1 Variants at *RNF212* and *RNF212B* are associated with 2 recombination rate variation in Soay sheep (*Ovis aries*).

3 Susan E. Johnston[1,*], Martin Stoffel[1], Josephine M. Pemberton[1]

4 [1]Institute of Evolutionary Biology, School of Biological Sciences,
5 University of Edinburgh, Edinburgh, EH9 3FL, United Kingdom
6 [*]Corresponding author: Susan.Johnston@ed.ac.uk

7 Abstract

8 [to be written]

Introduction

Meiotic recombination is ubiquitous in sexually-reproducing organisms and is an important driver of genomic diversity in eukaryotes (Barton and Charlesworth, 1998; Otto and Lenormand, 2002). The process of crossing-over ensures the proper disjunction of homologous chromosomes during meiosis (Hassold and Hunt, 2001), with most species having a minimum requirement of at least one obligate crossover per chromosome pair (Otto and Payseur, 2019). It also has the advantage of generating novel haplotypes and preventing the accumulation of harmful mutations in the genome, allowing populations to respond to selection faster and to a greater degree (Muller, 1964; Hill and Robertson, 1966; Battagin *et al.*, 2016). However, this can come at the cost of increased mutations at crossover sites (Halldorsson *et al.*, 2019) and the break up of favourable haplotypes (Barton and Charlesworth, 1998). Variation in recombination rate is likely to arise across taxa with variation in genome size and chromosome number (Stapley *et al.*, 2017); in addition, the evolutionary costs and benefits of recombination may vary depending on the selective context, again leading to the expectation of recombination rate variation within and between populations (Otto and Barton, 2001; Otto and Lenormand, 2002). Indeed, cytogenetic and linkage mapping studies have shown that recombination rates can vary by orders of magnitude within and between chromosomes (Myers *et al.*, 2005), individuals (Kong *et al.*, 2004), populations (Samuk *et al.*, 2020), sexes (Lenormand and Dutheil, 2005) and species (Stapley *et al.*, 2017).

In mammals, genomic studies in different species (i.e. humans, cattle, sheep, pigs and deer) have shown that recombination rate is often heritable, with a significant proportion of the additive genetic variation underpinned by loci of moderate to large effects (Sandor *et al.*, 2012; Kong *et al.*, 2014; Ma *et al.*, 2015; Kadri *et al.*, 2016; Petit *et al.*, 2017; Johnston *et al.*, 2018; Johnsson *et al.*, 2020). Genome-wide association studies (GWAS) have shown repeated associations with variants at ring finger protein 212 *RNF212* and its paralogue *RNF212B*, whose associated proteins are essential for the crossover designation process during meiosis (Reynolds *et al.*, 2013), as well as other loci with established roles in meiotic processes (e.g.

HEI10, *REC8* and *TOP2*; Sandor *et al.* 2012; Kong *et al.* 2014; Ma *et al.* 2015; Kadri *et al.* 2016; Petit *et al.* 2017; Johnston *et al.* 2018; Johnsson *et al.* 2020). This heritable, oligogenic architecture implies that recombination rate has the potential to evolve rapidly to selection if variation in recombination rate is associated with individual fitness (Otto and Barton, 2001; Stapley *et al.*, 2017). Therefore, detailed examination of the genetic architecture of recombination rate is a critical step in understanding how rates are evolving over different evolutionary timescales.

The Soay sheep (*Ovis aries*) are a Neolithic breed of domestic sheep that have been living wild on the St Kilda archipelago since the Bronze age. These sheep are a model system for studies of ecology and evolution, and have been intensively studied since 1985 (Clutton-Brock *et al.*, 2004). A previous study integrated pedigree information with genotypic data from ~39,000 SNPs to characterise autosomal crossover counts (ACCs) in ~3,300 gametes transmitted from parents to their offspring (Johnston *et al.*, 2016). A GWAS showed that a single region on chromosome 6, corresponding to the locus *RNF212*, explained around 47% of heritable variation in female ACC only. A broader-scale "regional heritability" analysis identified a tentative association in both sexes at a 1.09 Mb region on chromosome 7, containing the candidate loci *REC8* and *RNF212B*), but this was not significant in the GWAS or in sex-specific regional heritability models. Since then, addition of more SNP genotyped individuals and improvements in pedigree construction methodology (Huisman, 2017) now allows us to characterise recombination rate in a further ~3,900 gametes. Furthermore, advances in genotype imputation methods in complex pedigrees has allowed us to increase the SNP dataset for conducting genome-wide association study from ~39,000 to ~417,000 SNPs.

Here, we revisit our previous analysis to carry out GWAS of male and female recombination rate. Our motivation for this study was to: (a) confirm a sex-limited association with *RNF212* on chromosome 6; (b) fine-map the association on chromosome 7 and determine the action and effect size of this locus in both males and females; and (c) to identify any further associations that could not be detected using the previous dataset.

Materials and Methods

Study Population.

Soay sheep living within the Village Bay area on the island of Hirta, St Kilda, Scotland (57°49'N, 8°34'W) have been individually studied since 1985 (Clutton-Brock *et al.*, 2004). All sheep are ear-tagged at their first capture, including 95% of lambs born within the study area, and the majority of animals are recaptured on an annual basis. DNA samples are routinely obtained from ear punches and/or blood sampling. All animal work is carried out according to UK Home Office procedures and was licensed under the UK Animals (Scientific Procedures) Act 1986 (License no. PPL60/4211).

Genotype and Pedigree Data.

All sheep were genotyped at 51,135 SNP markers on the Illumina Ovine SNP50 BeadChip (Kijas *et al.*, 2009). Quality control was carried out using the check.marker function in GenABEL v1.8-0 with the following thresholds: SNP minor allele frequency (MAF) > 0.001, SNP locus genotyping success > 0.99, individual sheep genotyping success > 0.95; a total of 7,700 sheep and 39,398 SNPs remained. Pedigree relationships were previously inferred using data from 438 SNP loci in the R package Sequoia v1.02 (Huisman, 2017) and from field observations between mothers and their offspring born within the study area. Autosomal SNP genotypes were used to calculate the $\hat{F}_{III}$ genomic inbreeding coefficients using GCTA v1.90.2 (Yang *et al.*, 2011).

A further 189 individuals were genotyped at 430,702 polymorphic SNPs on the Ovine Infinium HD SNP BeadChip for imputation of genotypes into individuals typed on the 50K chip (see Johnston *et al.* 2016 for method and individual selection criteria and Stoffel *et al.* 2020 for full imputation method for this dataset). Briefly, SNP genotypes from the HD Chip were imputed into the SNP50 Chip typed individuals using AlphaImpute v1.98 (Hickey *et al.*, 2012; Antolín *et al.*, 2017) resulting in a dataset with 7,691 individuals genotyped at 417,373

SNPs, with a mean genotyping rate per individual of 99.5% (range 94.8% - 100%). As there were no X-chromosome SNPs in common between the two SNP Chips, only the SNP50 Chip SNPs were used for association analyses on the X chromosome. SNP positions for both chips were are known relative to sheep genome assembly Oar_v3.1 (GenBank assembly ID GCA_000298735.1).

Estimation of autosomal crossover counts and linkage maps.

Autosomal crossover positions were estimated using an identical protocol to that outlined in Johnston *et al.* (2016). Briefly, the method uses Ovine SNP50 BeadChip SNP data from a focal individual's parents, mate and offspring to characterise crossovers that occurred in the gametes transmitted from the focal individual to its offspring. Crossovers were determined using the software CRI-MAP v2.504a (Green *et al.*, 1990). Here, we estimated ACC for a further 3,908 gametes, leading to a dataset of ACC for 7,235 gametes transmitted from 1,632 unique focal individuals. Because of differences in the SNP quality control between analyses, we used 37,853 SNPs in common with the previous recombination rate study. Simulation studies showed that this method will identify $>99\%$ of crossovers, meaning that ACC estimated using this method can be used as a proxy for individual recombination rate. The method also provides linkage maps for each chromosome, allowing us to update the existing map for Soay sheep to include information from a much larger number of meioses.

Animal Models

A restricted maximum likelihood (REML) animal model approach (Henderson, 1975) was used to estimate components of phenotypic variance for ACC, including the additive genetic effect (i.e. heritability). A genomic relatedness matrix (GRM) was constructed for all polymorphic autosomal SNPs from the 50K chip using GCTA v1.90.2 (Yang *et al.*, 2011) and was adjusted with the argument *--grm-adj 0*, which assumes a similar frequency spectra of genotyped and causal loci. Related individuals were not removed from the GRM as there is high relatedness in the study population, and there are were no significant parental effects

or common environment effects on individual recombination rate (Johnston *et al.*, 2016). Animal models were constructed for both sexes combined, females only and males only, and were run in ASReml-R v4 (Butler *et al.*, 2009) in R v3.6.2. Fixed effects included sex (in the combined model) and the genomic inbreeding coefficient $\hat{F}_{III}$; initial models fit age in years fit as a linear covariate, but as it was not significant, it was removed from the final models (Figure 1). The significance of fixed effects was tested using Wald tests. Random effects included the additive genetic effect (as estimated using the GRM) and identity of the focal individual (i.e. "permanent environment") effect to account for repeated measures. Despite not being significant, the fixed effect $\hat{F}_{III}$ and the random permanent environment effect were retained in all models to account for potential underestimation of ACC due to long runs of homozygosity and pseudoreplication, respectively. Significance of the additive genetic effect was determined by dropping the additive genetic effect from each model and comparing with the full model using a likelihood ratio test distributed as χ^2_1. Bivariate models examining the genetic correlation r_A between male and female ACC were run using the CORGH error structure in ASReml-R (correlation with heterogeneous variances). Models were set with r_A to be unconstrained. To test whether r_A was significantly different from 0 and 1, the model was compared to models with r_A fixed at a value of 0 or 0.999 using likelihood ratio tests.

Genome-wide association studies.

Genome-wide association studies of ACC with the imputed SNP dataset were carried out using function *rGLS* in the package RepeatABEL v1.1.31 (Rönnegård *et al.*, 2016) in R v3.6.2. This approach fits both repeated measures and the GRM, the latter accounting for population structure. Association statistics were corrected for any further inflation by dividing by the genomic control parameter λ, calculated as the median observed χ^2_1 statistic divided by the median χ^2_1 expected under a null distribution. The significance threshold at $\alpha = 0.05$ was calculated by Stoffel *et al.* (2020) as $P < 1.28 \times 10^{-06}$, using a 'simpleM' approach using linkage disequilibrium information to determine the effective number of independent tests (Gao *et al.*, 2008).

For the most highly-associated SNPs, genotype effect sizes were estimated by fitting the SNP genotype as a three level factor in the animal models described above. Then, in separate models, the proportion of phenotypic variance explained by significant regions was then determined using a regional heritability approach. For these models, GRMs were constructed as above for the 20 SNPs from the SNP50 chip spanning the highest association SNP (i.e. 10 SNPs from either side of the association). In one case where the association was at the end of the chromosome, the last 20 SNPs on that chromosome were used to construct the GRM. These regional GRMs were then fit as additional random effects in the animal models described above, with their significance determined using likelihood ratio tests distributed as χ^2_1.

In order to identify potential candidate genes for ACC, sheep gene IDs SNPs in the broad associated regions were extracted from Ensembl (gene build ID Oar_v3.1.100) using the function *getBM* in the R package biomaRt v2.42.1 (Durinck *et al.*, 2009) in R v3.6.2. Gene orthologues in humans (*Homo sapiens*), cattle (*Bos taurus*), mouse (*Mus musculus*) and rat (*Rattus norvegicus*) were extracted using the function *getLDS*. The associated gene ontology (GO annotations) information for all genes and orthologues were then extracted, again using the *getBM* function. All gene names, phenotype descriptions, GO terms and definitions were then queried for terms associated with meiosis and recombination, using the R command *grep* with the text strings *meio* and *recombin*.

Data Availability

Raw data will be publicly archived. Code for the analysis is archived at https://github.com/susjoh/2020_Soay_Recomb_GWAS.

Results

Variation and heritability of autosomal crossover count.

Males had higher recombination rates than females, with 7.24 more autosomal crossovers observed per gamete (SE = 0.18. Wald χ^2_1 = 1613.6, P < 0.001; Figure 1). There was no association with ACC and age or the inbreeding coefficient $\hat{F}_{III}$ (P > 0.05; Figure 1). Females had higher phenotypic variance than males (female V_P = 32.41, male V_P = 24.162; Table 1). ACC was heritable in both sexes (h^2 = 0.148) and in males and females separately (h^2 = 0.118 and 0.181, respectively; Table 1); likelihood ratio tests confirmed that including the additive genetic effect significantly improved the animal models (χ^2_1 = 139.4, 19.72 and 119.9 for both sexes, males, and females, respectively; P < 0.001). The permanent environment contribution to phenotypic variance was not significant in any models (P > 0.05), with the remaining variance attributed to residual effects (Table 1). Bivariate models of ACC between males and females showed a positive genetic correlation (r_A = 0.555, SE = 0.138) that was significantly different from both r_A = 0 and 1 (χ^2_1 = 11.95 and 17.91, respectively; P < 0.001). Linkage map lengths from the ACC analysis in this study were strongly correlated with the previous maps (linear regression, adjusted R^2 >0.999, Figure S1). The updated map had 15,787 unique centiMorgan positions compared with 12,359 in the previous study; it is is provided in Table S2 and Figure S2.

Genome-wide Association Study.

There was a strong association between SNPs at the sub-telomeric region of chromosome 6 and female ACC (*oar3_OAR6_116402578*, P = 2.45×10^{-17}; Figures 2 & S4; Tables 2 & S3). This locus explained 45.9% of the heritable variation in female ACC, with a difference of 5 crossovers per gamete between the two homozygotes (Figure S3, Table 2). There was no significant association with this locus and male ACC, confirming that its action is likely to be sex-limited. The same SNP showed the highest genome-wide association in the previous study

(JOHNSTON *et al.*, 2016) and occurs 25.2 kb from the putative location of *RNF212*. It should be noted that *RNF212* is not annotated on the Oar_v3.1 sheep genome due to a gap in the sequence, but is predicted to be positioned between 116,427,802 and 116,446,904 bp (JOHNSTON *et al.*, 2016); *RNF212* is present in this position on the Oar_rambouillet_v1.0 genome assembly (GenBank assembly ID GCA_002742125.1). A second association was observed in both sexes on chromosome 7 (*oar3_OAR7_21368818* and *oar3_OAR7_21347355*, P = 6.11×10^{-13}, 1.83×10^{-10} and 7.20×10^{-7} in both sexes combined, females and males, respectively); the association statistic in males was also identical for locus *oar3_OAR7_21116299* (Figures 2 & S5; Tables 2 & S3). The action of this locus was similar between males and females, with a difference of ~4.5 crossovers per gamete between the two homozygotes (Figure S3, Table 2) and explaining between 19.7% and 24.8% of the heritable variation in ACC. These SNPs spanned the candidate locus *RNF212B* (21,241,831 to 21,273,165 bp). GO term analysis indicated that both *RNF212* and *RNF212B* were the closest genes to each hit that were associated with meiotic processes (Table S4; Figures S4 & S5).

Discussion

This study revisited a previous analysis investigating the genetic architecture of ACC in Soay sheep, with more than double the number of gametes and more than ten times the number of SNP loci. We confirmed that ACC is ~1.27 times higher in males than in females, and is not associated with age or inbreeding in neither males nor females. We identified two candidate genes associated with ACC: *RNF212* on chromosome 6, which is associated with female recombination rate only; and its paralogue *RNF212B* on chromosome 7, which is associated with recombination rate in both sexes. Both loci have repeatedly been associated with recombination rate variation in mammals, with all studies conducted to date implicating either one or both of these loci; these include humans (Kong *et al.*, 2008, 2014), cattle (Sandor *et al.*, 2012; Ma *et al.*, 2015; Kadri *et al.*, 2016), pigs (Johnsson *et al.*, 2020), deer (Johnston *et al.*, 2018) and domestic sheep (Petit *et al.*, 2017). Functional studies in mice indicate that the protein RNF212 is essential for crossover formation during meiosis, and that it has a dosage sensitive effect (Reynolds *et al.*, 2013). We observe an additive effect of alleles on ACC at both loci in this study, suggesting that a similar dosage sensitive effect is the mechanism driving rate differences in Soay sheep.

The current study confirms the previous association at *RNF212* and does not gain any novel insights into the nature of its association with ACC. However, the observation at *RNF212B* builds on tentative evidence of an association in this genomic region. In the 2016 study, a genome-wide regional heritability analysis identified an association in both sexes combined in a 1.09Mb region, which contained *RNF212B* and another candidate locus, meiotic recombination protein *REC8*. However, there was no corresponding association using GWAS and there was no significant effect observed in the sex-specific regional heritability analyses, providing limited insight into the action of this region on ACC variation. In this study, we have shown that adding a larger number of gametes and SNPs has increased the power to detect an association at *RNF212B*, and indicating that ACC variation is more likely to be attributed to *RNF212B* than *REC8* (although see further discussion below). The MAF of *RNF212B* is relatively low (0.082), suggesting that lower sample sizes in the 2016 study were not sufficient

to detect the effects of rare loci within the population. Using the SNP50 Chip alone, the increased number of gametes in current study would shown a significant on chromosome 7 in females (but not in males; Figure S5), however, the highest association would be around 345Kb away from the that observed in the current study. Our work shows that increasing marker density using imputation is likely to be an increasingly important approach in conducting GWAS in natural populations.

Soay sheep have extensive linkage disequilibrium (LD) throughout the genome due to recent population bottlenecks, a small effective population size and a high prevalence of inbreeding (Clutton-Brock *et al.*, 2004; Stoffel *et al.*, 2020). An advantage of this is that there is a higher change of typing SNP loci that are in LD with causal loci, and that genotype imputation can be carried out with high accuracy. However, a disadvantage is that high LD within the population makes it difficult to separate the effects of linked loci that also contribute to phenotypic variation in the population. The associated regions on both chromosomes 6 and 7 extend over several megabases, with several other loci associated with meiotic processes occurring within these regions (Figures S4 & S5, Table S4); we cannot rule out that these loci also affect variation in ACC in Soay sheep. Functional validation of the role of *RNF212* and *RNF212B* is not possible within this wild system as experimental manipulation and invasive sampling are prohibited. However, further studies in domestic sheep (e.g. Petit *et al.* 2017) and related species may further elucidate the role of loci within these regions and the relative importance of regulatory and protein-coding variation in driving recombination rate variation.

Overall, this study has highlighted the merits of genotype imputation and increasing sample sizes to determine the genetic architecture of recombination rate variation in a natural system. Identifying candidate loci and their sex-specific effect sizes provides a stepping stone for future studies investigating the evolution of recombination rates within this and related systems, through modelling the temporal dynamics of these loci, their origin and their association with individual fitness.

Acknowledgements

We thank Jill Pilkington, Ian Stevenson and all Soay sheep project members and volunteers for collection of data and samples. Camillo Bérénos, Philip Ellis and Hannah Lemon prepared DNA samples for Ovine SNP50 BeadChip genotyping and Jisca Huisman assisted with the pedigree construction. This work has made extensive use of the resources provided by the Edinburgh Compute and Data Facility (http://www.ecdf.ed.ac.uk/). Permission to work on St Kilda is granted by The National Trust for Scotland, and logistical support is provided by QinetiQ and Eurest. The Soay sheep project has been supported by grants from the UK Natural Environment Research Council, with the SNP genotyping funded by a European Research Council Advanced Grant to JMP. SEJ is supported by a Royal Society University Research Fellowship.

Author Contributions

S.E.J and J.M.P. conceived the study. J.M.P organised the collection of samples. M.S. conducted the genotype imputation. S.E.J. analysed the data and wrote the paper. All authors contributed to revisions.

273 Figures and Tables

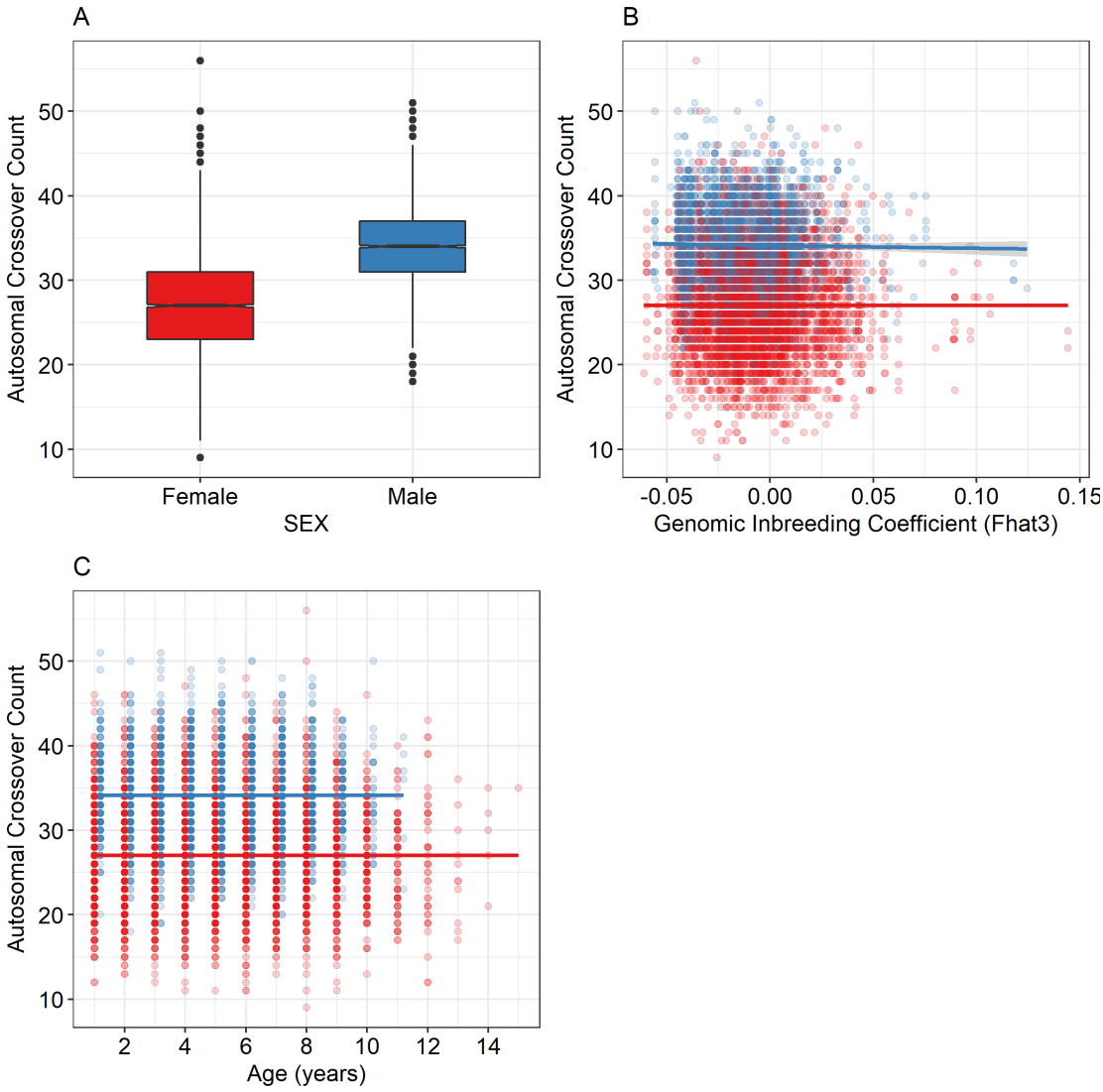

Figure 1: Association between autosomal crossover count and: A. sex; B. inbreeding coefficient $\hat{F}_{III}$; and C. age in years. Plots were made using the raw data. Points are red for females and blue for males. Lines in B. and C. are general additive model smoothing parameters as fit by the plotting library ggplot2 v3.3.2 (Wickham, 2016) in R v3.6.2. Points in plot C are slightly jittered between the sexes on the x-axis to more clearly visualise the sex differences with age. Sample sizes are provided in Table 1.

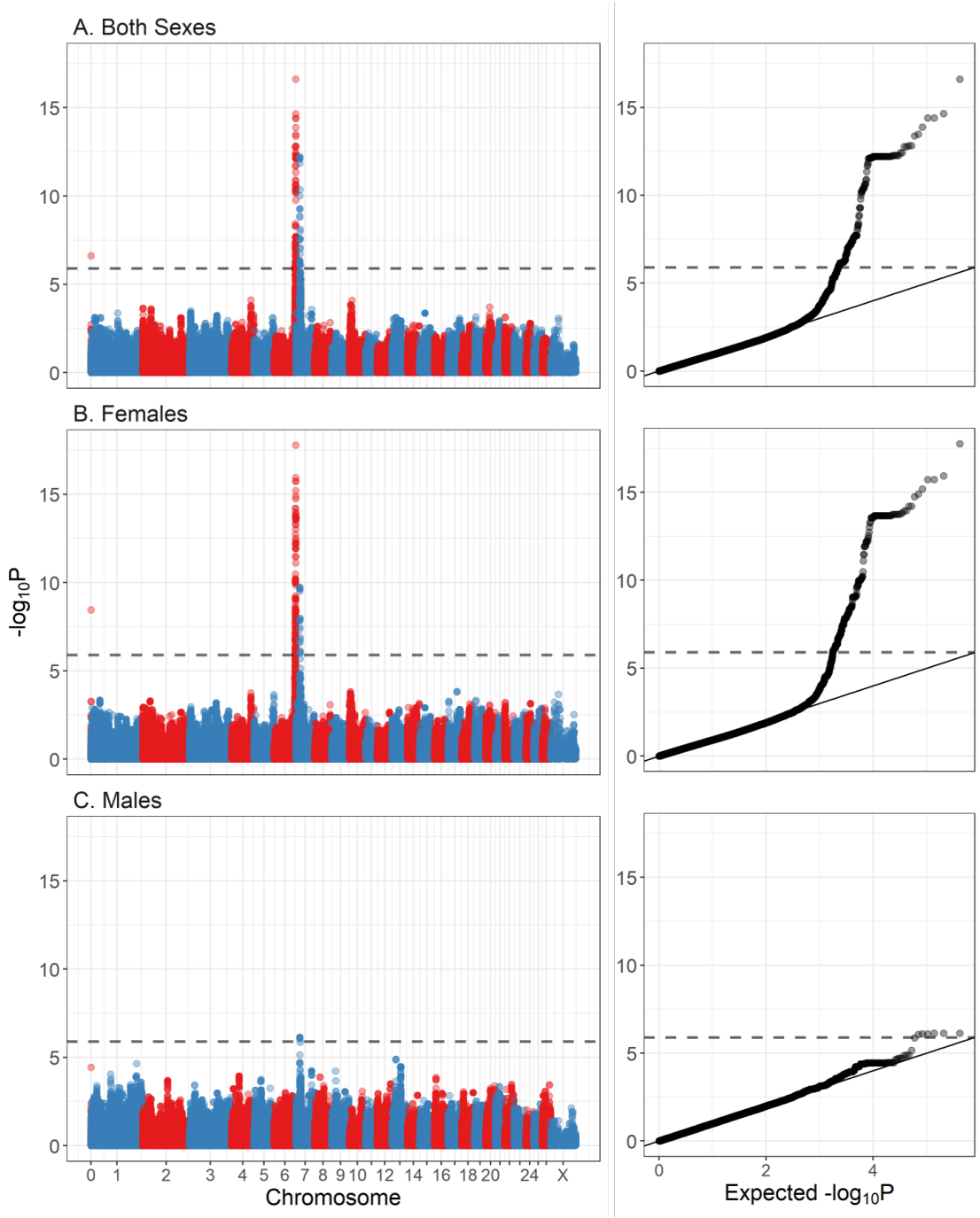

Figure 2: Genome-wide association plots of autosomal crossover counts for A. both sexes combined, B. females only and C. males only. The dashed line is the significant threshold equivalent to $\alpha < 0.05$. The left panels show the association for individual SNPs relative to their genomic position, with points colour-coded by chromosome. The right plots show the distribution of the observed $-log_{10}P$ values against those expected under a null distribution. All association statistics have been corrected for genomic control using the λ parameter. All results are provided in Table S3. Sample sizes are provided in Table 1.

Table 1: Data information and animal model results for autosomal crossover count (ACC) in Soay sheep. Numbers in parentheses are the standard error unless otherwise stated. N_{OBS} is the observed number of gametes, with numbers in parentheses the number of unique focal individuals. N_{xovers} are the total number of crossovers in the dataset. The mean ACC and the variance (V_{OBS}) were calculated from the raw data. V_P is phenotypic variance. h^2, pe^2 and e^2 are the narrow-sense heritability, the permanent environment effect, and the residual effect, respectively; all are calculated as the proportion of V_P.

Analysis	N_{OBS}	N_{xovers}	Mean	V_{OBS}	V_P	h^2	pe^2	e^2
Both Sexes	7626	227646	29.851	41.32	29.514	0.148	0.025	0.827
	(1483)				(0.588)	(0.019)	(0.013)	(0.014)
Males	3043	103914	34.149	23.974	24.162	0.118	0	0.882
	(434)				(0.72)	(0.032)	(0.024)	(0.02)
Females	4583	123732	26.998	32.439	32.41	0.181	0	0.819
	(1049)				(0.797)	(0.024)	(0.017)	(0.017)

Table 2: The most highly associated SNPs from GWAS of ACC in Soay sheep. MAF is the minor allele frequency (corresponding to the A allele for each locus). Sex indicates if the model was run in both sexes combined, or in females or males only. P is the P-value for the association statistic after correction using genomic control λ. Effect sizes are provided for animal models where the genotype fit as a fixed factor, and are given relative to the model intercept of 0 at genotype AA. Prop. V_A is the proportion of the additive genetic variance explained by the genomic region containing the SNP locus. All values in parentheses are standard errors. Full GWAS results are provided in Table S3. Effect sizes and associated sample sizes for each genotype are shown in Figure S3.

SNP Information	MAF	Sex	P	Effect AG	Effect GG	Prop. V_A
oar3_OAR6_116402578	0.275	Both Sexes	2.45×10^{-17}	2.201	3.837	0.399
Chr. 6				(0.332)	(0.361)	(0.118)
Position 116,402,578		Females	1.68×10^{-18}	3.038	4.985	0.459
Candidate Gene: *RNF212*				(0.397)	(0.423)	(0.115)
		Males	0.0317	0.168	0.884	0.049
				(0.577)	(0.607)	(0.067, ns)
oar3_OAR7_21368818*	0.082	Both Sexes	6.11×10^{-13}	-2.383	-4.520	0.242
Chr. 7				(0.922)	(0.952)	(0.102)
Position 21,368,818		Females	1.83×10^{-10}	-2.705	-4.757	0.197
Candidate Gene: *RNF212B*				(1.152)	(1.194)	(0.09)
		Males	7.20×10^{-07}	-2.479	-4.394	0.248
				(1.454)	(1.447)	(0.145)

* This association was identical to hits at SNP loci oar3_OAR7_21347355 at position 21,347,355 (in both sexes, males and females) and oar3_OAR7_21116299 at position 21,116,299 (in males).

References

Antolín, R., C. Nettelblad, G. Gorjanc, D. Money, and J. M. Hickey, 2017 A hybrid method for the imputation of genomic data in livestock populations. Genetics Selection Evolution : 1–17.

Barton, N. H., and B. Charlesworth, 1998 Why sex and recombination? Science **74**: 187–195.

Battagin, M., G. Gorjanc, A.-M. Faux, S. E. Johnston, and J. M. Hickey, 2016 Effect of manipulating recombination rates on response to selection in livestock breeding programs. Genet. Select. Evol. **48**: 44.

Butler, D. G., B. R. Cullis, A. R. Gilmour, and B. J. Gogel, 2009 Mixed Models for S language Environments: ASReml-R reference manual.

Clutton-Brock, T., J. Pemberton, T. N. Coulson, I. R. Stevenson, and A. D. C. MacColl, 2004 The sheep of St Kilda. In *Soay Sheep: Dynamics and Selection in an Island Population.2*. Cambridge University Press, 17–51.

Durinck, S., P. T. Spellman, E. Birney, and W. Huber, 2009 Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. Nature protocols **4**: 1184–1191.

Gao, X., J. Starmer, and E. R. Martin, 2008 A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. Genetic epidemiology **32**: 361–369.

Green, P., K. Falls, and S. Crooks, 1990 *Documentation for CRIMAP, version 2.4*. Washington University School of Medicine.

Halldorsson, B. V., G. Palsson, O. A. Stefansson, H. Jonsson, M. T. Hardarson, *et al.*, 2019 Characterizing mutagenic effects of recombination through a sequence-level genetic map. Science **363**: eaau1043.

299 Hassold, T., and P. Hunt, 2001 To err (meiotically) is human: the genesis of human aneu-
300 ploidy. Nat. Rev. Genet. **2**: 280–291.

301 Henderson, C. R., 1975 Best linear unbiased estimation and prediction under a selection
302 model. Biometrics **31**: 423–447.

303 Hickey, J. M., B. P. Kinghorn, B. Tier, J. H. J. V. D. Werf, and M. A. Cleveland, 2012
304 A phasing and imputation method for pedigreed populations that results in a single-stage
305 genomic evaluation. Genet Sel Evol **44**: 9.

306 Hill, W. G., and a. Robertson, 1966 The effect of linkage on limits to artificial selection.
307 Genet. Res. **8**: 269–294.

308 Huisman, J., 2017 Pedigree reconstruction from SNP data: parentage assignment, sibship
309 clustering and beyond. Mol. Ecol. Resour. **17**: 1009–1024.

310 Johnsson, M., A. Whalen, R. Ros-Freixedes, G. Gorjanc, C.-Y. Chen, *et al.*, 2020 Genetics
311 of recombination rate variation in the pig. bioRxiv .

312 Johnston, S. E., C. Bérénos, J. Slate, and J. M. Pemberton, 2016 Conserved genetic ar-
313 chitecture underlying individual recombination rate variation in a wild population of soay
314 sheep (Ovis aries). Genetics **203**: 583–598.

315 Johnston, S. E., J. Huisman, and J. M. Pemberton, 2018 A genomic region containing
316 *REC8* and *RNF212B* is associated with individual recombination rate variation in a wild
317 population of red deer (*Cervus elaphus*). BioRxiv .

318 Kadri, N. K., C. Harland, P. Faux, N. Cambisano, L. Karim, *et al.*, 2016 Coding and
319 noncoding variants in HFM1, MLH3, MSH4, MSH5, RNF212, and RNF212B affect re-
320 combination rate in cattle. Genome research **26**: 1323–1332.

321 Kijas, J. W., D. Townley, B. P. Dalrymple, M. P. Heaton, J. F. Maddox, *et al.*, 2009 A
322 genome wide survey of SNP variation reveals the genetic structure of sheep breeds. PLoS
323 ONE **4**: e4668.

324 Kong, A., J. Barnard, D. F. Gudbjartsson, G. Thorleifsson, G. Jonsdottir, *et al.*, 2004
325 Recombination rate and reproductive success in humans. Nat. Genet. **36**: 1203–1206.

326 Kong, A., G. Thorleifsson, M. L. Frigge, G. Masson, D. F. Gudbjartsson, *et al.*, 2014
327 Common and low-frequency variants associated with genome-wide recombination rate.
328 Nat. Genet. **46**: 11–6.

329 Kong, A., G. Thorleifsson, H. Stefansson, G. Masson, A. Helgason, *et al.*, 2008 Se-
330 quence variants in the RNF212 gene associate with genome-wide recombination rate. Sci-
331 ence **319**: 1398–1401.

332 Lenormand, T., and J. Dutheil, 2005 Recombination difference between sexes: a role for
333 haploid selection. PLoS biology **3**: e63.

334 Ma, L., J. R. O'Connell, P. M. VanRaden, B. Shen, A. Padhi, *et al.*, 2015 Cattle Sex-Specific
335 Recombination and Genetic Control from a Large Pedigree Analysis. PLoS Genet. **11**:
336 e1005387.

337 Muller, H., 1964 The relation of recombination to mutational advance. Mutat. Res. **1**: 2–9.

338 Myers, S., L. Bottolo, C. Freeman, G. McVean, and P. Donnelly, 2005 A fine-scale map
339 of recombination rates and hotspots across the human genome. Science **310**: 321–4.

340 Otto, S. P., and N. H. Barton, 2001 Selection for recombination in small populations.
341 Evolution **55**: 1921–1931.

342 Otto, S. P., and T. Lenormand, 2002 Resolving the paradox of sex and recombination. Nat.
343 Rev. Genet. **3**: 252–261.

344 Otto, S. P., and B. A. Payseur, 2019 Crossover Interference: Shedding Light on the Evolu-
345 tion of Recombination. Annual review of genetics **53**: 19–44.

346 Petit, M., J.-M. Astruc, J. Sarry, L. Drouilhet, S. Fabre, *et al.*, 2017 Variation in Recom-
347 bination Rate and Its Genetic Determinism in Sheep Populations. Genetics **207**: 767–784.

348 Reynolds, A., H. Qiao, Y. Yang, J. K. Chen, N. Jackson, *et al.*, 2013 RNF212 is a dosage-
349 sensitive regulator of crossing-over during mammalian meiosis. Nat. Genet. **45**: 269–78.

350 Rönnegård, L., S. E. McFarlane, A. Husby, T. Kawakami, H. Ellegren, *et al.*, 2016 Increas-
351 ing the power of genome wide association studies in natural populations using repeated
352 measures - evaluation and implementation. Methods in Ecology and Evolution **7**: 792–799.

353 Samuk, K., B. Manzano-Winkler, K. R. Ritz, and M. A. F. Noor, 2020 Natural Selection
354 Shapes Variation in Genome-wide Recombination Rate in Drosophila pseudoob-
355 scura. Current Biology **30**: 1517–1528.e6.

356 Sandor, C., W. Li, W. Coppieters, T. Druet, C. Charlier, *et al.*, 2012 Genetic variants in
357 *REC8*, *RNF212*, and *PRDM9* influence male recombination in cattle. PLoS Genet. **8**: 1–13.

358 Stapley, J., P. G. D. Feulner, S. E. Johnston, A. W. Santure, and C. M. Smadja, 2017
359 Variation in recombination frequency and distribution across Eukaryotes: patterns and
360 processes. Philos. Trans. R. Soc. Lond., B, Biol. Sci. **372**: 20160455.

361 Stoffel, M., S. Johnston, J. Pilkington, and J. Pemberton, 2020 Genetic architecture and
362 lifetime dynamics of inbreeding depression in a wild mammal. bioRxiv .

363 Wickham, H., 2016 *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

364 Yang, J., S. H. Lee, M. E. Goddard, and P. M. Visscher, 2011 GCTA: a tool for genome-
365 wide complex trait analysis. Am. J. Hum. Genet. **88**: 76–82.

Supplementary Information

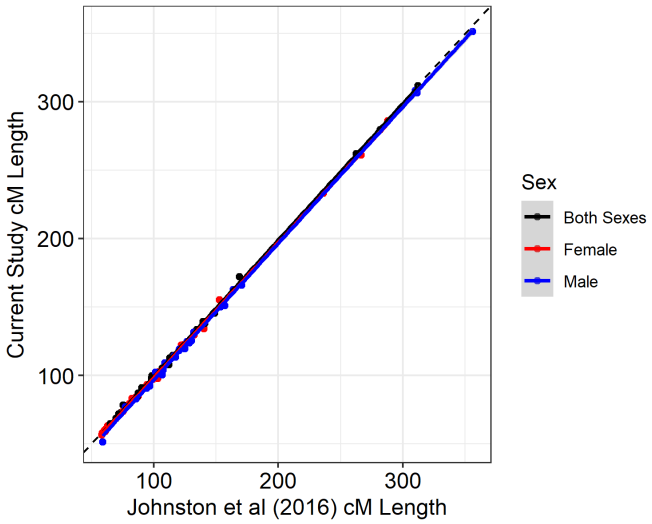

Figure S1: Correlation between linkage map lengths determined in the previous study (JOHNSTON *et al.*, 2016) and in the current study. Each point indicates a chromosome, and lines indicate linear regressions. Points are coloured by sex-averaged or sex-specific map types. The black dashed line indicates a perfect correlation with slope = 1 and intercept = 0.

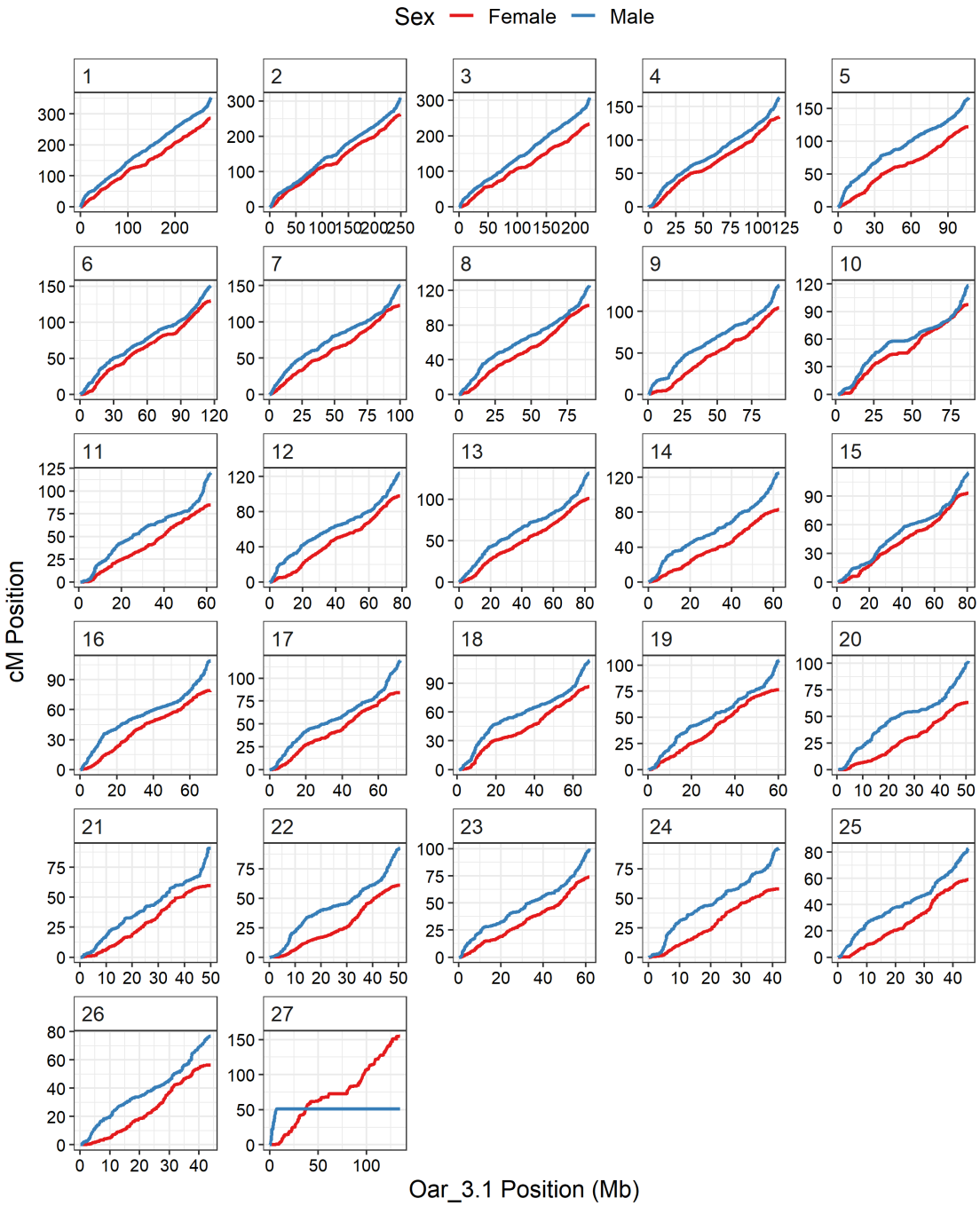

Figure S2: Sex-specific linkage maps for each sheep chromosome plotted relative to the sheep genome assembly Oar_v3.1. The underlying map data is provided in Table S2. Chromosome 27 is the X chromosome, with the short map segment in male sheep corresponding to the pseudoautosomal region (PAR).

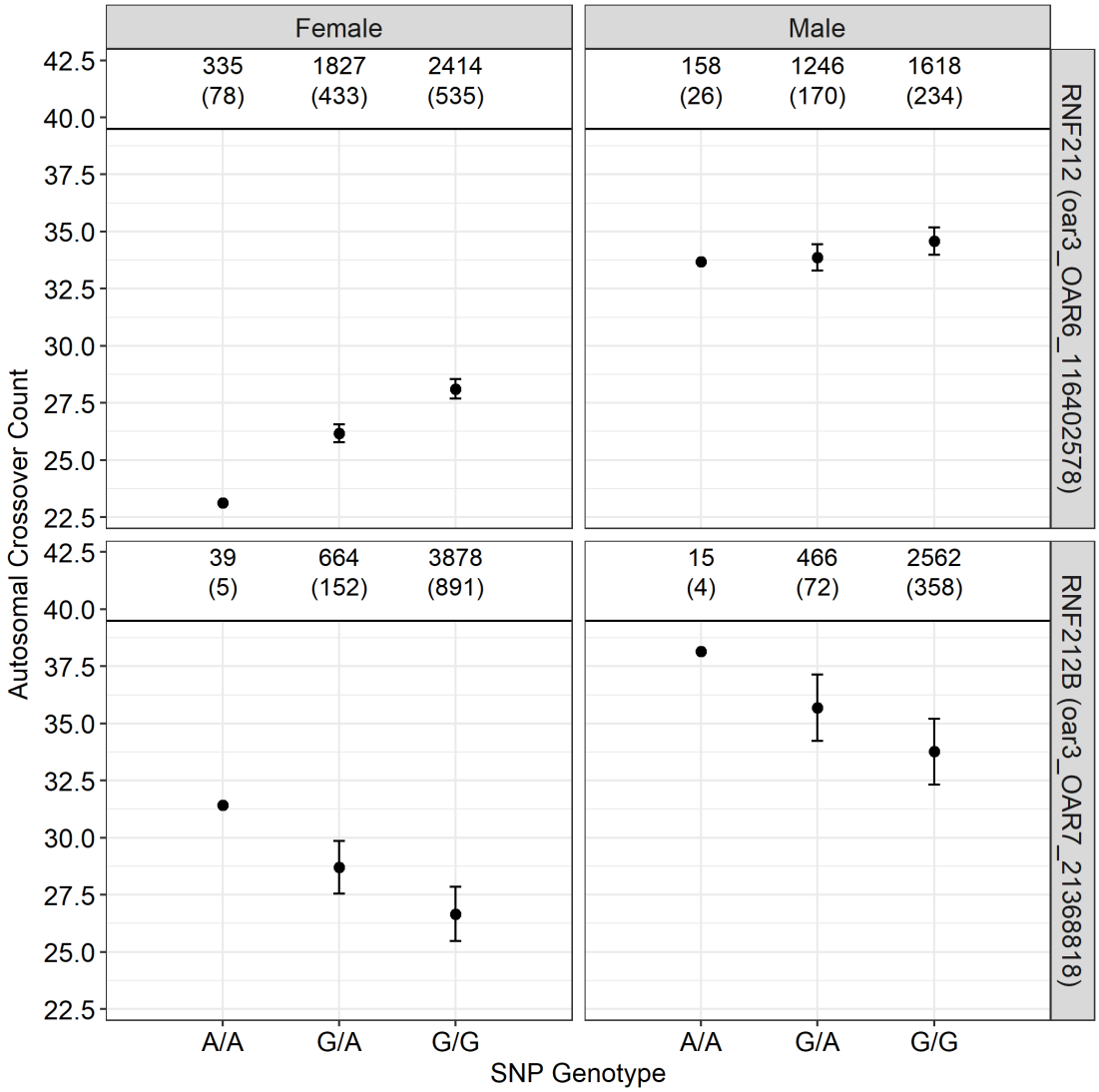

Figure S3: Effect sizes for the most highly associated SNPs shown for males and females, as calculated from animal models including the SNP genotype as a fixed factor. Genotype AA is the model intercept. Numbers indicate the number of observations at each genotype, with the number of unique focal individuals in parentheses.

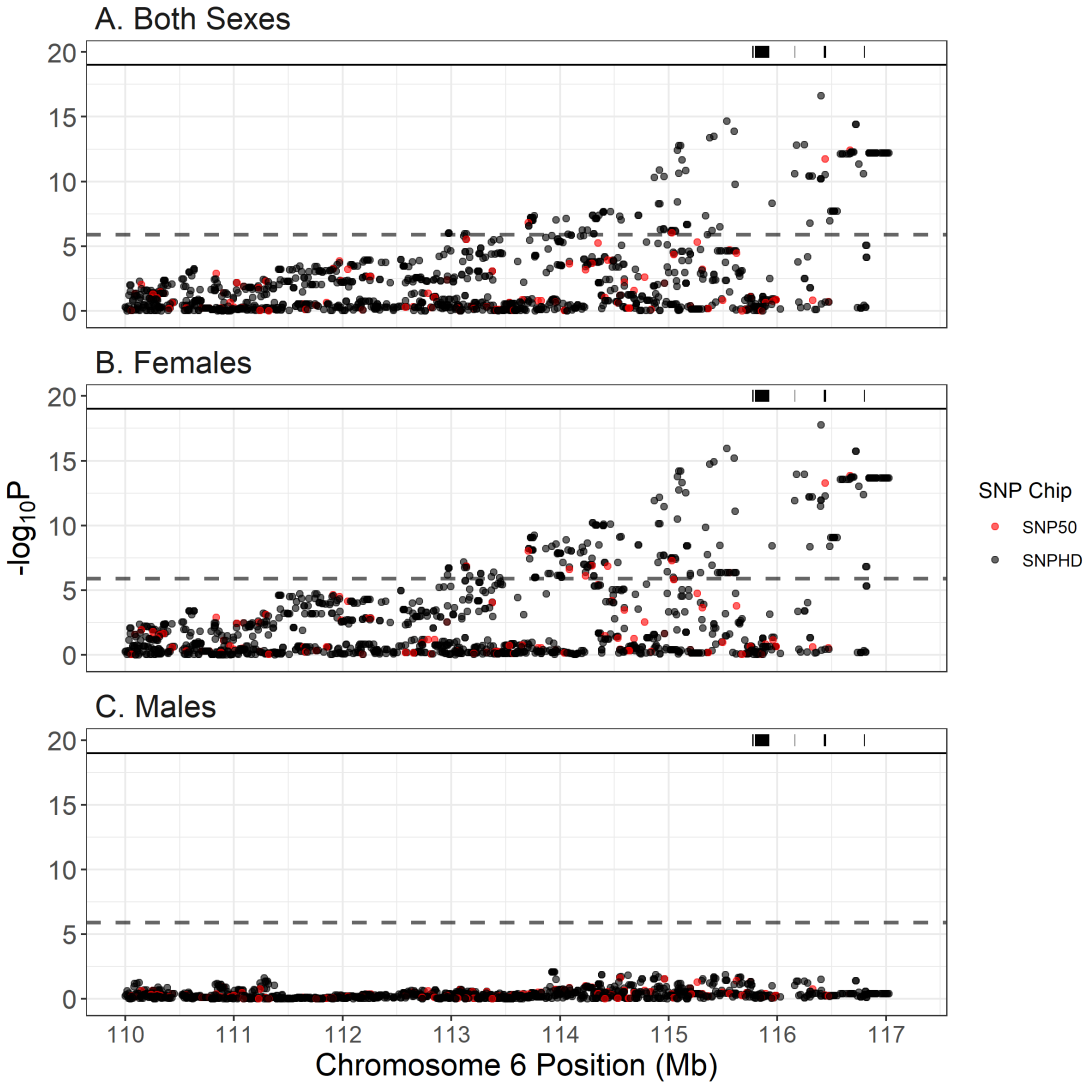

Figure S4: Regional association statistics for the significant region on chromosome 6. Each point represents a single SNP locus, coloured by their origin on the SNP50 (red) or SNPHD Chips (black, i.e. imputed genotypes). The positions of genes associated with meiotic processes are shown as black bars at the top of each plot (Gene information in Table S4).

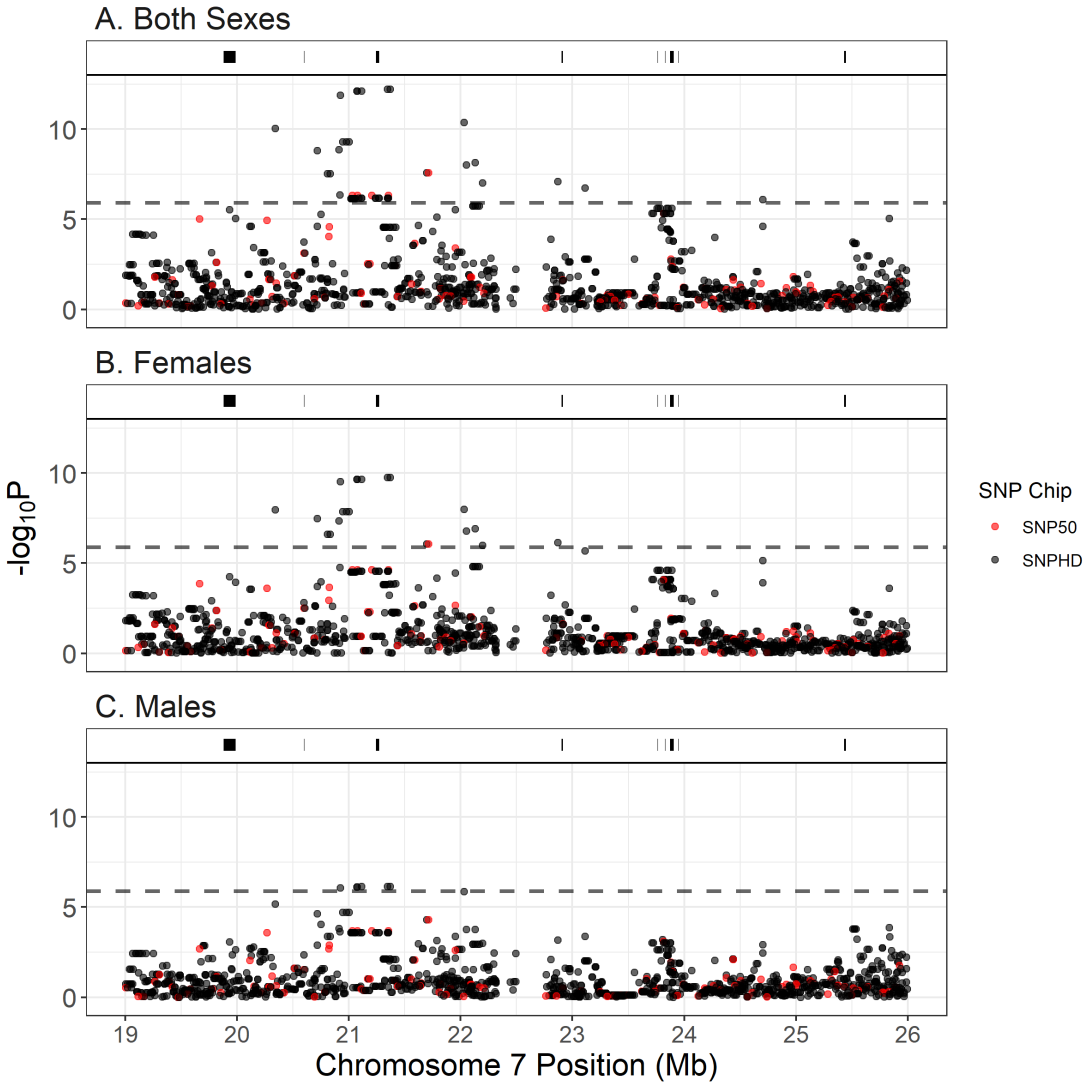

Figure S5: Regional association statistics for the significant region on chromosome 7. Each point represents a single SNP locus, coloured by their origin on the SNP50 (red) or SNPHD Chips (black, i.e. imputed genotypes).The positions of genes associated with meiotic processes are shown as black bars at the top of each plot (Gene information in Table S4).

Table S1: Fixed effects in animal models of autosomal crossover count. Models were run in both sexes, males and females. All Wald statistics were equivalent to a χ^2 test with 1 degree of freedom.

Model	Effect	Estimate	SE	Z Ratio	Wald Statistic	P
Both Sexes	(Intercept)	26.966	0.106	253.849	112596.078	0.0000
	Fhat3	1.148	3.692	0.311	0.097	0.7557
	Sex (Male)	7.236	0.180	40.161	1613.634	0.0000
Males	(Intercept)	34.178	0.144	236.710	61239.042	0.0000
	Fhat3	-3.402	5.872	-0.579	0.336	0.5623
Females	(Intercept)	26.970	0.113	238.303	59962.846	0.0000
	Fhat3	0.853	4.610	0.185	0.034	0.8533

Table S2: Linkage map information for the current study. Order is the order of markers on the chromosome. SNP.Name is the Ovine SNP50 BeadChip identifier. Chr is the chromosome number, where 27 is the X chromosome. cMPosition, r, cMdiff are the sex-averaged centiMorgan position, recombination fraction and centiMorgan difference with the following locus, respectively; similarly this is provided for Female and Male maps. Oar3_Chr and Oar3_Pos are the relative chromosome number and position relative to the sheep genome (Oar_v3.1), cMPosition.2016, cMPosition.Female.2016 and cMPosition.Male.2016 indicate the centiMorgan positions as determined by the previous study (JOHNSTON *et al.*, 2016).

[File: Table_S2_Linkage_Map.txt]

Table S3: Association statistics for genome-wide association studies of autosomal chromosome counts. Sex indicates analyses for both sexes, females only and males only. SNP.Name is the Ovine SNP50 BeadChip identifier. Chromosome and Position (bp) are given relative to the sheep genome Oar_v3.1. A1 and A2 reference and alternate allele at each SNP. effB is the slope of the effect of allele A2, with the standard error se_effB. Chi2.1df and P1df is the association chi-squared statistic and associated P-value, respectively, before correction with genomic control λ. Pc1df is the corrected P-value after genomic control. Exp is the corresponding P-value for that SNP locus assuming a null distribution of P-values (see Figure 2). Q.2 is the minor allele frequency. Cumu is the cumulative genomic distance.

[File: Table_S3_ACC_GWAS.txt]

Table S4: Gene positions in the GWAS significant regions obtained using biomaRt v2.42.1. ensembl_gene_id is the Ensembl identifier for the sheep gene. external_gene_name is the sheep gene name. chromosome_name, start_position and end_position are the chromosome number, start and stop positions for each gene, respectively. Meiotic indicates whether this gene and/or its orthologues in other species have GO terms associated with meiotic processes. Orthos indicates the unique gene names associated with orthologues of each gene.

[File: Table_S4_Gene_List_Table.txt]