of the aforementioned fitness components. And because of the discrepancy of breeding success between the sexes, I performed separate analyses for males and females. Maternity was inferred by observation of maternal care behaviour, and paternity was inferred by analyses on single nucleotide polymorphisms (SNPs), using maximum-likelihood implemented on CERVUS (Marshall et al. 1998).

Because older animals have obviously had more time to produce offspring, the multivariate regressions for selection gradient estimation were pooled-within age groups. The age groups were classified according to morphological differences as: yearlings (1-1.9 years), young adults (2-2.9 years), adults (3-7 years), and old animals (8 or more years). Adults with 3 to 7 years were grouped into the same category due to morphological similarity, since after reaching sexual maturity at 3 years, the sheep stop growing (Clutton-Brock et al. 2004b), thus maintaining a similar morphology. Animals older than 8 years, although morphologically similar to adults, have a much
higher breeding success (Table 1). I classified these older animals as a different age
group, as an effort to capture subtle morphological characteristics that might be
linked to longevity and high breeding success rates. From the data on breeding
success from each sheep, relative fitness was calculated by dividing the individual’s
LBS by the average LBS of all individuals from the same sex and age group (Table 1).

Table 1. Average values of Life-time Breeding Success (LBS) for each sex and each age group.

<table>
<thead>
<tr>
<th>Sex/Age group</th>
<th>Yearling</th>
<th>Young Adult</th>
<th>Adult</th>
<th>Old</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>0.2</td>
<td>0.95</td>
<td>6.04</td>
<td>17.1</td>
</tr>
<tr>
<td>Female</td>
<td>0.04</td>
<td>0.63</td>
<td>2.36</td>
<td>6.6</td>
</tr>
</tbody>
</table>

4.2.3 Shape data

My sample consists of 1,293 skulls, of which 616 are males and 677 are
females. A single observer (EMD) digitized 50 landmarks in three dimensions (Fig.
1) using a MicroScribe G2X. Each skull was digitized twice to correct for
measurement errors. The raw coordinate data was then imported into MorphiJ
(Klingenberg 2011), where all subsequent analysis took place. First, I ran a
generalized Procrustes fit to extract shape information. Since the skulls have object
symmetry, I only use the symmetric component of variation in our analyses
(Klingenberg et al. 2002). I then ran a Procrustes ANOVA to correct the data for
measurement errors, followed by a Principal Component Analysis (PCA). I used the principal components scores from the PCA in the selection analysis.

![Figure 1. Scheme showing landmark locations on A) dorsal view; B) lateral view; and C) ventral view.](image)

I use principal components (PCs) instead of the original data because of the high-dimensionality of skull shape. I could compromise the robustness of our analysis by not having a large enough sample size relative to the number of dimensions. But I still want to include the as much data on skull shape as possible. Therefore, in my selection analysis, I included the same amount of PCs as the square number of my sample size (35 PCs), so that the robustness of the analysis is not compromised.

4.2.4 Estimation of Selection Gradients
Selection can be directional, disruptive, or stabilizing, and they differ mainly in the way through which they alter the populations phenotype. Directional selection acts on the mean phenotype whereas disruptive and stabilizing selection act on all other moments of the phenotypic distribution (Lande and Arnold 1983; Endler 1986). One way of estimating the different modes of selection is through selection gradients. Selection gradients are the partial regression coefficients derived from multiple regressions of fitness on traits (Lande and Arnold 1983; Phillips and Arnold 1989; Klingenberg and Monteiro 2005). Unlike selection differentials, they estimate only the direct effects of selection in one trait, thus correcting for selection on correlated traits (Lande and Arnold 1983; Phillips and Arnold 1989). In order to estimate selection gradients for directional and stabilizing/disruptive selection (quadratic selection), I used the multiple regression techniques from Philips and Arnold (1989) as implemented in MorphoJ (Klingenberg 2011):

\[ w = \alpha + z^T \beta + z^T \gamma z \]

where \( z \) is the vector of the traits on which selection is being measured, in this case, shape changes in the form of Principal Components scores (PCs), \( w \) is relative fitness, \( \beta \) is the vector of directional selection gradients, and \( \gamma \) is the matrix of quadratic selection gradients and superscript \( T \) stands for matrix transposition. The directional selection gradient \( \beta \) was computed from the standardized partial regression coefficients of a linear regression of relative fitness on the first 35 Principal Components (to improve the robustness of the analysis as previously explained), whereas the quadratic selection gradient \( \gamma \) was calculated from the second-order standardized coefficient in a quadratic regression. Since the selection gradients \( \beta \) and
$\mathbf{y}$ are vectors, I standardized the gradients by their standard errors $\sigma$. I use this standardized version of the gradients, $\mathbf{\beta}_\sigma$ and $\mathbf{y}_\sigma$, to represent the magnitude of linear and nonlinear selection, respectively.

But selection gradients, although adequate to represent the magnitude of selection, should not be visualized as shape changes being selected, as they are in units of the inverse of shape (Klingenberg and Monteiro 2005). To be able to visualize the shape changes associated with linear and nonlinear selection, I used selection differentials as proposed by Klingenberg and Leamy (2001) and Klingenberg and Monteiro (2005). In order to test the statistical significance of the selection gradient estimates I ran permutation tests (1,000 randomization rounds).

4.2.5 Linear Selection

The linear relationship between fitness and a set of traits given by directional selection, has led to this form of selection being called linear selection (Simpson 1953; Spiess 1977; Lande and Arnold 1983). The directional selection gradient ($\mathbf{\beta}$) is the average slope of the straight line that best describes the dependence of fitness on a character $z$, after removing residual effects of other characters (Lande and Arnold 1983; Phillips and Arnold 1989; Blows and Brooks 2003).

4.2.6 Nonlinear Selection

The quadratic selection gradient ($\mathbf{y}$) represents in the curvature of the quadratic regression, showing the nonlinear relationship between fitness and trait ($z$), and thus, nonlinear selection. Stabilizing and disruptive selection are both classified as nonlinear selection, as it describes selection that causes a change in the second or higher moments of phenotypic distribution beyond those caused by directional

4.3 Results

The permutation test found significant linear selection in males ($\beta_1 = 0.83$, $p = 0.031$) (Fig. 2) and females ($\beta_1 = 0.67$, $p = 0.0077$) (Fig. 2), but no significant nonlinear selection in either sex ($p > 0.5$). In males, selection is acting in favour of wider molar teeth, in relation to the pre-molar teeth, a broader nasal bone (in relation to the overall width of the skull), a more square-shaped palate, and more forward facing orbits (Fig. 3). In females, selection favours a flattening of the posterior superior portion of the braincase, narrower nasal bones (in relation to the overall width of the skull), and higher maxillary bones in relation the total height of the skull, especially where the nasal bone meets the incisive bone (Fig. 4).
Figure 2. Scatterplots from regressions between relative fitness and symmetric component (shape) scores for males (♂) and females (♀).

Figure 3. Diagram showing the average skull shapes (average shape) and the skull shape associated with the highest fitness (selected shape), in dorsal, lateral and ventral views for males.
4.4 Discussion

When compared to selection estimates on morphometric traits in multiple natural populations (median $|\beta| = 0.17$) (Kingsolver et al. 2001), selection in skull shape in Soay sheep is not only significant, but also strong. There has been many studies estimating selection on an array of Soay sheep traits, including body size, body weight, incisor breadth arcade, horn morph, coat colour, etc (Illius et al. 1995; Moorcroft et al. 1996; Clutton-Brock et al. 1997; Coltman et al. 1999; Milner et al. 1999a; Robinson et al. 2006). The focus of most of these studies is on analysing the link between strength of selection and ecological events, such as population crashes (Illius et al. 1995; Moorcroft et al. 1996; Clutton-Brock et al. 1997; Milner et al.
1999a). When we compare standardized selection gradients ($\beta_\sigma$) between skull shape and other traits, we find that selection in skull shape is comparatively strong. In males, both incisor breadth and horn size have $\beta_\sigma = 0.39$ (Milner et al. 1999a; Robinson et al. 2006), whereas hindleg length has $\beta_\sigma = 0.99$ (Robinson et al. 2006). In females, there is no selection on incisor breadth arcade (Milner et al. 1999a), horn size has $\beta_\sigma = -0.009$, and hindleg length has $\beta_\sigma = 1.16$ (Robinson et al. 2006).

We found that selection favours particular skull shapes in males and females of Soay sheep. It could be that the selected shape confers an anatomic advantage related to physiology, mechanical processes, or anything that might improve individual fitness (Endler 1980). In males, there is another factor contributing to increased fitness: female choice and competition for female access. Soay sheep are promiscuous, but males can guard females for several hours, until they have mated multiple times (Grubb 1974b; Clutton-Brock et al. 2004b). During this period, the guarding males might be challenged by other males, and they usually engage in head butting and chasing behaviour (Grubb 1974b; Clutton-Brock et al. 2004b). The skull shape being selected in males, might be related to increased ability to defeat other males in combats, which would also explain why it differs from the skull shape selected in females. And perhaps, skull shape is not even the main target of selection, but only a by-product of selection on a different trait that is genetically related to the skull (Clutton-Brock et al. 1997). For example, if there is positive selection for testosterone levels in males, skull shape would be affected, and combat advantages would be acquired, such as an increase in dominance (Ruiz-de-la-Torre and Manteca 1999) and physical stamina (Malo et al. 2009).
The differences on selected skull shape between the sexes might be related with mating competition, which is restricted to males. But these differences in skull shape, could also be related to differences in grazing efficiency and selective grazing. Because males have larger body sizes and thus higher metabolic requirements, grazing short swards impose greater selection pressure on males to have flatter arcades (Illius and Gordon 1987; Illius et al. 1995). Illius et al. (1995) also suggested the existence of resource partitioning between the sexes, with males often grazing on the short but nutritious halophytic sward, but females seldom doing the same, although this has never been confirmed. There are many factors that might explain differences between selection in males and females, and it is very difficult to identify which specific factors contribute to selection. Selection towards maintaining or increasing sexual dimorphism has been found in several organisms (Endler 1980; Forsman 1995; Preziosi and Fairbairn 2000; Rice and Chippindale 2001; Schulte-Hostedde et al. 2004). Our method is efficient in giving the strength and direction of selection, but the factors driving selection, can only be speculated.

So far, all studies using geometric morphometrics data to estimate selection on shape has used only the first few (2-3) relative warps (RW) scores as shape variables (Herrera 1993; Gómez et al. 2006; Benítez-Vieyra et al. 2009; Gómez et al. 2009; Nattero et al. 2010). But the first relative warps, like the first principal components, represent the most varying aspects of shape only, thus omitting other, less variable traits (Bookstein 1989; Marcus et al. 1993). The problem with using only two or three RW, is that less varying traits are disregarded, even though they might be just the ones being selected. Another difference between our method and previous methods is ability to visualise of the selected shape changes. The previous
methods do not allow for visualization of the selected shape changes. Each RW was treated as an independent variable, so only the shape changes associated with that particular RW could be visualised. Our method analyses all the PCs together, thus combining the shape changes associated with all the PCs that contribute to an increased fitness. In that way, we can visualise all aspects of shape changes that are being selected, instead of seeing only the shape changes associated with one PC/RW. In this way, we are able to assess exactly which shape changes are most associated with high fitness, as well as visualise the selected shape.

Our analysis found no significant nonlinear selection on skull shape. Estimations of nonlinear selection are seldom significant (Endler 1986; Kingsolver et al. 2001), although several authors have argued that they might have been grossly underestimated by researchers (Kingsolver et al. 2001; Blows and Brooks 2003; Stinchcombe et al. 2008). There are two procedures that prevent nonlinear selection from being underestimated. One, is to apply canonical analysis prior to quadratic regression analysis (Phillips and Arnold 1989; Blows and Brooks 2003), and the other is to adjust the quadratic regression coefficients, multiplying it by two, to obtain accurate estimates of nonlinear selection (Stinchcombe et al. 2008). Even though our method applies all the methodological procedures suggested to effectively estimate nonlinear selection gradients (Blows and Brooks 2003; Stinchcombe et al. 2008), there is still a debate around the appropriate methods to estimate nonlinear selection (Morrissey 2014). Therefore, we cannot confidently state that nonlinear selection does not occur in skull shape of Soay sheep.

In this paper we have presented a new method for estimating multivariate selection using geometric morphometrics. We tested our method in one of the most
well-studied populations of wild mammals, thus providing us with reliable fitness
data and a large sample. Our results show a stronger linear selection in skull shape,
and most importantly, that the aspects of the selected shapes differ between the sexes.
For the first time, selection on skull shape has been estimated, and we hope to
encourage other researchers to apply geometric morphometrics on selection studies,
so that we can have a better understanding on the causes and consequences of
selection in more complex traits, such as skull shape.

4.5 Literature Cited
Benitez-Vieyra, S., A. M. Medina, and A. A. Cocucci. 2009. Variable selection patterns on
the labelllum shape of Geoblasta pennicillata, a sexually deceptive orchid. Journal of
Evolutionary Biology 22:2354-2362.
Naturalist 162:815-820.
Bookstein, F. L. 1989. Principal warps: thin-plate splines and the decomposition of
deformations. IEEE Transactions on Pattern Analysis and Machine Intelligence
11:567-585.
Pemberton, eds. Soay sheep: Dynamics and selection in an island population.
Cambridge University Press, Cambridge, UK.
T. H. Clutton-Brock, and J. M. Pemberton, eds. Soay Sheep: dynamics and
Pemberton, eds. Soay sheep: dynamics and selection in an island population
Cambridge University Press, Cambridge, UK.
mediated selection against inbred Soay sheep in a free-living, island population.
Evolution 53:1259-1267.


110


Evidence of climate-change-induced heterochrony in the skull of Soay sheep

Abstract

Heterochronic processes are known as any changes in shape caused by changes in developmental timing. Research on the population of Soay sheep on the island of St. Kilda, Scotland has shown that shorter and less severe winters caused by global warming are allowing slow growing, smaller animals to survive winter, causing both an increase in population size and a decrease in average body size. Several studies have already shown that size and shape are extremely correlated in mammals. In this study, I am looking for evidence of heterochronic processes affecting the skull shape of the St. Kildan sheep. To achieve that, I ran regressions between size and skull shape (allometry) and year of birth and skull shape, followed by an angular comparison between the vector of the two regressions. I found significant effect of allometry and year of birth in skull shape. I also found that shape changes throughout the years are related to the decreasing size of the sheep. This means that at least a portion of the sheep are undergoing heterochronic processes, since their skull shape has paedomorphic characteristics. The fact that climate change might be associated with this phenomenon in the Soay sheep sheds new light into the consequences of global warming, in a way that has not been explored before.

5.1 Introduction

Long-term studies have shown us that a population can undergo morphological changes after only a few generations (Hendry and Kinnison 1999; Schluter 2000; Kinnison and Hendry 2001; Grant and Grant 2002; Garant et al. 2004; Gilchrist et al. 2004; Hargeby et al. 2004; Hendry et al. 2006; Phillips et al. 2006; Strauss et al. 2006; Carroll et al. 2007). Although most of these studies take
place in laboratory settings, a few of them regard wild populations (for a review see Carroll et al. 2007). The forces driving phenotypic changes vary, but it usually involves natural selection caused by changes in the environment (Carroll et al. 2007). With animals either being introduced in a new habitat (Losos et al. 1997; Reznick et al. 1997; Hendry et al. 2000; Strauss et al. 2006), or the habitat itself changing due to anthropogenic effects (Palumbi 2001; Hendry et al. 2006; Strauss et al. 2006), or climatic events (Clutton-Brock and Albon 1983; Grant and Grant 2002; Stanseth et al. 2002; Babin-Fenske et al. 2008). The latter, in particular, have been linked to morphological abnormalities (Telemenco et al. 2013) and changes in body size (Clutton-Brock and Albon 1983; Grant and Grant 2002; Ozgul et al. 2009; Ozgul et al. 2010; Caruso et al. 2014).

In the past decades, several studies have shown how climate change can affect the phenology of animals (Beebee 1995; Crick et al. 1997; Bradley et al. 1999; Crick and Sparks 1999; Dunn and Winkler 1999; Roy and Sparks 2000; Gibbs and Breisch 2001; Beebee et al. 2002; Crick 2004; Both et al. 2005; Marra et al. 2005; Sparks et al. 2005; Menzel et al. 2006) and plants (Bradley et al. 1999; Menzel and Fabian 1999; Menzel 2000; Menzel et al. 2001; Ahas et al. 2002; Fitter and Fitter 2002; Luo et al. 2007; Franks and Weis 2008). Records of the effect of climate change on morphology are not as recurrent, but they have been found in Darwin’s finches (Grant and Grant 2002), stream beetles (Babin-Fenske et al. 2008), salamanders (Caruso et al. 2014).

A very good example of changes in climate affecting the mean body size of a population are the Soay sheep of St. Kilda (Ozgul et al. 2009). This particular population of sheep has been monitored since 1964, and since 1986, there has been
regular measuring and weighing of almost every individual (Clutton-Brock and Pemberton 2004a). Data on body weight shows a continuous decline between the years 1986 and 2006 provoked by reduced growth rates in lambs. This reduced growth rates, in turn, are correlated with a decrease in the North Atlantic-Oscillation (NAO) index. This means that mild, rainy winters have contributed to a reduction in the average body weight of the population (Ozgul et al. 2009). So far, there are only records of changes in body weight and body size in the Soay sheep, even though there is a known correlation between size and shape (allometry) in mammalian anatomy, especially in the skull (Gould 1966, 1971; Klingenberg 1998; Goswami and Prochel 2007; Lieberman 2011).

One of the best studied and most widespread mechanisms for producing variation in size and shape is heterochrony. Heterochrony encompasses all changes in developmental rates and/or timing that generate a change in shape, but size and shape are rarely isolated from each other, and heterochronic processes often cause alterations in size (Gould 1977; Klingenberg 1998). Heterochronic changes can come by as changes in growth rates, changes in time of cessation of somatic growth, and changes in time of developmental onset. In most cases, heterochronic processes are observed between different taxa, where the descendent species is paedomorphic or peramorphic in relation to the ancestral species. Paedomorphosis is the retention of ancestral juvenile characteristics in adult descendants, whereas the opposite process is peramorphosis (Alberch et al. 1979).

Studies linking heterochrony with climatic events, on the other hand, are quite scarce. Heterochronic processes being induced by the modern ice age during the Late Neogene has been hypothesized in hominins (Vrba 1996). Also, in an
extinct species of salamander, *Brachycormus noachicus*, resulting from the Oligocene cooling (Roček 1995, 1996). But in both cases, the heterochronic events occur in evolutionary time-scales, in a phylogenetic context.

This study will test if heterochronic processes are behind the morphological changes found in the wild population of Soay sheep (*Ovis aries*) from St. Kilda in the course of 30 years (Clutton-Brock and Pemberton 2004b), which show a decrease in body size during the same period (Ozgul et al. 2009). This particular Soay sheep population is ideal for this kind of study, not only due to the availability of long-term data, but also because there is already an indication of morphological changes caused by an increase in temperatures (Ozgul et al. 2009). Through my analysis, I will be able to assess if and how climate change affects skull morphology in the Soay sheep. Long-term studies of wild mammals are quite scarce, and I believe this particular study will improve our understanding on the consequences of global warming on the phenotype of herbivores.

### 5.2 Methods

The Soay sheep population of St. Kilda has been monitored since 1964 and the skulls of deceased animals were collected from 1976 until 2006 (Clutton-Brock and Pemberton 2004). All the skulls collected in St. Kilda were then brought to and deposited at the National Museums of Scotland, in Edinburgh. Every individual was tagged a few days after birth, and its birth date recorded, as well as the date of death (Boyd *et al.* 1964), thus allowing me to know the age of every animal in our sample. Shape data were obtained from the skulls via geometric morphometrics. In order to
do that, one of the authors (EMD) digitised 50 landmarks with a Microscribe G2X
(Fig. 1). Each skull was digitized twice to correct for measurement error, and all the
specimens analysed had information regarding sex, age and year of birth. The total
sample size was 1,528 skulls.

The raw landmark coordinates were imported into MorphoJ (Klingenberg
2011), where all the subsequent analyses were performed. Because the skulls have
object symmetry, I used only the symmetric component of variation in the further
analyses (Klingenberg et al. 2002), and from now on I will refer to it as 'shape'.
Figure 1. Scheme showing landmark locations on A) dorsal view; B) lateral view; and C) ventral view.
To test for changes in size during the study period, I ran an Analysis of Variance (ANOVA) between centroid size and year of birth. To establish if there were any shape changes at all during the 30 years the skulls were collected, I ran a pooled within age-group regression between shape and the year of birth of each specimen (Drake and Klingenberg 2008; Weisensee and Jantz 2011).

To test the influence of size on shape (ontogenetic allometry) I ran a regression between shape and log centroid size. I used centroid size as a proxy for body size because it is the most accurate and reliable measure of size available for this data. Furthermore, centroid size has been used in many previous geometric morphometrics studies as a proxy for body size (e.g. Frost et al. 2003; Merino et al. 2005; Figueirido et al. 2013; Klingenberg and Marugán-Lobón 2013). A permutation test was carried out to test statistical significance of the regressions. Since sexual dimorphism and ontogeny-related changes are found in the skull shape of Soay sheep, both regressions were pooled-within sex, and the regression of shape against years was also pooled-within age groups (yearlings, young adults and adults). The age groups were assigned based on preliminary analysis (Principal Component Analysis) that revealed which ages (in years) are morphologically similar to each other. Yearlings are animals older than one year, but younger than 2 years, young adults are animals between two to three years and adults are animals with three years or more.

To assess whether the changes in skull shape were related to the changes in size I compared the angles between the vectors of the two regressions. The angle between the vector can be calculated as $\alpha = \arccos \left( \frac{\mathbf{a}^T \mathbf{b}}{\| \mathbf{a} \| \| \mathbf{b} \|} \right)^{0.5}$, where $\mathbf{a}$ and $\mathbf{b}$ are vectors, "arccos" is arc-cosine, and "$^T$" stands for matrix transposition (Klingenberg and Marugán-Lobón 2013). To estimate statistical significance from
the comparison of vectors, I used a closed-form formula used by Li (2011) to compute the area of a hypersphere cap. For more details on the vector comparison methods see Klingenberg and Marugán-Lobón (2013).

5.3 Results

The ANOVA between centroid size and year of birth showed a significant (P < 0.001) relationship between the two variables, with an overall decrease in average centroid size throughout the years (Fig. 2). There is also more variation on centroid sizes after the year 1985. The regression between shape and year of birth is also significant (P < 0.0001), showing that the Soay sheep have changed during the course of these 30 years, even though year of birth accounts for only 0.82% of shape variation. The morphological changes associated with year of birth include (from earlier to later times), narrower nasal bones, a more round-shaped palate as well as an elongation of the anterior region of the palate. One of the most conspicuous differences is in the posterior edge of the maxilla, just after the last molar. In earlier years, this part of the maxilla is curved, forming a wide angle with the zygomatic process, whereas later on, the same part forms a more accentuated angle with the zygomatic process (Fig. 3). Because the regressions were pooled-within age groups, the "Birth Year" axis contains numbers ranging from -20 to 15, representing earlier and later years, respectively.
Figure 2. Graph from the analysis of variance (ANOVA), showing the relationship between centroid size and year of birth. The black dots mark the average centroid size, and the horizontal bars represent the standard errors. The red circles around the black dots mark the years in which the population crashed.
Figure 3. At the top, scatterplot of regression between year of birth and skull shape with the amount of variation explained by year of birth (0.82%), and the significance of the test ($P<0.0001$). At the bottom, figures showing the skull shape associated with the earliest (right) and latest (left) years. To improve visualization, the changes were exaggerated five-fold.
Ontogenetic allometry is significantly ($P < 0.0001$) present in the Soay sheep skull, and accounts for 31% of shape variation. Smaller animals have narrower and shorter nasal bones in comparison with the braincase, larger orbits in comparison with the size of the skull, wider frontal bones (in relation to nasal bones), broader zygomatic arches, and more prominent occipital condyles (Fig. 4).

The angular comparison between the shape vs. size regression and the shape vs. year of birth regression shows significant similarity among its vectors ($P < 0.00001$) with a $60^\circ$ degree difference between them. In order to correctly interpret the degrees of difference between vectors, we must take into account the dimensionality of the data. The number of dimensions can be calculated by the expression $3k + 2l - 4$, where $k$ is the number of pairs of landmarks (landmarks that can be found in both sides of the skull), and $l$ is the number of unpaired landmarks (Klingenberg et al. 2002; Klingenberg 2015). The subtraction of four degrees of freedom is due to constraints imposed by the Procrustes fit such as scaling, position (all configurations must align with the median axis) and rotation (Klingenberg et al. 2002; Klingenberg 2015)
Figure 4. At the top, scatterplot of regression between log centroid size and skull shape with the amount of variation explained by log centroid size (31%), and the significance of the test ($P<0.0001$). At the bottom, figures showing the skull shape associated with the smallest (right) and largest (left) centroid sizes.
5.4 Discussion

My results show significant ontogenetic allometry in the Soay sheep skull, along with a significant shape change throughout 30 years, I also found that these changes resemble in several ways the changes related to smaller sizes. The influence of decreasing size on skull shape was also statistically proven by comparing the vectors of regressions between shape vs. size and shape vs. year of birth. A 60° angle might seem, at first, a great amount of divergence, but is in fact, inconspicuous when we take into consideration the number of dimensions in the shape space. For this particular study I have 20 paired landmarks and 10 unpaired (midline) landmarks, and thus 76 dimensions.

Finding significant phenotypic changes in such short period of time (evolutionary time) is not unusual, as many studies show (Hendry and Kinnison 1999; Filchak et al. 2000; Hendry et al. 2000; Schluter 2000; Kinnison and Hendry 2001; Grant and Grant 2002; Gilchrist et al. 2004; Hargeby et al. 2004; Garant et al. 2005; Schwarz et al. 2005; Hendry et al. 2006; Phillips et al. 2006; Strauss et al. 2006; Carroll et al. 2007; Hendry et al. 2007) (See Carroll et al. 2007 for a review). Evolution on ecological time scales has even been responsible for substantial differentiation among populations within the same species, in some cases enough to represent the early stages of speciation (Filchak et al. 2000; Hendry et al. 2000; Schwarz et al. 2005; Hendry et al. 2007). And since in most cases, natural selection is the main factor driving rapid phenotypic change (Hendry and Kinnison 2001), micro-evolutionary studies can be crucial for making inferences on the role of selection in nature (Carroll et al. 2007).
But there are some examples of rapid phenotypic change in which climate change, and not selection, is the protagonist (Ozgul et al. 2009; Ozgul et al. 2010). Data on hind leg length collected for 20 years (from 1985 to 2005) show that the average body size of the Soay sheep population has been gradually decreasing, despite positive selection for body size (Gratten et al. 2008; Ozgul et al. 2009). We also found the same gradual decrease occurring in centroid size. Over the same period of time, the North Atlantic Oscillation (NAO) has also diminished (Hurrell 1995), meaning milder, shorter winters that allows small, slow growing animals to survive thus increasing the amount of smaller animals in the population. This caused a decline in average body size (Ozgul et al. 2009) and consequently, changes in skull shape, as suggested by my results.

The increased variation in centroid size found in years after 1985 can be explained by the structure of my sample after crash years. The proper monitoring of the island population began in 1985 (Clutton-Brock et al. 2004), and prior to that year, collection of skulls were sporadic, with mostly big rams being sampled, and very few lambs, after 1985, all deceased animals found had their skulls collected (Clutton-Brock et al. 2004). The years with comparatively small average centroid size are usually years in which the population crashed (marked with red circles in Fig. 2), and most of the skulls collected belonged to yearlings (which have higher mortality) (Clutton-Brock et al. 2004), hence the decrease in average centroid size.

The presence of allometry comes as no surprise, since allometry is pervasive among vertebrates (Gould 1966, 1971; Klingenberg 1996; Weston 2003; Christiansen and Adolfsson 2005). The allometric pattern seen in Soay sheep, is found in many mammalian taxa, where smaller animals tend to have shorter snouts
and larger braincases (Zelditch et al. 1992; Abdala et al. 2001; Flores et al. 2003; Weston 2003; Berge and Penin 2004; Giannini et al. 2004; Cardini and O'Higgins 2005; Marroig and Cheverud 2005; Cardini and R. W. Thorton 2006; Flores et al. 2006; Giannini et al. 2010; Cardini and Polly 2013). The shape changes in the skull throughout the study period, as we can see, that have also followed this allometric trajectory.

Any changes in shape caused by changes in developmental timing (growth rates, time of development onset, and end of somatic growth) comprise what we know as heterochrony (Gould 1966; Alberch et al. 1979; McNamara 1986; McKinney 1988; McKinney and McNamara 1991; Klingenberg 2000). Data on body weight gain in the first years of Soay sheep’s lives, have shown that a portion of the population is growing more slowly in years with mild winters (Ozgul et al. 2009). These milder winters have extended the season of grass growth while reducing the length of time that individuals depend on stored fat reserves (Ozgul et al. 2009). This means that the sheep do not need to gain weight as fast as they used to, since they will spend less time depending on fat reserves. Since the sheep are growing more slowly than they used to (see Figure 4 in Ozgul et al. 2009) (and we have observed that the skulls are paedomorphic) thus retaining its juvenile characteristics, I can presume that the Soay sheep are undergoing an heterochronic process known as neoteny (Alberch et al. 1979).

Heterochrony is one of the most recurrent paths to morphological differentiation (McNamara 1988; McNamara and McKinney 2005), and it has been documented in several vertebrates (e.g. (Vrba 1996; Bonnett et al. 2014; Franz-Odendaal and Adriaens 2014; Koyabu and Son 2014; Sheil et al. 2014; Tsai and
Fordyce 2014). And yet, this is the first study to find evidence of heterochrony within a population of wild ungulates. I have shown that rapid changes in temperature can affect not only distribution and life-history aspects of animal populations but also the phenotype of large mammals such as sheep, by altering the their growth rate. The fast pace in which skull morphology in the Soay sheep is changing gives us a new, broader perspective of the extreme effects global warming can have on wild populations.

5.5 Literature cited


129
Cercopithecinae) inferred from geometric analysis of landmark data. The Anatomical Record 275A:1048-1072.


Kinnison, M. T. and A. P. Hendry. 2001. The pace of modern life II: From rates of contemporary microevolution to pattern and process. Genetica 8:145-164.


6. Chapter 6:

Testing the influence of environment, life-history, and genetics on fluctuating asymmetry levels

Abstract

Fluctuating asymmetry has been linked to developmental instability, which in turn can be caused by environmental stress. But FA studies are mostly done in laboratory conditions, and rarely account for the heritable portion of FA. In this study, I am looking for correlations between environmental and/or physiological conditions and the levels of FA of skull shape of a wild population of Soay sheep. I found that the quality of vegetation found in the sheep's territory is significantly correlated to the amount of asymmetry in its skull. Parasite load, breeding success and horn length also significantly affect FA levels on the skull. Other factors such as climate and population density do not seem to influence FA levels. Heritability of FA is low, with $7.51 \times 10^4$ for Procrustes scores and $3.02 \times 10^5$ for Mahalanobis scores. The additive genetic effect and dominance are not significantly different from zero. Other studies corroborate the results found here, where FA is affected by environmental conditions and its heritability is either very low, or zero.

6.1 Introduction

The subtle differences between left and right sides of bilateral characters, or fluctuating asymmetry (FA), has long been related to developmental instability (Van Valen 1962; Palmer and Strobeck 1986; Parsons 1992; Polak 2003; Van Dongen 2006; Graham et al. 2010). Developmental instability is an organism's sensitivity to random perturbations during development and it can be measured as the amount of phenotypic variation resulting from these random perturbations (Polak 2003). Because the left and right sides of organisms share the same genome, and are usually under the same environmental conditions, the differences between the sides are
attributed to random perturbations during development (Klingenberg 2015). These random perturbations can stem from poor environmental conditions (Badyaev et al. 2000; Marchand et al. 2003; Oleksyk et al. 2004; Sánchez-Chardi et al. 2013; Lazić et al. 2015; Lezcano et al. 2015; Maestri et al. 2015), and low genetic quality, usually linked to homozygosis (Debat et al. 2000; Rego et al. 2006; Carreira et al. 2008; Albarrán-Lara et al. 2010; Mikula et al. 2010; Nouvellet et al. 2011; Quinto-Sánchez et al. 2015).

Fluctuating asymmetry of shape has often been associated with various measures of stress, individual quality, and fitness (Badyaev et al. 2000; Marchand et al. 2003; Oleksyk et al. 2004; Sánchez-Chardi et al. 2013; Lazić et al. 2015; Lezcano et al. 2015; Maestri et al. 2015). Most of these studies focus on either comparing levels of FA from different populations, from distinct environments, or different experimental conditions. Studies estimating the inheritance of FA, in the other hand, are quite scarce. Inheritance of FA is difficult to estimate partly due to the statistical power (large sample sizes) required for such analyses (Van Dongen 2006, 2007), and partly due to the known developmental origins of FA (Palmer 1994; Polak 2003). Some researchers claim that FA is influenced by non-additive genetic effects, such as dominance and especially epistatic interactions among genes (Leamy et al. 2002; Leamy and Klingenberg 2005; Leamy et al. 2005; Leamy et al. 2015). They also state that, except for unusual populations events, epistasis supresses additive genetic variation in FA, hence the low levels of heritability found so far (Leamy and Klingenberg 2005). Studying levels of shape FA and its heritability in the wild is even more difficult. This is because large sample sizes, pedigree data, as well as data on
environmental variables that might be linked to FA, are considerably harder to obtain for wild populations.

It is important to fully understand the extent of the genetic basis of FA because only then, we will be able to properly use FA as a measure of developmental instability (Leamy and Klingenberg 2005; Van Dongen 2006). Estimating inheritance of FA will allow us to know the amount of genetic variance of developmental instability, and thus, if it can respond to selection (Pélabon et al. 2004; Leamy and Klingenberg 2005; Van Dongen 2006). This is particularly important, since developmental instability has been suggested to play a major role in sexual selection and mate choice (Van Dongen 2006).

In this study, I calculated FA levels of skull shape in one of the most well-studied populations of wild mammals in the world, the Soay sheep of St. Kilda. This population has been monitored for 30 years, and a series of different environmental, morphological, ecological and behavioural data has been gathered throughout this period (Boyd et al. 1964; Hutchings et al. 2002; Clutton-Brock et al. 2004b; Crawley et al. 2004). I used this abundance of data to assess the correlation between FA and several extrinsic factors, including population density, vegetation quality, climate, and intrinsic factors, such as birth weight, parasite burden, horn size and breeding success. I also use the available pedigree data to estimate heritability of FA and the effect of genetic dominance on these estimates.
6.2 Methods

The Soay sheep population under study here, come from St. Kilda archipelago, in Scotland (Clutton-Brock and Pemberton 2004b). Every sheep is tagged and weighed individually within a few days after birth, and a skin sample is taken for paternity analysis (Boyd et al. 1964). The population has been monitored since 1964 and the skulls of deceased animals were collected from 1976 until 2006. All the skulls collected in St. Kilda were deposited at the National Museum of Scotland, in Edinburgh.

6.2.1 Fluctuating asymmetry scores

We used geometric morphometrics techniques to obtain fluctuating asymmetry scores for each individual. A single observer (EMD) digitized 50 landmarks in three dimensions (Fig. 1) using a MicroScribe G2X. The total sample size was 1,528 skulls, and each skull was digitized twice to correct for measurement errors. Skulls have object symmetry, which means that each half of the skull is a mirrored image of the other (Klingenberg et al. 2002). This allowed most of our landmarks to be placed correspondingly in both sides of the skull. The raw landmark data was then imported into MorphoJ (Klingenberg 2011), where we ran a Procrustes fit to extract shape information, followed by a Procrustes ANOVA (Klingenberg and McIntyre 1998; Klingenberg et al. 2002), from which we obtained Procrustes and Mahalanobis FA scores (Klingenberg and Monteiro 2005). The differences between left and right sides are calculated as the Procrustes distances between the original configuration of landmarks, and a reflected and relabelled configuration (Komori et al. 2009; Muñoz-Reyes et al. 2012; Abend et al. 2015;
Klingenberg 2015; van Dongen 2015). But quantifying FA as Procrustes distances does not take into account that FA is not distributed equally in all directions. Calculating FA as Mahalanobis distances, on the other hand, scales each direction of FA distribution by the inverse of its variation of asymmetries, so that the new space has an equal amount of FA in every direction (Klingenberg and Monteiro 2005; Klingenberg 2015). In our analyses we used both Procrustes and Mahalanobis distances.

Figure 1. Scheme showing landmark locations on A) dorsal view; B) lateral view; and C) ventral view.
6.2.2 Environmental & Morphological Variables

Ever since the Soay sheep population of St. Kilda started being monitored, in 1964, several environmental, ecological, and morphological data were recorded. In this section, I describe the methodology involved in obtaining these data, as detailed in previous studies on the Soay sheep of St. Kilda.

Birth weight - Most lambs are weighed on the second day after their birth, but some lambs might only be caught up to a week after they are born. In 1988 and 1989 a sample of lambs was caught and reweighed every two days, thus providing data on body mass increase during the first 6 weeks (Robertson et al. 1992). These growth rates were subsequently used to estimate birth weight, reducing the observed weight by 108g for every twenty-four-hour period since birth (Clutton-Brock and Pemberton 2004a). Upon the digitizing process I was able to match the skull being measured with the birth weight recorded on the project’s database.

Horn length & Parasite burden - Every year between July and August, over 50% of the sheep living in the Village Bay area are captured. They are weighed, their horn and leg length are measured, samples of wool, blood and feces are collected, ectoparasites are counted, and the sheep are released. As a measure of parasite burden I used only Strongyle faecal egg count (FEC), since they are the most prevalent and abundant nematodes found in gastrointestinal tracts of Soay sheep (Gulland and Fox 1992; Coltman et al. 1999; Craig et al. 2006). Although faecal egg count is not a direct count of parasites, previous work on Soay sheep has shown a strong correlation between FEC and parasite burden (Wilson et al. 2004).
Breeding success - I calculated the number of offspring an animal produced during its life-time through pedigree information. Maternity was inferred by observation of maternal care behaviour, and paternity was inferred by analyses on single nucleotide polymorphisms (SNPs) (Kijas et al. 2009), using maximum-likelihood implemented on CERVUS (Marshall et al. 1998). For each lamb, a log-likelihood ratio score was calculated for each possible father, and criteria for assigning paternity to the male with the highest score with 95% and 80% confidence were obtained by simulation (Pemberton et al. 1999; Pemberton et al. 2004). Because males have considerably higher potential breeding success than females, I analysed the two sexes separately.

Population density - Estimations on population density were obtained through censuses that are carried out on the Village Bay area three times per year, with ten replicates each time (Clutton-Brock and Pemberton 2004a). The numbers of Soay sheep in St. Kilda has varied wildly during the study period. Between 1985 and 2006, total numbers fluctuated from 663 to 222, while in the study area (Village Bay) those numbers varied from 211 and 591 (Clutton-Brock et al. 2004a). In some years there is a high mortality of sheep during winter, in which nearly half of the population dies (Clutton-Brock et al. 2004a). These high mortality years, or crash years, are directly linked with high population densities and winter severity (Clutton-Brock et al. 2004a).

Vegetation quality - Various plant communities are widespread throughout the whole Village Bay study area (Gwynne et al. 1974; Clutton-Brock et al. 2004b; Crawley et al. 2004). The community that is most widely used by sheep, and thus granted as the one with highest vegetation quality is the Holcus-Agrostis habitat (Clutton-Brock et
But different parts of the study area have more or less abundance of high quality grass. In this study, I used the percentage coverage of *H. lanatus* as a measure of vegetation quality of the area (Regan *et al.* *in press*). The percentage coverage of the species in each area was estimated by eye (to the nearest 5%). The estimation of cover by eye was then calibrated against sorted biomass samples (Regan *et al.* *in press*). In order to estimate the amount of *H. lanatus* available for each individual, we calculated the average coverage of *H. lanatus* in each area, weighed by the number of times the sheep was seen in them.

**NAO index** - The North Atlantic Oscillation (NAO) (Rogers 1984) has a great influence on the island’s weather (Clutton-Brock *et al.* 2004b). Fluctuations in NAO provide an index of weather conditions, where high NAO years have warm, wet winters, whereas low NAO years have colder, drier winters (Forchhammer *et al.* 1998a; Belgrano *et al.* 1999; Clutton-Brock *et al.* 2004b). In St. Kilda, high values of NAO index are correlated with reduced survival in the Soay sheep population, because storms restrict movements and increase heat loss (Forchhammer *et al.* 1998b; Post and Stenseth 1998; Milner *et al.* 1999a; Milner *et al.* 1999b; Post and Stenseth 1999; Post and Forchhammer 2002; Clutton-Brock *et al.* 2004b). Measures of NAO index are available at [https://climatedataguide.ucar.edu/climate-data/hurrell-north-atlantic-oscillation-nao-index-station-based](https://climatedataguide.ucar.edu/climate-data/hurrell-north-atlantic-oscillation-nao-index-station-based).

**6.2.3 Statistical analyses**

I used multiple regression analyses to test the significance of correlations between FA Procrustes and Mahalanobis scores and: *Holcus* coverage; NAO index (on year of birth and death); birth weight; breeding success; horn length; parasite
burden; and population density. Because FA levels increase with age and the population levels are growing throughout the study period, age and year of birth were included as covariates in every regression analyses. I then computed the partial correlation scores and \( p \) values for the variable of interest.

I also ran Analyses of variance (ANOVA) to test for differences in FA levels between sheep that live on different areas of the island, and sheep born or deceased in years that were categorized according to population crashes. Those years can be pre-crash, which entails animals born in the spring just before the crash (which happens during winter); post-crash (animals born in the following spring, after the crash); and the years in between, in which the population is growing, but haven’t reached its highest numbers yet. The main differences between these classes of years are population density which influences food availability and thus, body condition of the ewes.

6.2.4 Quantitative genetics

To calculate the additive genetic and dominance effects of FA scores (Procrustes and Mahalanobis), I performed a Restricted Maximum-Likelihood analysis based on an animal model (Lynch and Walsh 1998; Kruuk 2004; Kruuk and Hill 2008; Wilson et al. 2010) using Wombat software (Meyer 2007). Sex, horn morph, and age were set as fixed effects, individuals and dominance as random effects, while centroid size was set as a covariate. The fraternity coefficient matrix, or dominance matrix, was built using ‘nadiv’ package in R (Wolak 2012). I used likelihood ratio tests to determine if there is significant dominance effect in the
model (Lynch and Walsh 1998). The heritability *sensu stricto* was estimated from the ratio of the additive genetic variance to the total phenotypic variance: \( h^2 = \frac{V_A}{V_T} \).

### 6.3 Results

Of all the variables tested, centroid size, age, birth year, vegetation quality, and breeding success in males significantly correlate with both Procrustes and Mahalanobis FA scores (Table 1). Parasite load was only correlated to Mahalanobis FA scores, whereas horn length in males was only correlated with Procrustes FA scores. Year of birth, vegetation quality and breeding success have a negative effect on FA scores, whereas age, horn length, and parasite load have a positive effect (See Table 1 for partial correlations and *p* values).

Table 1. Partial correlation scores and their respective *p* values for each variable tested for correspondence with Procrustes and Mahalanobis FA scores.

<table>
<thead>
<tr>
<th>Variables/ FA scores</th>
<th>Procrustes</th>
<th>Mahalanobis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Centroid Size</td>
<td>0.09 (p=0.0003)</td>
<td>0.07 (0.004)</td>
</tr>
<tr>
<td>Age</td>
<td>0.1 (p=0.00002)</td>
<td>0.08 (0.0009)</td>
</tr>
<tr>
<td>Birth year</td>
<td>-0.06 (0.01)</td>
<td>-0.07 (0.002)</td>
</tr>
<tr>
<td>Vegetation quality</td>
<td>-0.08 (p=0.0012)</td>
<td>-0.10 (p=0.0002)</td>
</tr>
<tr>
<td>Birth weight</td>
<td>-0.04 (p=0.21)</td>
<td>-0.03 (p=0.3)</td>
</tr>
<tr>
<td>Male Horn length</td>
<td>0.08 (p=0.03)</td>
<td>0.06 (0.3)</td>
</tr>
<tr>
<td>Female Horn length</td>
<td>0.06 (p=0.19)</td>
<td>0.05 (p=0.31)</td>
</tr>
<tr>
<td>Parasite load</td>
<td>0.05 (p=0.08)</td>
<td>0.07 (p=0.02)</td>
</tr>
<tr>
<td>Male Breed. success</td>
<td>-0.12 (p=0.04)</td>
<td>-0.18 (p=0.001)</td>
</tr>
<tr>
<td>Female Breed. success</td>
<td>-0.05 (p=0.41)</td>
<td>-0.09 (p=0.13)</td>
</tr>
<tr>
<td>Population density</td>
<td>0.01 (p=0.69)</td>
<td>0.003 (p=0.88)</td>
</tr>
<tr>
<td>NAO index_birth</td>
<td>0.029 (p=0.26)</td>
<td>0.028 (p=0.26)</td>
</tr>
<tr>
<td>NAO index_death</td>
<td>0.037 (p=0.14)</td>
<td>0.026 (p=0.3)</td>
</tr>
</tbody>
</table>

The Analysis of variance on different status of death year (whether the year was pre-crash, post-crash, or in between crashes) showed no significant (p=0.47) differences in FA levels between the groups. The ANOVA on birth year status (Fig. 2), on the other hand, showed significant (p=0.018) effect on FA levels. A post-hoc Tukey test between the different statuses of birth year found that animals born in between crash years had significantly smaller Procrustes FA levels than animals born in post-crash years (p=0.03). When using the same test on Mahalanobis scores, I found that animals born in between crash years are significantly less asymmetrical than animals born in pre- and post-crash years (p=0.02 and p=0.01, respectively). The ANOVA on different areas of the island also showed a significant (p=0.04) effect on FA levels, but the post-hoc Tukey test was unable to find significant differences between FA levels of animals occupying different areas.
Figure 2. Results from Analysis of Variance (ANOVA) showing the respective mean (black circles) and standard deviations (vertical bars) of FA scores (Procrustes and Mahalanobis) for animals born in pre-crash, between-crashes, and post-crash years.

Estimates of the heritability and variance components for additive genetic effect and dominance are given in Table 2. Neither additive genetic nor dominance effect were statistically different from zero. The Likelihood-ratio test on Procrustes scores returned a difference of 0.006399 between the reduced and complete (taking dominance into account) models, and on Mahalanobis scores, of 0.001193. A chi-
square test on both values returned non-significant $p$ values: 0.99 and 0.97 for Procrustes and Mahalanobis FA scores, respectively.

Table 2. Values of heritability, additive genetic variance ($V_A$), dominance effect variance ($V_D$), and their respective Standard Errors (SE$_A$ and SE$_D$, respectively) for both Procrustes and Mahalanobis FA scores.

<table>
<thead>
<tr>
<th></th>
<th>Heritability</th>
<th>$V_A$</th>
<th>$V_D$</th>
<th>SE$_A$</th>
<th>SE$_D$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Procrustes</td>
<td>7.51 x 10^{-4}</td>
<td>5.5 x 10^{-5}</td>
<td>3.42 x 10^{-3}</td>
<td>0.09</td>
<td>0.057</td>
</tr>
<tr>
<td>Mahalanobis</td>
<td>3.02 x 10^{-4}</td>
<td>2.5 x 10^{-5}</td>
<td>1.58 x 10^{-4}</td>
<td>0.097</td>
<td>0.011</td>
</tr>
</tbody>
</table>

6.4 Discussion

6.4.1 Fluctuating asymmetry and fitness

Our results show that animals with higher breeding success rates, or fitness, were significantly more symmetric. Higher breeding success has also been associated with decreased levels of FA in antler length in red deer (Kruuk et al. 2003); mandible shape in shrews (Badyaev et al. 2000); and in several other species (Leung and Forbes 1996; Møller and Thornhill 1998; Møller 1999). Because only males seem to show a negative correlation between FA and breeding success, we can conclude that either more symmetrical males have a better chance at winning contests with other males, and thus getting access to females, or that more symmetrical males are selected by females. There are some indications of female choice for symmetrical ornaments (Møller 1990; Møller and Hoglund 1991; Møller 1992a), which does not apply for this study, since the trait in question is not visible. In this case, we can presume that lower levels of FA in successful males are associated with better environmental conditions during their development that allowed such males to be bigger, stronger, and thus, have higher competitive ability (Møller 1997).
The ability to avoid parasitism is also considered as a measure of fitness, related to survival. In the population of Soay sheep from St. Kilda, strongyle intestinal nematodes are known to play a major role in the mortality of sheep (Gulland and Fox 1992; Coltman et al. 1999; Craig et al. 2006). And as in many other species (Escós et al. 1995a; Escós et al. 1995b; Møller 1995; Folstad et al. 1996), we found an association between FA levels and parasite load in the Soay sheep. There have been extensive discussions on the mechanism behind the correlation between parasitism and developmental instability (Møller 1996a; Thornhill and Møller 1997). One hypothesis is that animals with high FA levels are more susceptible to parasites (Møller 1992b; Polak 1993; Møller 1996b, a), since poor environmental conditions reduces the efficiency of the immune system, while also increasing FA (Chandra and Newberne 1977; Gershwin et al. 1985; Lagesen and Folstad 1998). Another hypothesis is that the physiological strains caused by parasites increase FA levels (Folstad et al. 1996; Møller 1996a; Thornhill and Møller 1997).

6.4.2 Fluctuating asymmetry and environment

The quality of vegetation found in a sheep’s home range appears to affect skull symmetry, even though the environmental variations among these areas are subtler when compared to previous FA studies. We also found that animals born in years between crashes have less asymmetrical skulls. Those years are the ones with better environmental quality, since population growth is coupled with food availability (Milner et al. 1999b). The years preceding crashes are the ones with least food availability, and thus, worst environmental quality. And the years following crashes, although having the lowest population densities, do not present ideal
environmental conditions, since plant communities are still recovering from the previous year's resource depletion (Crawley et al. 2004). Usually, higher asymmetry levels are found in habitats disturbed by pollution, radiation or in urban areas (Badyaev et al. 2000; Marchand et al. 2003; Oleksyk et al. 2004; Sánchez-Chardi et al. 2013; Lazić et al. 2015; Lezcano et al. 2015; Maestri et al. 2015). Our results show that even subtle differences in environmental quality found within the same habitat can increase asymmetry levels in the skull.

6.4.3 Fluctuating asymmetry and size

The relationship between age and FA could be an artefact stemming from the association between age and the overall size of an animal, even though this variable was included in the analysis. In our analysis, age and the size of the skull, or centroid size, are correlated ($r=0.63$, $p<0.0001$), and both are also correlated to FA levels (Table 1). The positive correlation between a structure's size and its asymmetry has been discussed at length in the literature (Palmer and Strobeck 1986; Palmer 1994; Leung 1998). It is believed that bigger individuals, or traits, have more 'opportunity' to become asymmetric (Leung 1998), and the same rationale may be applied to older individuals. Horn size is also linked to centroid size, but since it is only significantly correlated with Procrustes FA scores, and the associated $p$ value is only marginally significant, the relationship between horn size and FA levels is inconclusive.

6.4.4 Inheritance of fluctuating asymmetry

Our quantitative genetics analysis found low heritability for both Procrustes and Mahalanobis FA scores. And the estimated variances for additive genetic effect
and dominance effect are indistinguishable from zero (Kruuk et al. 2003). We also failed to find significant effect of dominance in our model.

There are two main reasons because heritability of FA is hardly ever found in any trait or organism. One of the reasons could be the inherent developmental nature of FA, and the other reason is the high statistical power required, but seldom achieved, to run the appropriate analyses (Klingenberg 2015). It is already hard enough to obtain ideal sample sizes of FA measurements that minimize the impact of measurement errors, but it is even harder to obtain enough data to run quantitative genetics analysis with FA scores (Van Dongen 2007; Klingenberg 2015). Even studies with large sample sizes, such as ours, still lack statistical power to estimate heritability of FA, since simulations show that a sample size of 5,000 and higher is required to obtain realistic estimations (Van Dongen 2007).

Although the majority of studies, including ours, found low or no heritability of fluctuating asymmetry (Radwan et al. 2003; Leamy and Klingenberg 2005; Carter and Houle 2011; Tsujino and Takahashi 2012, 2014), few studies, by using ideal sample sizes or different methods, were able to find genetic basis of FA in drosophila (Breuker et al. 2006; Tsujino and Takahashi 2012) and mouse skulls (Willmore et al. 2006), and even effects of dominance and epistasis (Klingenberg and Nijhout 1999; Klingenberg et al. 2001; Leamy et al. 2002; Leamy and Klingenberg 2005; Leamy et al. 2005; Leamy et al. 2015). So we cannot possibly affirm that there is no heritability (nor dominance) of skull shape FA in the Soay sheep, since our results may merely reflect the lack of statistical power of our analyses.
6.5 Literature Cited


7. General Conclusions

This thesis explores some of the factors behind variation of skull shape in a population of Soay sheep. To do that, I combined geometric morphometric and quantitative genetic methods to analyse patterns of variation, genetic basis and evolution of skull shape. The amount of morphological, ecological and genetic data available for this particular population of Soay sheep, provided an invaluable source of information, which certainly differentiates this study from previous studies on variation of skull shape.

In the second chapter of the thesis, I analysed magnitudes and patterns of integration and modularity in different aspects of skull shape variation (individual, genetic and fluctuating asymmetry). In all of the aspects of skull shape variation, we found low levels of integration- in comparison to previous studies (Lieberman et al. 2000; Drake and Klingenberg 2010). Also in all aspects of variation, we found that variation is distributed in two modules, face and neurocranium. The fact that we found the same patterns of modularity across individual, genetic, and FA variation indicates a high degree of interaction between the processes behind variation in skull shape (Klingenberg 2003, 2005, 2008). This pattern of covariation between face and neurocranium can be explained by their different developmental origins (neural crest and paraxial mesoderm, respectively) (Couly et al. 1992; Jiang et al. 2002; Noden and Trainor 2005). The low integration and modularity found in the skull, might affect the patterns of variation in a broader, evolutionary context (Wagner 1996).

The third chapter contains analyses that estimate the response to selection after hypothetical selection scenarios. The results found in the third chapter corroborate some of the conclusions from the previous chapter. I found that certain
shape changes, like the retraction of the nose are evolutionarily unconstrained. This requires a certain independence between parts of the skull, as suggested by the presence of modules. And although the neurocranium constitutes a module, when selection acts to reduce the braincase, there is a strong correlated response for an elongation of the face. This means that even though traits within the neurocranium covary more within than between modules, there is still genetic integration constraining the reduction of braincase. Constraining certain shape changes while leaving others unconstrained, could be of extreme evolutionary value (Wagner and Altenberg 1996; Eroukhmanoff 2009).

After seeing how the skull responds to hypothetical selection regimes, we wanted to know: which skull shape is actually being selected on the Soay sheep? We answer this question in the fourth chapter, where we use a new method of estimating selection from geometric morphometrics data. Selection on skull shape is strong, when compared to other traits in the Soay sheep (Milner et al. 1999; Milner et al. 2004; Robinson et al. 2006). And the skull shape selected in males differs from the one selected in females. In males, selection favors broader nasal bones (in relation to the width of the whole skull) and wider molar teeth, in relation to the pre-molar teeth, whereas in females, it favours narrower nasal bones (also in relation to the rest of the skull) and flattened braincases. But as we discovered in the third chapter, the response to selection does not always portray the shape that is under selective pressures. It is possible that some of the outcome of selection we see is actually a correlated response to selection in a different trait. It is also possible that skull shape is not the prime target of selection, only the by-product of selection in other morphological, physiological and even behavioural traits. Because of that, we must be
cautious when trying to determine what is causing selection towards a particular shape.

Other factors besides selection can also affect skull shape, and sometimes, changes can occur in the phenotype, even when they are not being selected for. This is the case for the Soay sheep from St. Kilda. In the fifth chapter, we found evidence that climate change might have given rise to a heterochronic process, more specifically, paedomorphosis. Milder winters, combined with the early arrival of spring, have allowed slow growing, small sheep to survive, thus decreasing the mean body size of the population (Ozgul et al. 2009). We found that sheep are not only decreasing in size (on average), but their skull shape is also becoming more juvenilized, or paedomorphic (McKinney 1988; McKinney and McNamara 1991; Klingenberg 1998). When we compare the shape changes occurring throughout these 40 years, we see that the shape changes associated with a paedomorphic skull do not resemble in any way the selected skull shape seen in chapter three. This shows that, in some cases, environmental factors may play a bigger role in phenotypic evolution than natural selection.

Another way environmental factors affect the phenotype, is by producing random variations in skull symmetry known as fluctuating asymmetry. FA has often been related to developmental instability caused by stress during development, or poor genetic quality (homozygosity) (Badyaev et al. 2000; Marchand et al. 2003; Oleksyk et al. 2004; Sánchez-Chardi et al. 2013; Lazić et al. 2015; Lezcano et al. 2015; Maestri et al. 2015). In the sixth chapter, we found that low quality of vegetation, high parasite load, and low breeding success are related to high levels of FA. In concordance with previous studies (Radwan et al. 2003; Leamy and
Klingenberg 2005; Carter and Houle 2011; Tsujino and Takahashi 2012, 2014), we could not find any significant heritability or dominance in FA scores, but this might be due to a lack of statistical power in our analysis.

Throughout five chapters, I have assessed levels of integration and patterns of modularity, measured the effects of genetic integration in the outcome of selection, estimated magnitude and direction of natural selection, analysed how the shape is actually changing throughout the years, and measured fluctuating asymmetry levels, found which factors affect them, and how heritable they are. The main conclusion of this thesis is that development seems to play a major role in skull shape variation on the Soay sheep. The modularity analysis shows that development affects how the variation within the skull is distributed (between face and neurocranium), which is an indication of the patterns of variation in an evolutionary context (Wagner 1996). And indeed, we saw that the shape changes undergone by the skull of Soay sheep throughout 30 years are related to changes in developmental timing. The skull is becoming more paedomorphic, which entails a comparatively larger neurocranium and shorter snout (face) (Cardini and Polly 2013). We also saw that the skull shape being selected does not have paedomorphic characteristics. This means that the way the skull shape has been changing, is not the skull shape being selected for. These results suggest that development, and not selection, is the main factor behind variation of skull shape.
7.1 Literature Cited


