ST. KILDA SOAY SHEEP PROJECT: ANNUAL REPORT 2021

J.G. Pilkington¹, X. Bal¹, M. Bjornson-Sutherland², E.G. Chapman¹, D. Childs², J. Dorrens¹, Y. Corripio-Miyar³, A. Fenton¹⁰, M. Fraser⁸, A. Free¹¹, H. Froy⁹, L. Harrington¹², 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. Mittell⁵, M. Morrissey⁵, P. Navarro¹, D. Nussey¹, R. J. Pakeman⁷, A. Pedersen¹, J.M. Pemberton¹, S. Ravindran¹, J. Slate², L.A. Seeker¹, I.R. Stevenson⁶, M.A. Stoffel¹, A. Sweeny¹, S.L. Underwood¹, H. Vallin⁸, K. Watt¹, E. Wiersma¹, 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.
¹²Institute for Research in Immunology and Cancer, Montreal.

POPULATION OVERVIEW	2
REPORTS ON COMPONENT STUDIES	4
Linking vegetation funtional traits to parasite burdens	4
Logging location and activity in Soay sheep: pilot studies	6
Causes and consequences of early reproduction in male Soay sheep	9
The association between telomere length, parasite burdens and antibody immune response in t	wild
Soay sheep	12
Long-term temporal variation in gastrointestinal parasite infections in St Kilda Soay sheep	14
Can we see a missing fraction in the Soay sheep?	16
Quantifying variation between and within coat colour morphs in Soay sheep	18
Embryonic semi-lethal mutations in the Soay population	22
Partitioning genetic variation into familial and common genetic variance	24
PUBLICATIONS.	27
ACKNOWLEDGEMENTS	28
APPENDIX A: PERSONNEL NEWS & SCHEDULE OF WORK	29

Population Overview

The sheep population on Hirta entered 2021 at a relatively low level and there was moderate mortality over winter with 53 tagged animals being found dead with in the study area. Lambing began on the 4th of April with 76.76% of lambs born surviving (Fig. 1).



Figure 1. The temporal distribution of lamb births during 2021.

In December 2021, 795 tagged sheep were believed to be alive on Hirta, of which 557 regularly used the study area, an increase of 8.34% 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 2021.



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 2132 tagged and untagged sheep, with the details displayed in Table 1. The total population had increased by 54.60% since summer 2020 when it was 1379. This gives a delta (calculated as $\ln (N_{t+1}/N_t)$) of +0.436. 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 19th 2021. Coat colours are DW = dark wild, DS = dark self, LW = light wild, and LS = light self.

Location	Females				Males				Lambs	Total
	DW	DS	LW	LS	DW	DS	LW	LS		
Conachair/Oiseval	201	11	49	2	53	3	18	0	178	515
Mullach Bi / Cambir	315	10	100	3	58	2	19	0	301	808
Ruaival/Village	315	9	76	6	92	0	28	1	282	809
Total	831	30	225	11	203	5	65	1	761	2132

Reports on component studies

Linking Vegetation Functional Traits to Parasite Burdens.

Ellis Wiersma, Robin Pakeman, Jill Pilkington, Dan Nussey and Amy Sweeny.

Functional traits are the characteristics or features of a species or individual that influences the role it plays within an ecosystem. These allow us to generate broader ecological predictions without reference to specific species. The functional traits of vegetation can influence both the habitat and the habitat use of other species, particularly with regards to disease. Certain vegetation traits could favour host immunity (i.e., nutritive value), while others may favour parasite survival or transmission (i.e., microclimate conditions). Variation in the vegetation community can also have different environmental effects depending on an individual's age, and habitat use can also influence host susceptibility to disease. The Soay sheep of St. Kilda are an ideal system to investigate how an individual's habitat may influence their parasite burdens because of the availability of long-term census data, individual parasite data and vegetation data.

Using data on the distribution of plant species and databases of plant traits, we developed maps of the distribution of six main functional traits: canopy height (CanHt), leaf dry matter content (LDMC), leaf area (LeafSize), specific leaf area (SLA), moisture (F), and nitrogen (N). Due to high correlations between LDMC, F, SLA, and LeafSize vegetation traits, these traits were compiled into a composite variable (PC1) representing digestibility using principal component analysis. For each individual sheep-year combination, we calculated a core home range from census data. This was then overlaid onto maps of vegetation functional traits to obtain an average value for each vegetation trait in question for each individual. Hereafter, we associated these trait values with strongyle worm fecal egg counts (FEC), accounting for age, sex, condition, density, and spatial autocorrelation. We used FEC data from 1988 to 2018 from a total of 2698 individuals for 5510 unique individual-year measurements.

Lamb FEC counts were associated with digestibility, with low PC1 values representing greater digestibility (Fig. 5). This means that lambs with higher parasite counts tend to have home ranges containing more palatable vegetation. This relationship could be indicative of density dependent parasite transmission: areas with more palatable vegetation have higher sheep numbers grazing on them and thus greater transmission of nematode parasites. Adult and yearling FEC counts were not associated with any vegetation traits. In all, functional traits are a useful way of incorporating habitat and vegetation into interactions between host and parasite but must be used appropriately. For example, age class was an important factor influencing parasite burdens, suggesting that individuals may interact with their environment differently depending on their age.





Figure 4. Vegetation trait maps of the St. Kilda study area in Village Bay showing variation in trait values where CanHt (canopy height) and F (moisture) are individual vegetation traits and PC1 is a composite variable of LDMC, N, SLA, and LeafSize vegetation traits (positive values representing less digestible vegetation).

0.6

0.5

0.4

0.3

0.2

2 1 0 -1 -2 -3 -4 -5



Figure 5. Effect of vegetation functional trait PC1 on strongyle faecal eggcounts in lambs from the final model output. The functional trait values show a significant decrease in strongyle burdens as trait values increase. Slope: -0.098; 95% confidence interval: (-0.155, -0.045); p = 0.000354.

Logging location and activity in Soay sheep: pilot studies.

Vivienne Litzke and Michael Morrissey.

It is increasingly apparent that variation among sheep in how they use the different parts of the study area can be important. In order to support future studies with higher resolution data on what sheep do in different areas of the study area, we have been developing a biologger module to characterise activity and location of St Kilda Soay sheep.

The logger is composed of two sensors: a GPS unit and an accelerometer, as well as associated hardware for powering and controlling these sensors, and a micro-SD card to save data. The GPS unit provides high precision location data, as well as precise time and date. The accelerometer records motion in three dimensions, which can be used to reconstruct behavioural states (e.g., foraging, walking, resting).

Following initial pilot deployments on domestic sheep in Fife, a small pilot study on St Kilda was undertaken. In July and August of 2021, Vivienne Litzke, Xavier Bal, and Jill Pilkington deployed dataloggers on six adult female Soay sheep and recorded 343 sheep-hours of GPS and accelerometry data. Video was collected of collared sheep for subsequent transcription of activity data. Processing of the resulting data is in progress as part of Vivienne Litzke's PhD studies.



Figure 6. Representative time series of accelerometer data for sheep engaged in contrasting activities. These traces represent the raw data to be used to develop the statistical model to automatically discern activities from accelerometry data without video. All activities were recorded on ewe CR526 on 5^{th} August 2021.



Figure 7. Distribution of average net acceleration during different behaviours of two Soay ewes from pilot studies in 2021. The measure of activity is the mean of the norm of the Euclidean length of the acceleration vector throughout a 3 second activity bout.



Figure 8. Representative GPS tracks for ewe CR526 during the day of 6 August (red) and evening to morning of 6-7 August 2021 (blue). Note that the GPS is interrogated much more frequently during the day. CR526 spent the early morning of the 6^{th} at the bottom of Signals Meadows, reaching the street by about 13:30 UTC. She spent most of the rest of the day above the street and the night near the site of the old water tanks, before returning to the bottom of Signals Meadow in the early morning of the 7^{th} .

Preliminary data analyses reveals substantial promise to recover detailed location (+time) and activity data in St Kilda Soay sheep. Preliminary inspection of accelerometry time-series indicates good concordance with videos (Fig. 6), and simple measures of overall acceleration discern between the main behavioural states (Fig. 7). A measure of total activity level discerns major types of active vs. inactive activities (Fig. 7), and we expect that more detailed models of movement patterns will allow further resolution of specific behaviours.

Similarly, the GPS data reliably recover location and time (Fig. 8). Work is ongoing on the GPS data to investigate how to better filter out improbable location measurements. Nonetheless, it seems clear that the GPS-based location data is of sufficient quality to resolve fine-scale temporal behavioural data with location data on a very fine scale, for example, within the spatial scale at which vegetation communities vary within the Village Bay study area.

Ongoing work will use the joint accelerometry and video data to generate a statistical classification algorithm to automate the assignment of behavioural states from accelerometry data (i.e., ultimately without the need for video). A larger future pilot study will seek to characterise (a) the degree to which individuals differ in behavioural time budgets (b) how the activity budgets of individual sheep relate to their space use, and (c) how individuals vary in their behaviour and space use in response to different weather conditions.

Causes and consequences of early reproduction in male Soay sheep.

Elisabeth Chapman and Josephine Pemberton.

Background

Life history trade-offs in which a favourable change in one trait affects other traits detrimentally, are widespread. For example, animals must decide when to invest in reproduction, potentially risking growth, survival, and future offspring. Generally, natural selection favours males that reproduce quickly, maximising their reproductive output. In contrast, females tend to invest in offspring quality rather than quantity and often reproduce later. The Soay sheep of St Kilda mature quickly, with both sexes able to breed as young as seven months old, in their first rut. Understanding which individuals reproduce at this stage will reveal what aspects of individual and annual variation affect early life reproductive success in the species. We studied early life reproduction in male Soay sheep, asking what individual and year-level factors affect the ability of males to breed in their first year and how these interact. We then investigated the effects of reproducing early on survival and future reproduction.

Factors influencing reproductive success

We used data collected on St Kilda between 1986 and 2019 including individual August body size and condition, population factors and genetic paternity inferences. In this data set, 10.8% of male lambs obtained at least one paternity in their first rut. We ran binomial generalised linear models (GLMs) and followed a process of model simplification to determine which factors were important in predicting reproductive success of male lambs. We found that individuals with large testes and those born into low-density populations were most likely to

breed in their first year (Fig. 9). This is consistent with previous work that has shown young males to be more successful in low-density populations as in these years there are relatively few adult males and so there is less intense competition for females during the rut. The significance of testes size also reflects the importance of sperm competition in the population. Females mate with multiple males throughout the rut so males with larger testes are more likely to sire offspring by producing more sperm, outcompeting other males within the female. We also found that individuals with a twin had lower first year success than singletons (Fig. 9). This may be due to sharing of the mother's resources leading to twins being smaller than singletons. Body weight did not contribute to breeding success when testes size and twin status were in the model, and nor did horn type play a role.



Figure 9. Factors affecting first year reproductive success in males. A shows the effect of population size. B shows the effect of twin status. C shows the effect of testes circumference. Here reproductive success is measured as the proportion of males that had offspring in their first year. Dots represent individuals that either bred successfully (1) or did not (0).

Effect of reproducing on survival

Next we used a second binomial GLM to determine how reproducing in their first rut affected survival of male lambs. Individuals were categorised as either surviving or not surviving their first winter, as this is the period of highest mortality in the population. We found that individuals that reproduced in their first year had higher survival than those that did not breed (Fig. 10). Additionally, heavier individuals were more likely to survive in both high- and low-density years (Fig. 10), though survival was highest in low-density years. This is consistent with the idea that high-quality, heavy individuals are those most likely to reproduce *and* survive the harsh winters.



Figure 10. Factors affecting first year survival in males. **A** shows that individuals that reproduced in their first year had higher survival. **B** shows the interaction between weight and population density, with heavy individuals and those born into low-density years having higher first year survival.

Effect of reproducing on future reproduction

Finally, we looked at how reproducing early affects future reproduction. We looked at the number of offspring surviving males had *after* their first year. This data was zero-inflated, with most males not reproducing again. We ran a Hurdle model to account for this, which is made up of two parts. Firstly, data was analysed with a binomial distribution, categorising individuals as having offspring or having zero offspring following their first year. Following this, the model looked only at individuals that did have offspring and analysed these with a Poisson distribution. We found that reproducing early affected neither presence nor number of future offspring. This contrasts with the expectation that early reproduction might depress future reproduction.

Conclusion

Our results show that early reproductive success is driven by individual variation and environmental conditions, with high quality individuals being most successful and low-density populations facilitating reproductive success of young males. Early reproduction was not associated with reduced survival and had no impact on later-life reproductive success on surviving individuals. Any trade-off between early breeding and later survival and breeding appears to be masked by differences in individual quality.

The association between telomere length, parasite burdens and antibody immune response in wild Soay sheep.

Sanjana Ravindran, Zhulin Chen, Hannah Froy, Sarah Underwood, Jennifer Dorrens, Luise Seeker, Kathryn Watt, Rachael Wilbourn, Jill Pilkington, Lea Harrington, Josephine Pemberton and Daniel Nussey.

Repetitive sequences of DNA called telomeres are present at the end of our chromosomes to protect these ends from degrading - similar to the plastic tips that prevent shoelaces from fraying. Telomeres tend to get shorter after every cell replication cycle and over time, cells with telomeres that get too short 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 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, parasite burdens and immunity in birds and some mammals, providing some evidence that individuals who have high parasite burdens suffer subsequently shortened telomeres. We have used measurements of TL taken from August blood samples collected from both male and female lambs and adults over 18 years of the study period on St. Kilda to investigate the relationship between TL, strongyle worm parasite burdens and antibody immune response. Figure 11 describes the potential predicted associations between these different parameters.



Figure 11. Illustration of different potential hypotheses underlying the relationship between telomere length measurement, parasite burdens and immune response. Being infected with parasites may negatively impact TLs ("cost of infection" hypothesis) and mounting an immune response may be energetically and oxidatively costly also resulting in shortened TLs ("cost of immunity" hypothesis). In addition, broad ageing-related processes could lead to shortened TL which could, in turn, result in immunosenescence ("immunosenescence" hypothesis).

As depicted in Figures 12 and 13, we found that –

- a) No association between TL and strongyle parasite burdens in lambs
- b) No association between TL and immune response (IgG antibody levels) in lambs
- c) Positive association between TL and strongyle parasite burdens in adults
- d) No association between TL and immune response (IgG antibody levels) in adults

We found no support for the 'cost of parasitism' and 'cost of immune response' hypotheses in lambs finding no association of TL with strongyle parasite burdens or IgG antibody levels (antibody immune response), after accounting for variation in summer bodyweight in lambs (Fig. 12). We also found no evidence for any sex-specific costs of parasitism and immunity on TL in lambs. In adults, only a weak, positive association between TL and parasite burdens was found at the within-individual level with no support for the 'immunosenescence' hypothesis as no association between TL and IgG antibody immune response was found (Fig. 13). The lack of any relationship between TL and antibody immune response levels in both lambs and adults suggests that immune cell TL and antibody immune response likely operate under different, independent mechanisms. The positive covariance between TL and strongyle worm burdens in adults could potentially be due to covariances with other environment-related variables such as habitat quality or investment in reproduction during spring.

Our results suggest no support for any hypotheses linking TL, parasitism and immunity and that TL is positively correlated with overall strongyle parasite burdens in the Soays. Further studies integrating different molecular processes such as oxidative stress levels, glucocorticoid levels, TL with life-history traits like immunity and survival are needed to ultimately understand the underlying molecular mechanisms linking different life history traits like immunity and survival.



Figure 12. Relationship of TL with *a*) Strongyle FEC and *b*) IgG-Tc antibody levels of Soay sheep lambs (831 and 1103 observations respectively) - Points represent raw data and lines represent the median line (in bold) and 2000 draws from the posterior distribution



Figure 13. Correlations of telomere length with strongyle FEC and IgG-Tc (n=1320 observations from 565 individuals) at different hierarchical levels in Soay sheep adults (aged 3 years and above). Correlations were estimated as the mode of the posterior distribution with 95% HPDIs from multivariate Bayesian mixed-effects models. Estimates for among-individual, among-year and residual correlations displayed.

Long-term temporal variation in gastrointestinal parasite infections in St Kilda Soay sheep.

Adam Hayward, Dylan Childs, Andy Fenton, Mariecia Fraser, Fiona Kenyon, Tom McNeilly, Dan Nussey, Robin Pakeman, Amy Pedersen, Josephine Pemberton, Jill Pilkington and Amy Sweeny.

Gastrointestinal parasites, including nematode worms and single-celled coccidia of the genus *Eimeria*, are important parasites of Soay sheep on St Kilda and domestic sheep on the mainland. These parasites are associated with intestinal pathology, weight loss, and even death. The dynamics of these parasites, both in sheep and other hosts, are dynamic across the seasons, with factors such as weather and host/parasite biology thought to be important in determining this. Across years, however, the picture is more one of stability: in studies of wild rodents, for example, the prevalence (proportion of animals infected) and abundance (average number of parasites within a host) are remarkably stable across periods of ten years. Longer-term data sets spanning up to and over 20 years are, however, very rare. Examining long-term temporal changes in parasite dynamics is important, both for understanding the factors that regulate host-parasites of livestock. Mathematical modelling, for example, predicts an increase in the transmission of nematode parasites in domestic sheep in the future as the climate warms, but there are no long-term data to test this prediction with. The >20 years of parasitology data collected on St Kilda Soay sheep is an excellent opportunity to explore how parasite infections

have changed over time and to examine the drivers of these changes. Here, we address the first task, describing change over time.

Our recent analyses suggest that there has been a linear increase in the prevalence of both strongyles and coccidia across time, as well as a linear increase in the abundance of coccidia. More complex models, however, suggested that the patterns were better described by both non-linear effects and interactions with age class; thus, although linear changes were present, they did not explain the data better than more complex scenarios.

For example, prevalence of strongyles changed in a non-linear fashion that was dependent on age (Fig. 14A, below): in lambs (green) there was more of a directional increase, while in adults (purple), the increase across the 1990s has levelled off. Meanwhile, the prevalence of coccidia (Fig. 14B, below) has followed a non-linear pattern, potentially driven by the low prevalence in 2005.



Figure 14. (A) Prevalence of strongyles changed in a non-linear fashion that differed between lambs (green) and adults (purple); (B) prevalence of coccidia followed a non-linear pattern.

Meanwhile, there is evidence to suggest that the changes in abundance have been entirely driven by lambs: for both strongyles (Fig.15 left below) and coccidia (Fig.15 right below), abundance has increase in lambs (green), but not in adults (purple). The increase is especially striking for the coccidia.



Figure 15. In both strongyles and coccidia, abundance increased across the study period in lambs (green) but not adults (purple), an increase which is especially striking in the coccidia.

Further analyses will refine these results, exploring the differences among age class in more detail and exploring more nuanced non-linear functions. Future work will aim to determine the potential drivers of temporal variation, including climate and variation in patterns of selection.

Can we see a missing fraction in the Soay sheep?

Elizabeth Mittell and Michael Morrissey.

It is fairly common to see a discrepancy between our observations and our predictions of microevolutionary change in wild populations. For example, despite estimates of selection (S) and heritability (h^2) of adult body weight in the Soay sheep suggesting there should be a positive response to selection and an increase in size, the sheep appear to be getting smaller. Two previous studies have suggested there is no genetic change in body weight. It is therefore an open question as to why adult body weight might not be evolving in the Soay sheep.

One potential explanation for evolutionary stasis is known as the missing fraction problem. The missing fraction refers to the *fraction* of the population that is *missing* for a trait (e.g., adult body size) when we measure it because individuals died before it was expressed. This is only a *problem* if the missing fraction are missing not at random. Missing not at random means that the fraction of the population that is missing when we make measurements *disproportionately* contains individuals that would have been similar to each other. For example, when the missing fraction contains mostly individuals that would have become large adults, fewer large adults will be seen when measuring body size compared to if all individuals had survived. If we can understand and estimate the missing fraction, we might improve both our predictions and understanding of microevolutionary change in wild populations.

The two key factors that influence how big the missing fraction might be are: (a) the strength of viability selection on a correlated trait(s); and (b) the strength of that correlation. The dilemma for this kind of problem is: how can we measure adult traits (and therefore any correlations between traits) for individuals that never became adults? And therefore, how can

we estimate the missing fraction? This is a question that we have been investigating in the Soay sheep.

Inference of genetic selection differentials

Our method for investigating the missing fraction uses information on adult traits, the knowledge that individuals existed (i.e. who died young/survived to become adults), and, who they were related to. This method relies on genetic information but does not require us to know why individuals died before becoming adults, or what they would have been like had they lived to adulthood. We run bivariate animal models to obtain estimates of the genetic covariance between first year survival and adult traits, which we can then use to estimate genetic selection differentials. These genetic selection differentials can be used to improve our predictions of micro-evolutionary change in adult traits.

Here we found a negative genetic covariance between first year survival and adult body weight (-0.082 [-0.025; -0.106 95%CI]). This suggests that adult body weight could be constrained by one or more traits that are being selected early in life that are correlated with adult body weight: i.e. correlated early-life trait(s) are the reason why individuals who would be larger adults are less likely to survive to adulthood and are not seen in the population as adults.

From these data we obtain a genetic selection differential of -0.44 kg. Our current prediction of adult body weight based solely on measurements of adult body weight, and selection on it, made later-in-life is +0.57 kg (using a relationship called the univariate breeder's equation, UVBE). By combining the estimate of selection later-in-life (UVBE = +0.57 kg) and early-in-life (-0.44 kg), we go from predicting an increase in size through time (+0.57 kg), to little change in adult body weight (+0.13 kg;). This is closer to the no evolutionary change found by previous authors, but it also adds some insight into the potential cause. Our result suggests that viability selection is strongest on juveniles that would go on to become the largest in the population, which is resulting in a missing fraction of larger adults (and therefore an incorrect estimation of the phenotypic response to selection when it is not considered).

This example illustrates how if we are not seeing the micro-evolutionary change predicted in wild populations, we can go some way to get an understanding of what the underlying cause may be by including genetic selection differentials in our predictions.

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

Mark Bjornson-Sutherland, Dylan Childs and Jon Slate.

Soay sheep on St Kilda can be one of two colours, dark brown (dark) or light tawny (light) (Fig. 16). When two or more discrete colour forms, or morphs, are found in a wild population this is known as a colour polymorphism. These are determined by heritable genetic variation and are often explained by simple Mendelian genetics, where a few alleles (gene variants) at limited loci (genes) result in the displayed colour (phenotype). For a polymorphism to be maintained in a population they must be subject to some form of balancing selection or one morph would be lost to genetic drift or directional selection. Although selection on colour phenotypes can occur by direct selection, such as crypsis or mate choice, it can also occur through indirect selection when morphs differ in traits other than colour, such as body size and

immune function. It is often these traits that are subject to selection and manifests as selection on colour itself.



Figure 16. Light and dark coat colour polymorphism in Soay Sheep. The TYRP1 gene controls coat colour. GG and GT genotypes are dark, TT genotype is light.

The coat colour polymorphism in Soay sheep is controlled by 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 (Fig. 16). Despite GG and GT individuals displaying identical phenotypes, a previous study by Gratten et al. found differences in other traits between the two, as well as TT individuals, suggesting 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). The T allele and TT light sheep had increased in the population between 1985 and 2005, however recent projection models suggest that selection for the T allele will have reversed in favour of the G allele as population density in the study area increases. Selective pressures differ between sexes and at different life stages in Soay sheep, e.g. winter survival is lowest in lambs and a few large dominant rams sire a large proportion of offspring produced each year. Selection on TYRP1 may be sex and/or age dependent, something not addressed in the previous study.

This study utilises an extra 15 years of data to investigate if the T allele is still increasing or if selection has reversed, as predicted by projection models. It also looks to determine correlations between genotype and a range of morphological and fitness traits in males and females separately and across age classes.

Morphological and fitness traits

Animal models were run for each trait (Table 2), with separate models run for each sex and age category.

Table 2. Traits tested in animal models for correlation with TYRP1. All models were fitted separately for males and females and for the age categories in brackets.

TRAITS

Weight (new-borns, lambs, yearlings, adults,)
Hind leg length (lambs, yearlings, adults)
Survival (first year, adults)
Annual recruits
Lifetime Reproductive Success

Weight differences were seen in males only, with TT homozygotes lighter in weight than both GT and GG genotypes at birth and in adulthood (Fig. 17), this trend was maintained but nonsignificant in lambs and yearlings. Hind leg length showed a very similar relationship with TYRP1 genotype as weight, indicating this is an overall size difference rather than condition dependent weight in males.



Figure 17. TYRP1 genotypes differ in a) Birthweight, b) adult weight, c) annual recruitment and d) annual survival. All relationships are shown relative to GG homozygotes.

Annual survival in adults was lower in TT males compared to GT and GG males (Fig. 17) and this effect was reduced, but remained significant, after accounting for weight; suggesting survival differences are only partially explained by weight. There was no effect of TYRP1 genotype on annual survival in adult females. First year survival was higher in TT homozygotes in both males and females than GG homozygotes; GT heterozygotes appear intermediate but not statistically so (Fig. 18).

Annual recruits is a measure of the number of offspring an individual produces each year that go on to survive their first year. In males, TT homozygotes had the lowest likelihood of surviving to breed and, if they do, the fewest number of surviving offspring (recruits). GT heterozygotes were intermediate and GG homozygotes had the greatest likelihood of recruiting and largest number of recruits (Fig. 17). In females there was no effect of TYRP1 genotype on recruitment. Lifetime reproductive success (LRS) was lower in male TT homozygotes than the other genotypes. There was no difference in LRS in females.



Figure 18. Light lambs (TT) have improved first year survival in both males (left) and females (right). Relationships are shown relative to GG homozygotes.

Temporal and age class changes in TYRP1 allele frequencies

A method known as "gene-dropping" was used to test if changes in TYRP1 allele frequencies over time suggested selection for one allele or were more likely due to random chance. Essentially, this method uses the actual family tree (pedigree) to generate random, but feasible, changes in allele frequencies specific to the population. If the actual change in TYRP1 alleles sits in the 5% most extreme of these random values then it is strongly indicative of selection. The G allele increased in frequency in the population between 1985 and 2020, however this was not more than we would expect by chance (6.5%) (Fig. 19).



Figure 19. The G allele has increased since 1985 (red line), however this is not more than may happen by chance (shaded black area).

To test if TYRP1 allele frequencies varied between age classes a permutation test (x10, 000) was used to randomly assign alleles to individuals, maintaining the observed overall frequencies. Individuals maintained normal life history traits other than TYRP1 genotype and frequency of alleles was recorded at each age class (age in years). Counts were made for males and females separately and tested for ages 0-1 and 2+, as the fitness models suggested different alleles may be favoured in lambs and adults. As suggested by first year survival models, the G allele decreased between ages 0 and 1 in both males and females more than we would expect (Fig. 20). For adult males there was an increase in the G allele from ~45% in 2 year olds to over 60% in older age classes, although this change was not greater than could occur randomly. In females there was a more modest increase from ~ 45% to ~50% that did not differ greatly from the random samples.



Figure 20. The G allele decreases in frequency between 0 and 1 in both males and females. Although the G allele then increases after age 2, this is not more than might happen by chance.

Summary

This study shows the importance of considering selective pressure differences between sexes and ages. It finds fitness differences between coat colour morphs and TYRP1 genotypes that are sex and age dependent. While there was no differences in adult females, adult TT light males appear to have a significant fitness cost. Light coloured males are smaller at birth and in adulthood, have lower LRS, annual recruitment and adult survival. Contrary to this, first year survival is higher in light coloured males and females. First year survival is generally low and therefore may be under strong selection, potentially balancing the fitness cost of being light coloured in adult males. Although the mechanism underlying fitness differences between colour morphs remain uncertain, it is clear they differ in multiple traits that help maintain this polymorphism.

Embryonic semi-lethal mutations in the Soay population.

Martin Stoffel, Susan Johnston, Jill Pilkington and Josephine Pemberton

The replication of DNA is not perfect, causing all living species to accumulate deleterious mutations over time. It has been shown that most of these deleterious mutations are recessive, so that they only express their full effects when an individual inherits two copies of the same mutation from both mother and father, causing the genotype to be homozygous. These mutations can even express their detrimental effects during embryonic development and possibly prevent an individual from being born. However, we would generally expect that those large effect mutations are rare, as natural selection would quickly remove them from the population. In this project, we used detailed genomic data to scan the genome of wild Soay sheep for embryonic lethal mutations.

We scanned the genomes of mother-father-offspring triplets. If a certain genomic region contains an embryonic deleterious mutation, we would expect living offspring to generally be heterozygous in this region, as homozygous individuals would have a high embryonic mortality. Indeed, we identified three regions in the genome where we observe substantially fewer homozygous individuals than expected by chance (Figure 1A). In these regions, 27%-47% fewer homozygous offspring are born (Figure 1B), suggesting the presence of semi-lethal deleterious mutations reducing embryonic survival.



Change in predicted first-year survival

Figure 21: A) A genome-wide scan for embryonic lethal mutations identified three haplotypes (blue points) for which offspring homozygosity is substantially reduced, and which might carry embryonic semi-lethal mutations. B) Expected and observed number of homozygous offspring for each haplotype. C)Frequency of these haplotypes in the Soay population over the last three decades (blue lines) and expectation if haplotypes would change frequency purely due to chance, i.e. genetic drift (grey). C) An analysis of postnatal survival shows that none of the haplotypes is associated with lower survival in lambs.

After we identified the haplotypes carrying semi-lethal embryonic mutations, we can ask whether they change in frequency in the Soay sheep population (Figure 1C). The haplotypes on chromosome 7 and 18 show a downward trend. When we compare these frequency changes to expectations from random genetic drift (Figure 1C, grey lines), we see that such a strong decline in frequency is unlikely due to chance and thereby potentially caused by selection against individuals carrying these haplotypes. In contrast, the haplotype on chromosome 5 seems to have a fairly stable frequency in the population, despite carrying a putative semi-lethal mutation. This is an expected phenomenon in small population where the effects of random genetic drift often outweigh the effects of selection, causing even deleterious mutations to persist in the population.

Lastly, as some living Soay sheep are carrying these haplotypes (i.e. deleterious mutations) in their homozygous state, we explored their effects on first year survival. We found that none of the haplotypes had an impact on survival in lambs (Figure 1D), indicating that the deleterious effects occur mainly in the pre-natal stage. Overall, this analysis links back to previous work on inbreeding, suggesting that embryos from related parents will have a lower probability of survival due to a higher chance of carrying embryonic deleterious mutations in a homozygous state.



Figure 22. A lamb that made it. Photo by Martin Stoffel

Partitioning genetic variation into familial and common genetic variance.

Caelinn James, Sara Knott, Pau Navarro, and Josephine Pemberton.

Animal models can be used to estimate the heritability (h^2) of a trait – that is, how much of the trait variation in a population is due to additive genetic variation. Additive genetic variation is estimated by fitting a genetic effect in the model, and we can fit the genetic effect by using a matrix of relatedness between individuals. Relatedness can be estimated either from a pedigree or from genetic data, however estimates of h^2 often differ between the two methods. Pedigree-based h^2 can be inaccurate due to the fact it cannot calculate the exact proportion of genome sharing between individuals, and it may also include missing or erroneous pedigree links, whilst genetic-based h^2 can miss variance due to ungenotyped variants and variants of a low frequency, such as those that run only in families. The difference between pedigree-based h^2 estimates and genetic-based h^2 estimates is termed the "missing heritability".

In order to recover the missing heritability due to rare variants that run in families, we can run an animal model with a second genetic relatedness matrix (GRM) that only includes relationships above a certain threshold. This method also allows the partitioning of the additive genetic variation into variance caused by common variants in the population and variance caused by rare familial variants. This gives us more information about the underlying genetic architecture of a trait, which is especially useful for traits for which we have yet to find associations with specific variants.

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). For each trait we ran four models; i) a model using the pedigree to estimate relatedness but no thresholded GRM, ii) a model using genetic data to estimate relatedness but no thresholded GRM, iii) a model using genetic data to estimate relatedness and a thresholded GRM including only relationships with a relatedness higher than 0.0625, and iv) a model using genetic data to estimate relatedness higher than 0.1. The thresholds were chosen as they both included all close family relationships (parent-offspring, siblings and half-siblings), but the 0.0625 threshold included a higher number of more distant relationships (e.g. third cousins).

The results varied between traits; for some traits such as lamb hindleg and adult foreleg, we were able to recover some - if not all - of the missing heritability when fitting a secondary thresholded GRM in our models, whilst for other traits such as lamb August weight, the thresholded GRM had no effect (Fig. 23). The difference in results between the different traits is understandable given that there are different underlying genetics influencing the traits, and the models have provided us with information which will be useful in the future when studying these traits further.







Figure 23. Variance explained by the full GRM (blue) and the thresholded GRM (pink) for each trait. Each trait has four models; i) an animal model using the pedigree to estimate relatedness but no thresholded GRM ("Pedigree"), ii) an animal model using genetic data to estimate relatedness but no thresholded GRM ("Full GRM"), iii) an animal model using genetic data to estimate relatedness and a thresholded GRM including only relationships with a relatedness higher than 0.0625 ("0.0625"), and iv) an animal model using genetic data to estimate relatedness with a relatedness higher than 0.1 ("0.1"). Error bars represent standard errors for each variance estimate.

PUBLICATIONS

- Ashraf, B., D. C. Hunter, C. Bérénos, P. A. Ellis, S. E. Johnston, J. G. Pilkington, J. M. Pemberton, and J. Slate. 2021. Genomic prediction in the wild: A case study in Soay sheep. Molecular Ecology <u>https://doi.org/10.1111/mec.16262</u>
- Crawley, M. J., R. J. Pakeman, S. D. Albon, J. G. Pilkington, I. R. Stevenson, M. B. Morrissey, O. R. Jones, E. Allan, A. I. Bento, H. Hipperson, G. Asefa, and J. M. Pemberton. 2021. The dynamics of vegetation grazed by a food-limited population of Soay sheep on St Kilda. Journal of Ecology 109:3988-4006. doi:10.1111/1365-2745.13782
- Dicks, K. L., J. M. Pemberton, K. T. Ballingall, and S. E. Johnston. 2021. MHC class IIa haplotypes derived by high-throughput SNP screening in an isolated sheep population. G3-Genes Genomes Genetics 11:11. doi:10.1093/g3journal/jkab200
- Froy, H., S. L. Underwood, J. Dorrens, L. A. Seeker, K. Watt, R. V. Wilbourn, J. G. Pilkington, L. Harrington, J. M. Pemberton, and D. H. Nussey. 2021. Heritable variation in telomere length predicts mortality in Soay sheep. Proceedings of the National Academy of Sciences of the United States of America 118:8. doi:10.1073/pnas.2020563118
- Huang, W., K. L. Dicks, K. T. Ballingall, S. E. Johnston, A. M. Sparks, K. Watt, J. G. Pilkington, and J. M. Pemberton. 2022. Associations between MHC class II variation and phenotypic traits in a free-living sheep population. Molecular Ecology 31:902-915. doi:10.1111/mec.16265
- Huang, W., K. L. Dicks, J. D. Hadfield, S. E. Johnston, K. T. Ballingall, and J. M. Pemberton. 2022. Contemporary selection on MHC genes in a free-living ruminant population. Ecology Letters:11. doi:10.1111/ele.13957
- Huang, W., and J. M. Pemberton. 2021. Within-trio tests provide little support for postcopulatory selection on major histocompatibility complex haplotypes in a freeliving population. Proceedings of the Royal Society B-Biological Sciences 288:6. doi:10.1098/rspb.2020.2862
- Huang, W., J. G. Pilkington, and J. M. Pemberton. 2021. Patterns of MHC-dependent sexual selection in a free-living population of sheep. Molecular Ecology 30:6733-6742. doi:10.1111/mec.15938
- Paquet, M., J. Knape, D. Arlt, P. Forslund, T. Part, O. Flagstad, C. G. Jones, M. A. C. Nicoll, K. Norris, J. M. Pemberton, H. Sand, L. Svensson, V. Tatayah, P. Wabakken, C. Wikenros, M. Akesson, and M. Low. 2021. Integrated population models poorly estimate the demographic contribution of immigration. Methods in Ecology and Evolution 12:1899-1910. doi:10.1111/2041-210x.13667

- Ravindran, S., H. Froy, S. L. Underwood, J. Dorrens, L. A. Seeker, K. Watt, R. V. Wilbourn, J. G. Pilkington, L. Harrington, J. M. Pemberton, and D. H. Nussey. 2021. The association between female reproductive performance and leukocyte telomere length in wild Soay sheep. Molecular Ecology:13. doi:10.1111/mec.16175
- Sparks, A. M., S. E. Johnston, I. Handel, J. G. Pilkington, J. Berry, J. M. Pemberton, D. H. Nussey, and R. J. Mellanby. 2021. Vitamin D status is heritable and under environment-dependent selection in the wild. Molecular Ecology:15. doi:10.1111/mec.16318
- Stoffel, M. A., S. E. Johnston, J. G. Pilkington, and J. M. Pemberton. 2021. Genetic architecture and lifetime dynamics of inbreeding depression in a wild mammal. Nature Communications 12:10. doi:10.1038/s41467-021-23222-9
- Stoffel, M. A., S. E. Johnston, J. G. Pilkington, and J. M. Pemberton. 2021. Mutation load decreases with haplotype age in wild Soay sheep. Evolution Letters 5:187-195. doi:10.1002/evl3.229

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 volunteers for the year Alasdair Lemon, Joris Philip and Christian Roots, without whom the fieldwork for 2021 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

Elisabeth Chapman and Ellis Wiersma, whose reports appear here, conducted their Edinburgh honours projects on Soay Sheep Project data.

Wei Huang was awarded his PhD on MHC variation in Soay sheep.

Alex Sparks rejoined the Project as a postdoc with Jon Slate in Sheffield.

Schedule of work on St Kilda

Spring

From April 6th until May 26th, Jill Pilkington, Xavier Bal, Hannah Lemon, Martin Stoffel and Amy Sweeny conducted mortality checks, collected two rounds (April and May) of faecal samples for Ecology Within, carried out ten population censuses and monitored and recorded the lambing dates for the whole lambing season. 241 lambs were born to 224 ewes; these figures include 1 set of triplets (one of which survived), 15 definite sets of twins (13 ewes held both lambs and 2 ewes held one lamb and lost one). Other dead lambs processed may be twins but this will be determined by genetics at a later date.

Summer

Jill Pilkington, Xavier Bal and Vivienne Litzke returned to Hirta on July 19^{th} to carry out ten population censuses, conduct mortality searches (yielding 8 tagged dead animals) and prepare for the main catch-up of study area sheep. Six accelerometers were also attached to sheep and then recovered, recording GPS and accelerometry data. The catch-up took place from August $11^{th} - 25^{th}$ and was conducted by a team of 12 additional project members and volunteers. 268 sheep were caught and processed, of which 102 were lambs (49 males and 53 females), 29 were yearlings (7 male and 22 females), 25 were adult males and 112 were adult females. All animals were weighed and measured to monitor growth, and sampled for parasite and genetic analyses. 24 sheep were retagged because of damaged or missing tags. 9 previously untagged lambs, 3 untagged yearlings and 3 untagged adults were caught and processed. Robin Pakeman collected vegetation data. Xavier Bal and Hannah Lemon remained on Hirta until the 8th September and Jill Pilkington until the 13th September to complete mothering-up newly tagged lambs, pasture larvae counts and maintaining equipment used during catch up ready for the next field trip.

Autumn

From October 22nd to December 10th Jill Pilkington, Xavier Bal, Christian Roots and Joris Philip carried out ten population censuses, monitored the mating period, capturing and processing 20 incoming tups, 17 resident tups and 25 resident ewes. 25 previously darted, non-resident tups were seen in the study area during this rut. 7 dead sheep were found.