ST. KILDA SOAY SHEEP PROJECT:
ANNUAL REPORT 2020

J.G. Pilkington¹, X. Bal², D. Childs³, J. Dorrens¹, Y. Corripio-Miyar³, A. Fenton¹⁰, M. Fraser⁸, A. Free¹¹, H. Froy⁹, L. Harrington¹⁷, A. Hayward³, H. Hipperson², W. Huang¹, C. James¹, S.E. Johnston¹, F. Kenyon³, S. Knott¹, H. Lemon¹, V. Litzke⁵, D. McBean³, L. McNally¹, T. McNeilly³, E. Mittell², M. Morrissey⁵, P. Navarro¹, D. Nussey¹, R. J. Pakeman⁷, A. Pedersen¹, J.M. Pemberton¹, S. Ravindran¹, J. Slate², L.A. Seeker⁹-I.R. Stevenson⁶, M.A. Stoffel¹, M, Sutherland², A. Sweeny¹, S.L. Underwood¹, H. Vallin⁸, K. Watt¹, R.V. Wilbourn¹.

¹Institute of Evolutionary Biology, University of Edinburgh.
²Department of Animal and Plant Sciences, University of Sheffield.
³Moredun Research Institute, Edinburgh.
⁴Royal (Dick) School of Veterinary Studies, University of Edinburgh.
⁵School of Biology, University of St. Andrews.
⁶Sunadal Data Solutions, Penicuik.
⁷James Hutton Institute, Craigiebuckler, Aberdeen.
⁸Institute of Biological, Environmental & Rural Sciences, Aberystwyth University.
⁹Norwegian University of Science and Technology, Trondheim.
¹⁰Institute of Integrative Biology, University of Liverpool.
¹¹Institute of Quantitative Biology, Biochemistry and Biotechnology, University of Edinburgh.

POPULATION OVERVIEW .............................................................................................................. 2
REPORTS ON COMPONENT STUDIES .......................................................................................... 4
Vegetation ................................................................................................................................. 4
The Ecology Within: The impact of gut ecosystem dynamics on host fitness in the wild .......... 5
The association between reproduction and telomere length in wild Soay sheep females .... 8
Inbreeding depression across the lifetime.................................................................................... 10
Quantifying variation between and within coat colour morphs in Soay sheep .................... 11
The genetics of body size using high density genetic markers .............................................. 16
Sexual selection on MHC variation in Soay sheep................................................................. 18
PUBLICATIONS .......................................................................................................................... 20
ACKNOWLEDGEMENTS .......................................................................................................... 20
APPENDIX A: PERSONNEL NEWS & SCHEDULE OF WORK .................................................. 21
Population Overview

The sheep population on Hirta entered 2020 at a moderately high level and there was relatively high mortality over winter with 127 tagged animals being found dead within the study area. Lambing began on the 4th of April with 66.25% of lambs born surviving (Fig. 1).

Figure 1. The temporal distribution of lamb births during 2020.

In December 2020, 680 tagged sheep were believed to be alive on Hirta, of which 514 regularly used the study area, an increase of 1.17% using the study area since the previous year. The age distribution of the population is shown in Figure 2 and changes in sheep numbers in the study area over time are shown in Figure 3.

Figure 2. Age distribution of tagged Soay sheep presumed to be alive at the end of 2020.
Figure 3. The number of sheep counted on the whole island and the number of tagged sheep regularly using the study area since 1985.

One whole-island count yielded 1379 tagged and untagged sheep, with the details displayed in Table 1. The total population had decreased by 23.8% since summer 2019 when it was 1810. This gives a delta (calculated as \( \ln \left( \frac{N_{t+1}}{N_t} \right) \)) of -0.272. The whole island count is also shown in Figure 3.

Table 1. Demographic and geographic distribution of sheep observed during the count of Hirta on August 24th 2020. Coat colours are DW = dark wild, DS = dark self, LW = light wild, and LS = light self.

<table>
<thead>
<tr>
<th>Location</th>
<th>Females</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DW</td>
<td>DS</td>
<td>LW</td>
<td>LS</td>
<td>DW</td>
<td>DS</td>
<td>LW</td>
<td>LS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conachair/Oiseval</td>
<td>161</td>
<td>7</td>
<td>40</td>
<td>1</td>
<td>52</td>
<td>2</td>
<td>7</td>
<td>1</td>
<td>108</td>
<td>379</td>
<td></td>
</tr>
<tr>
<td>Mullach Bi / Cambir</td>
<td>205</td>
<td>4</td>
<td>63</td>
<td>6</td>
<td>35</td>
<td>2</td>
<td>18</td>
<td>1</td>
<td>171</td>
<td>505</td>
<td></td>
</tr>
<tr>
<td>Ruaival/Village</td>
<td>188</td>
<td>7</td>
<td>49</td>
<td>5</td>
<td>56</td>
<td>0</td>
<td>18</td>
<td>1</td>
<td>171</td>
<td>495</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>554</td>
<td>18</td>
<td>152</td>
<td>12</td>
<td>143</td>
<td>4</td>
<td>43</td>
<td>3</td>
<td>450</td>
<td>1379</td>
<td></td>
</tr>
</tbody>
</table>
Reports on component studies

Vegetation.

Robin Pakeman

The March trip hit the buffers as St Kilda was locked down. Fortunately, Xavier Bal was already in residence on the island, so many of the measurements including sward height, herbage quality, biomass and productivity could still be recorded.

The highlight of the summer trip was the removal of the old-style pyramid cages as they were getting increasingly battered and maintenance was a continuing commitment. In their place are some cages with a welded steel mesh, which should stand up to the weather and the sheep. They are even harder to spot than the pyramids and so don’t distract visually from the landscape. In all other ways the now standard vegetation monitoring protocol was carried out.

Figure 4. The new-style grazing exclosure now set up on Hirta.
The Ecology Within: The impact of gut ecosystem dynamics on host fitness in the wild.


In March 2019, we began a study into the gut communities of the Soay sheep on St Kilda, which is funded by a large NERC grant. The study relies on the collection of faecal samples on a seasonal basis from Soay sheep that are part of the long-term study in Village Bay – this has been occurring four times per year (March, May, July and October) and through heroic efforts in the field over the last 12 months (led by Xavier Bal and Jill Pilkington) this sampling has continued as planned despite the COVID pandemic. The samples collected in October 2020 represented two full years of sampling for this project with a grand total of 1,702 samples collected. See Table 2 for break down by season and age class.

**Table 2.** Number of faecal samples collected from known sheep over two years of study. Note that generally we have not sampled lambs before three months of age, except in May 2019 when a number of younger lambs were sampled for the gut microbiota analysis.

<table>
<thead>
<tr>
<th></th>
<th>2019</th>
<th></th>
<th>2020</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>March</td>
<td>May</td>
<td>July</td>
<td>Oct</td>
</tr>
<tr>
<td>Lambs</td>
<td>49</td>
<td>120</td>
<td>122</td>
<td></td>
</tr>
<tr>
<td>Adults</td>
<td>161</td>
<td>162</td>
<td>121</td>
<td>122</td>
</tr>
</tbody>
</table>

Most of the individuals sampled to date have been repeatedly sampled over the two years, allowing us to understand the factors influencing changes in the gut community over time within individuals and to separate this from consistent among individual differences in gut community structure. We have now collected samples from 398 different individuals, 171 of whom have been sampled four or more times, and 95 sampled in all eight seasons.

**Microbiota:** One major strand of work within this project is to use these faecal samples to study variation in the gut ‘microbiota’: the community of commensal bacteria living within the gut that play an essential role in allowing the sheep to break down their food and shape immune development, among other functions. Despite delays caused by laboratory closure during the COVID pandemic, we have now finished extracting DNA from all 1,702 faecal samples collected to date and, having completed pilot studies that demonstrate we have a robust and repeatable approach to measuring the gut microbiota using next-generation sequencing, we are
currently preparing these DNA extracts for sequencing. This will produce data which will allow us to examine how the composition of the gut bacteria community in the Soay sheep varies with age, season, sex and across space within the study area.

**Parasitology:** The project also aims to develop our understanding of the complex community of parasitic nematode worms which infect the Soay sheep. Past work on St Kilda has identified 11 species of these worms in the gastrointestinal tracts of dead animals, and the long-term study routinely collects faecal samples in August from study animals and uses counts of morphologically distinct eggs collected from the faeces to monitor these parasites. Samples collected in 2019 and 2020 have all been analysed in this way, to determine the number of eggs present across a range of parasite groups and we have analysed how these counts vary with season, age and sex. Our results, focussing on the most prevalent group of Strongyle worms reveals a dramatic increase in egg counts in spring in ewes that give birth compared with the ewes that are barren (Fig. 5). This fits with observations in Soay and domestic sheep that worm burdens increase in the period around birth, as the ewes’ immune function is suppressed through the later stages of gestation and early in the lactation period. Our results show that male Soays have a different seasonal pattern to females, showing a sustained rise in worm eggs numbers through spring and summer. Among adults, raised Strongyle egg counts in males compared to females are only present in summer months (Fig. 5).

![Graph showing parasite egg counts by season and sex](image)

**SexRepro**
- FemaleLamb
- FemaleNoLamb
- Male
Figure 5. Figure showing seasonal variation in Strongyle faecal egg counts (FEC) across five sampling seasons, among different reproductive groups of adult sheep (red: ewes that gave birth to at least one lamb in Spring 2019; pink: barren ewes that did not produce a lamb; blue: males). Points are means with standard error bars around them, with raw data presented underlain.

The microscopy-based counting of worm eggs groups five different species of Strongyle worm together into a single count, because the eggs of these different species cannot be differentiated by eye. We are using the samples collected as part of this study and next-generation sequencing approaches to understand the dynamics of these different Strongyle parasites. This involves culturing faecal samples so that worm eggs develop into early stage larvae, from which we then extract DNA and use meta-barcoding methods to estimate the proportions of different species of worm. Efforts to do this with the samples listed in Table 2 are well underway, and our first batch of sequence data was delivered to us in January 2021 and is currently being analysed.

Immunology: Alongside efforts to understand the dynamics of the bacteria and worms living in the guts of the Soay sheep, we have been using blood samples collected during the August catch to implement new immunological assays and better understand how variation in the immune response shapes infection with different kinds of parasites. Samples collected in August 2019 have been used to measure how different types of immune response, mediated by specialised white blood cells called ‘T helper cells’, relate to infection with two major groups of gut parasite: the Strongyle worms and single-celled protozoans called Coccidians. From past work by immunologists in the laboratory, we expect that so-called ‘T helper 1’ type immune responses are most important for dealing with single-celled parasites, while so-called ‘T helper 2’ type responses are essential for resistance to bigger parasites like worms. Our initial results provide rare support for the different roles of these two types of T helper responses from a wild population. We have found that sheep with an apparently strong T helper 1 profile have lower Coccidia counts, while those with a strong T helper 2 profile have lower Strongyle worm counts. We will continue these assays of the T helper response using samples collected in August 2020 and complement these with further assays of various types of antibodies, and plan to relate these measures to our data on the communities of bacteria and worms in the guts of the sheep as they become available.

Diet / vegetation: A key aim of this project is to develop techniques to monitor the diet composition of the Soay sheep, using next-generation sequencing to look at the different plant species in the faecal samples we collect on St Kilda. To validate such an approach, a small experimental population of Soay sheep has been established at Pwllpeiran Upland Research station (part of the University of Aberystwyth), where animals can be fed diets of known and controlled plant constituents. Trials with the 15 Soays have been ongoing, faecal samples have been collected, and informatic tools developed to select the best genetic regions in plants to allow identification of different species using meta-barcoding approaches. Laboratory work is ongoing to extract DNA from these samples and prepare them for sequencing, and the data produced will help validate and calibrate methods which will then be used on the samples collected on St Kilda (Table 2) to examine variation in diet in the Soay sheep.
The association between reproduction and telomere length in wild Soay sheep females.

Sanjana Ravindran, Hannah Froy, Sarah L. Underwood, Jennifer Dorrens, Luise A. Seeker, Kathryn Watt, Rachael V. Wilbourn, Jill G. Pilkington, Lea Harrington, Josephine M. Pemberton and Daniel H. Nussey

Think of your shoelaces and the plastic tips that prevent the laces from fraying – in a similar manner, a repetitive sequence of DNA called a telomere is present at each end of your chromosomes which protects these chromosomal ends. Unfortunately, telomeres tend to get shorter over time and when they get too short, the cells containing them can become dysfunctional and die. In humans and other vertebrates, we can measure an individual’s average telomere length (TL) in the laboratory from DNA contained in the cells in blood samples. Blood TLs tend to decline with age and shorten when individuals are exposed to stress, whilst having short TL seems to predict reduced prospects of survival in birds and mammals. Although TL is clearly an interesting physiological biomarker, one important question is what it can tell us about an animal’s physiological state and about life history trade-offs. A growing number of studies have investigated the relationship between blood TL and reproduction in birds and mammals, some providing evidence that individuals that invest more in energetically costly reproductive activities suffer subsequently shortened telomeres. We have used 1982 measurements of TL taken from August blood samples collected from 630 females over 18 years on St. Kilda to test whether raised investment in reproduction in spring predicts shorter TL in summer. Figure 6 describes the setup of our study.
Figure 6. Illustration of different reproductive investment scenarios and their relationship to timing of telomere length measurement in this study. The scenarios are as follows: A) female does not reproduce in a given spring (and is caught and sampled in the following August and has telomere length measured); B) female gives birth to 1 or 2 lambs that die as neonates (female pays cost of gestation but no or little cost of lactation); C) female gives birth to 1 or 2 lambs that survived to weaning age (~4 months; pays both gestation and lactation costs). Line drawing of sheep by Becky Lister-Kaye.

Figure 7. A) Relationship between TL & reproductive investment – Points and boxplot represent raw data (n=1982 observations from 630 Soay ewes) while density plot represents mean and 95% CI from posterior distribution; B) Pairwise contrast plot showing the difference in posterior median estimate between the different reproductive investment categories.

As illustrated in Figure 7, we found that –

a) Mothers whose lambs died shortly after birth had shorter telomeres compared to mothers who did not give birth to lambs.
b) Mothers who successfully weaned their lambs had longer telomeres compared to mothers whose lambs died.
c) Giving birth to twins had no effect on telomere length.

So indeed, we do observe a trade-off between reproduction and telomere length since investing in reproduction resulted in shorter telomeres but only in the case of mothers whose lambs died. Although we predicted that mothers who invest in lactation and successfully wean their lambs will have even shorter telomeres, we actually found the opposite – these mothers who were able to successfully wean their lambs had longer telomeres compared to mothers whose lambs died as neonates. We also found that it did not make a difference whether or not a mother had given birth to twins in terms of effects on telomere length. What could all this mean? Well, it could be that mothers whose lambs die as neonates are (physiologically speaking) low-quality mothers and pay a cost of reproduction that is reflected in their telomere length while mothers who successfully wean their lambs are high-quality females who are better able to tolerate any negative effects of reproduction on telomere length. These high-quality mothers are also usually the ones who give birth to twins which might explain why we don’t observe an effect of twinning on telomere length. Our results suggest the relationship between reproduction and TL, in wild mammals at least, can be complex and that different aspects of reproductive investment and individual physiological condition may influence telomere dynamics in different ways.

**Inbreeding depression across the lifetime.**

Martin A. Stoffel, Susan E. Johnston, Jill G. Pilkington and Josephine M. Pemberton

When related individuals mate, their offspring are likely to have reduced fitness. This phenomenon, termed inbreeding depression, is universal in the animal and plant kingdoms and its origins are genetic: All living things have complex genomes, where mutations happen regularly. Those new mutations are more often than not deleterious. However, fortunately, most of these deleterious mutations are also recessive, which means that they only express their deleterious effects when an individual inherits the same mutation from each parent, i.e. when the individual is homozygous for that mutation. For closely related parents, it becomes much more likely that their offspring inherits the same mutations from both parents, and because many of those will have negative effects, the individual’s fitness is likely to be reduced.

We investigated this phenomenon in Soay sheep, which are an excellent system to study inbreeding depression, because we know the exact pedigree, genetic make-up and fitness for thousands of individuals over more than three decades. We quantified individual inbreeding coefficients on the genetic level using a measure called $F_{ROH}$, which quantifies the proportion of an individuals' genome which is homozygous. Using statistical models, we can then predict the survival probability for each individual given its inbreeding coefficient. In Figure 8, we can see that in Soay sheep lambs a 10% increase in inbreeding coefficient can decrease the probability to survive by more than 20%.
However, the impact of inbreeding also becomes weaker across life. In sheep aged three and older, being inbred does not have a big effect in the survival probability anymore (Fig. 8). According to a theory called the "mutation accumulation theory", inbreeding depression is actually expected to be worse in late life, because many mutations might only show their bad effects later in life. So why do we find the opposite pattern? While we are not 100% sure, it might simply be because many very inbred individuals die early in life, so older individuals are more similar in their inbreeding coefficients and it becomes statistically harder to find differences between them.

Quantifying variation between and within coat colour morphs in Soay sheep.

Mark Sutherland, Jill G. Pilkington, Dylan Childs and Jon Slate.

Colour polymorphisms describe the presence of two or more discrete colour phenotypes, known as morphs, occurring within a population. Although some colour variation can be due to the environment, discrete phenotypes are usually determined by heritable genetic variation. Colour phenotypes are often explained by simple Mendelian genetics, where a few alleles (gene variants) at limited loci (genes) result in the displayed phenotype. This simple genetic architecture and the conspicuous nature of colour polymorphisms makes them an ideal system for studying evolution in wild populations, such as Soay sheep.

Colour polymorphisms can be maintained by direct selection, such as crypsis or mate choice, or by indirect selection when morphs differ in traits other than colour, such as body size and immune function, known as trait covariation. It is often these traits that are subject to selection and manifests as selection on colour itself. Assigning discrete phenotypes to individuals can reveal some of these trait covariations. However, quantifying colour variation within and between morphs may reveal further trait correlations as well as genetic and environmental determinants of colour. Together these can be used to further understanding of how selection acts to maintain polymorphisms.
Soay sheep exhibit a discrete light tawny or dark brown coat colour polymorphism that is known to be controlled by the TYRP1 gene. A single change in the genes code from a G to a T is responsible for the different phenotypes, individuals with two copies of the G allele (GG) or one of each (GT) are dark brown and those with two copies of the T allele (TT) are light tawny. A study by Gratten et al. found that light sheep increased in frequency in the Village Bay population between 1985 and 2005 and suggested positive selection due to increased fitness of both TT (light coats) and GT (dark coats) individuals compared to GG (dark coats). GG and GT individuals are considered to be phenotypically indistinguishable, yet they have different fitnesses. This suggests that selection is not acting on coat colour itself, but on something else that is affected by the TYRP1 gene or a nearby gene (or genes).

Alternatively though, perhaps there are subtle difference in coat colour between dark GG and dark GT sheep, and selection is acting on those differences. To date, there has been no attempt to measure colour variation within the dark and light morphs. Here we look at standardised photography as a method for quantifying colour variation in Soay sheep and if this can improve our understanding of its genetic architecture.

**Methods**

Wool samples from the flank region were obtained from 163 individuals during the 2009 August catch. Qualitative phenotype is recorded for each individual as standard data recording practice of the project; here we had 145 dark and 18 light individuals. Samples were placed on a matt green background and photographs were captured using a Canon EOS 1300D with standardised settings. An X-rite Colour Checker Board was included in each image and used to standardise colour and lighting. Images were processed in Adobe Photoshop CC 2019 to extract red, green, blue (RGB) and hue, saturation, lightness (HSL) values for each wool sample.

![Figure 9. Photographic measurement of coat colour. (1) Wool clippings were taken from the mid flank. (2) Samples placed on a high contrast green background. (3) X-rite ColorChecker® passport is included to generate a unique colour profile. (4) Photograph taken using standardised camera settings. (5) Colour profile applied and white balance standardised. (6) Average RGB (red, green, blue) and HSL (hue, saturation, lightness) recorded for each sample.](image)
All 163 individuals in this study have been genotyped at ~258K genome-wide Single Nucleotide Polymorphism (SNP) markers including the critical TYRP1 gene. A genome wide association study (GWAS) was carried out to investigate associations between coat colour measurements and each SNP locus. This helps us identify the genomic location of loci that may be involved in coat colour.

Results
Principal component analysis (PCA) can be used to transform a large set of variables into a smaller set of principal components (PC) that explain much of the variation in the data. Here, PC1 accounted for 74.1% in colour variation and was significantly different between light TT and dark GT and GG sheep, GT individuals did not differ significantly from GG individuals (Fig. 10). When this is broken down to individual measures we can see that TT individuals differ from the other genotypes in every measure and GT individuals are intermediate between GG and TT individuals for hue, this is the first evidence that the two dark genotypes (GG and GT) may be phenotypically different (Fig. 11).

Figure 10. Principal component analysis of colour measurements showing light TT sheep are significantly different to GT and GG sheep
Figure 11. TYRP1 associations with colour measurements. RGBav, and lightness, measurements were higher in TT individuals compared to both GG and GT individuals, where there was no significant difference (a). RG ratio was lower in TT individuals compared to both GT and GG (b). Hue was highest in TT genotypes and lowest in GG, with GT being intermediate (c). Saturation was lower in TT genotypes compared to GG and GT (d).

An initial GWAS was carried out not controlling for TYRP1 genotype and found significant associations for all colour measures and SNPs surrounding the TYRP1 loci on chromosome 2 (Fig. 12). This is reassuring as it shows that colour measurements are a true representation of colour by identifying the gene known to account for the polymorphism. Subsequent GWAS were carried out controlling for TYRP1 genotype to try and identify further genes that may have a smaller effect on coat colour. GWAS results from PC1 and RGBav (lightness) measurements found two potential loci of interest located on chromosomes 1 and 5. Although they do not reach the genome wide significance threshold, they passed the suggestive value (Fig. 13). This GWAS suffers from a small sample size and it is hoped that these results will be bolstered with future sampling.
Figure 12. GWAS for PC1 not controlling for TYRP1 genotype. The significant association on chromosome 2 maps to the TYRP1 loci.

Figure 13. GWAS of colour measures. RGBav and PC1 found clusters of SNPS on chromosomes 1 and 5 approaching genome wide significance (red line) and passing the suggestive value (blue line)

Summary
These results confirm that the light and dark coat colour is a discrete polymorphism. However there is within-morph variation that is not accounted for by TYRP1 genotype. By using standardised digital photography to quantified coat colour further genomic regions that may influence variation in coat colour have been identified. Further sampling is needed to improve the power of this study and to better understand coat colour selection in Soay sheep.
The genetics of body size using high density genetic markers.

Caelinn James, Sara Knott, Pau Navarro and Josephine Pemberton

Until recently, genetic analyses of the Soay sheep have been carried out using around 39,000 polymorphic single nucleotide polymorphisms (SNPs) which have been genotyped in most study individuals. A further 191 individuals have also been genotyped at even higher density, around 397,000 polymorphic SNPs. Using this new data and the pedigree of Soay sheep, Martin Stoffel (see earlier report) was able to impute all these extra markers in all individuals, to create a high density SNP data set.

The enhanced genetic data led us to focus on two questions. Firstly, what effect does the increased SNP density have on heritability estimates traits such as body size? And secondly, can we recover new SNP-body-size associations when using the high versus low density data?

For both questions we focused on six measures of body size: birth weight, August weight, August hind leg length, August fore leg length, post mortem metacarpal length, and post mortem jaw length. With the exception of birth weight, we looked at each measure in both lambs (less than one year old) and in adults (two years or older).

Heritability estimates

For each trait and age class we ran an animal model to estimate how much variation in each trait in the population is due to additive genetic variance - this is known as the heritability, or $h^2$. Relatedness was estimated using Genomic Relationship Matrix (GRM), which uses the SNP data (low or high density) to estimate how closely related each pair of individuals are. Covariates were fitted to account for variation in the trait caused by non-genetic effects – for example, for jaw length in adults, we also fitted sex, age at death (in years) and year of birth. We then ran each animal model twice – once using the low density genotypes and once using the high density genotypes – in order to compare how the heritability estimate changes between the different SNP densities. We found that for each trait, there was little difference between the heritability estimates when we used the low density genotype data and when we used the high density genotype data, indicating that the low density SNP data adequately captures all the additive genetic variance for these traits.
Heritability estimates for all 6 traits in lambs and adults. Estimates calculated using the low density SNPs are shown in blue, estimates using high density SNPs are shown in orange. Error bars show the standard errors of the estimates.

Genome Wide Association Studies

We then carried out Genome Wide Association Studies (GWAS) for each trait, again first using the low density genotypes and then using the high density genotypes. We fitted the same covariates as we used during the heritability estimation, and to account for relatedness in the population we also fitted the GRM. We found that for some traits, we identified new SNP-trait associations with the high density genotypes that were not significant in the low density genotype data. For example, in birth weight no SNPs were significant using the low density genotypes, but two SNPs on chromosome 7 are significant when using the high density genotypes, showing that the high density data set is indeed giving us more power to detect loci underpinning body size variation.

Manhattan plots showing the p value of the association of each SNP with birth weight using the low density genotype data (left) and the high density genotype data (right). The red line shows the significance threshold, with SNPs above the line being considered significant at the genome-wide level.
Sexual selection on MHC variation in Soay sheep.

Wei Huang, Jill G. Pilkington and Josephine M. Pemberton

The Major Histocompatibility Complex (MHC) is one of the most polymorphic gene clusters in vertebrates and plays an essential role in adaptive immunity. Although pathogen-mediated balancing selection is believed to be the major force shaping MHC diversity, sexual selection may also contribute to the maintenance of MHC diversity. MHC-dependent sexual selection could occur via three mechanisms: selection could favour particular MHC variants, MHC diversity or MHC compatibility between two parents. However, at present there is no consensus as to which of these mechanisms are involved and their importance. Previous studies have often suffered from limited genetic and behavioural data and small sample size, and were rarely able to examine all the mechanisms together, determine whether signatures of MHC-based non-random mating are independent of genomic effects or distinguish whether MHC-dependent sexual selection takes place at the pre- or post-copulatory stage.

In this study, we used a free-living population of Soay sheep (Ovis aries) on the island of Hirta, St Kilda to investigate MHC-dependent sexual selection. Recently, using genotyping-by-sequencing, a total of eight MHC class II variants have been identified in the study population and a large number of individuals alive between 1989 and 2012 have been typed. In addition, genomic pairwise relatedness and a genomic estimate of individual inbreeding is available for most individuals. This data combined with a large number of consort and parentage records enabled us to test MHC-dependent sexual selection more thoroughly than before using Monte Carlo simulations to determine expectations under random mating. If there was any sign of MHC-based mating, we also used generalized linear mixed models (GLMMs) to examine whether this signature was independent of a genomic effect.

We found one of the MHC variants (called C) was disfavoured in comparison with random expectation in both consort males and fathers (Fig. 16a). We found MHC homozygote females were over-represented in consort pairs, but this pattern was not observed in actual mothers (Fig. 16b). We found the average number of shared MHC variants between parents was lower compared with null expectation, but this pattern was not observed in the consort dataset (Fig. 16c). Finally, we found evidence of inbreeding avoidance, as the mean pairwise genomic relatedness in the consort and observed parentage datasets were significantly lower than expected under random mating (Fig. 16d). When fitting MHC and genomic effects in the same model of consort or parentage, we could not demonstrate an independent effect of disassortative mating based on MHC variant sharing, but we found the deviation towards MHC homozygote females in the consort data was independent of genome-wide heterozygosity. Our results suggest that multiple mechanisms of MHC-dependent sexual selection could act simultaneously in Soay sheep and that it is necessary to have an exhaustive examination of all possible mechanisms when investigating MHC-dependent sexual selection.
Figure 16. Results of the Monte Carlo simulation. Histograms represent the result of random simulations with dotted black lines representing the critical p-values. The red and blue dashed blue lines show the observed value in the consort and observed parentage dataset respectively. (a) Males carrying variant (haplotype) C are rarer than expected in both the consort and parentage dataset. (b) MHC homozygote females are commoner than expected in the consort but not the parentage dataset. (c) Parents share fewer MHC variants than expected. (d) Partners are less related than expected in both the consort and the parentage datasets.
PUBLICATIONS


And one paper about St Kilda mice from Tom Black’s project:

ACKNOWLEDGEMENTS

We are very grateful to the National Trust for Scotland for permission to work on St Kilda, and for their assistance in many aspects of the work. The project would not be possible without the generous assistance and support of MOD, QinetiQ and Elior staff stationed on St Kilda and Benbecula and servicing the island. We are particularly grateful to Susan Bain, the Western Isles Manager for the NTS.

We are also grateful for the help of our one and only volunteer for the year Christian Roots, without whom the fieldwork for 2020 would not have been possible. Thank you.

Our research is supported by grants, fellowships and studentships from the Natural Environment Research Council, the Biotechnology and Biological Sciences Research Council, the European Research Council, the Royal Society and the Leverhulme Trust.

APPENDIX A: PERSONNEL NEWS & SCHEDULE OF WORK

Personnel News

Mark Sutherland started his PhD in Sheffield in autumn 2019. During 2020 Elizabeth Mittell joined St Andrews as a Postdoc and Vivienne Litzke started her PhD there, both working with Michael Morrissey.

Schedule of work on St Kilda

Winter - Spring

From the 24th January till the 6th February Jill Pilkington collected repeat samples from many study area females, including every study area female lamb, for pregnancy testing.

From March 3rd until May 28th, Xavier Bal single handedly (rest of lambing team stopped from travelling to St Kilda due to Covid) conducted mortality checks, collected two rounds (March and May) of faecal samples for Ecology Within, carried out two population censuses and monitored and recorded the lambing dates for the whole lambing season. 240 lambs were born to 221 ewes; these figures include 19 definite sets of twins (14 ewes held both lambs, 3 ewe held one lamb and lost one, and 2 ewes lost both lambs). Other dead lambs processed may be twins but this will be determined by genetics at a later date.

Summer

Jill Pilkington, Xavier Bal, Hannah Lemon and Mark Sutherland returned to Hirta on July 29th to carry out ten population censuses, conduct mortality searches (yielding 12 tagged dead animals) and prepare for the main catch-up of study area sheep. The catch-up took place from August 18th – 28th and was conducted by a team of 6 additional project members. 119 sheep were caught and processed, of which 51 were lambs (27 males and 24 females), 12 were
yearlings (3 male and 9 females), 6 were adult males and 50 were adult females. All animals were weighed and measured to monitor growth, and sampled for parasite and genetic analyses. 11 sheep were retagged because of damaged or missing tags. 53 previously untagged lambs, 4 untagged yearlings and 9 untagged adults were caught and processed. In addition, before the catch team arrived and after they left, 18 lambs (15 previously untagged), 8 yearlings (1 previously untagged) and 10 adults were caught and processed. Robin Pakeman collected vegetation data. Xavier Bal and Hannah Lemon remained on Hirta until 16th September to complete mothering-up newly tagged lambs, pasture larvae counts and maintaining equipment used during catch up ready for the next field trip. Jill Pilkington remained on Hirta till the Rut trip due to Covid, in case the team could not return if a lockdown was enforced.

**Autumn**

From October 22nd to December 12th Jill Pilkington, Xavier Bal and Christian Roots carried out ten population censuses, monitored the mating period, capturing and processing 27 incoming tups, 13 resident tups and 13 resident ewes. 26 previously darted, non-resident tups were seen in the study area during this rut. 3 dead sheep were found.