Norwegian University of Life Sciences
Faculty of Biosciences
Department of Animal and Aquacultural Sciences

Philosophiae Doctor (PhD)
Thesis 2022:17

Genetic variation in recombination rates and genetic shuffling in pigs, cattle and Atlantic salmon

Genetisk variasjon i rekombinasjonsrater og genetisk omstokking hos gris, storfe og Atlanterhavslaks

Cathrine Brekke

Genetic variation in recombination rates and genetic shuffling in pigs, cattle and Atlantic salmon

Genetisk variasjon i rekombinasjonsrater og genetisk omstokking hos gris, storfe og Atlanterhavslaks

Philosophiae Doctor (PhD) Thesis

Cathrine Brekke

Norwegian University of Life Sciences
Faculty of Biosciences
Department of Animal and Aquacultural Sciences

Ås 2022

Thesis number 2022:17
ISSN 1894-6402
ISBN 978-82-575-1892-9

PhD Supervisors

Prof. Peer Berg
Department of Animal and Aquacultural Sciences
Norwegian University of life Sciences
P.O. Box 5003
1432 Ås, Norway

Dr. Susan E. Johnston
Institute of Evolutionary Biology
School of Biological Sciences
University of Edinburgh
Edinburgh EH9 3FL
United Kingdom

Dr. Arne B. Gjuvsland
Norsvin SA
Storhamargata 44
2317 Hamar, Norway

Dr. Tim Martin Knutsen
AquaGen AS, avd. Ås
Arboretveien 6
1430 Ås, Norway

Prof. Sigbjørn Lien
Department of Animal and Aquacultural Sciences
Norwegian University of life Sciences
P.O. Box 5003
1432 Ås, Norway

PhD Evaluation committee

Ass. Prof. Nadia D. Singh
University of Oregon
77 Klamath Hall
1210 University of Oregon
Eugene, OR97403-1210

Prof. Dirk-Jan de Koning
Swedish University of Agricultural Sciences
Inst för HGEN
Box 7023
750 07 Uppsalla

Committee coordinator:
Prof. Hans Magnus Gjøen
P.O. Box 5003 NMBU
NO-1432 Ås, NORWAY

Acknowledgements

I want to thank my main supervisor, Peer Berg, for taking me on for this project. Your door is always open, and you always make me feel like a colleague rather than a student, which I appreciate. Thank you for discussions, help with software, models, putting me into contact with other people in the field and last but not least for all the chats about our dogs. I also want to thank you for pushing me to make my own decisions, when I hand in my thesis today, this work truly feels like my own.

A very special thank you to Susan Johnston, whom I was only supposed to visit for two months in February 2020, but who have included me as part of her lab through zoom ever since the pandemic hit. Your knowledge in the field has greatly improved this project! Thank you for your ideas and advice, excitement about results, thorough revisioning and attention to detail. I also want to thank you and the rest of the lab for making me feel I was part of a group during almost two years of working from home. The discussions, laughs, software tips and paper sharing in the group has been essential for staying motivated. Thank you!

I want to thank Geno SA, Norsvin SA and AquaGen SA for providing the data for this project. Arne Gjuvsland has been my contact from Geno and Norsvin, thank you for making it easy to navigate in these datasets. You have been helpful inn all steps of this project, from understanding the data and what methods and software to use, to interpreting results and revision of manuscripts. I have really enjoyed working with you!

Tim Martin Knutsen has been my contact from Aquagen. Thank you for showing interest in this project from the beginning and for all the time you have spent trying to put together a dataset fit for this project. I have really enjoyed discussions with you about recombination, but also on the topic of doing a phd and everything that comes with it. I have learned important researcher street knowledge from you like do's and don'ts on a computer cluster. I hope to collaborate more in the future somehow! I also want to mention Kim Erik Grashei in Aquagen, for interesting discussions on recombination in salmon and sketching a triploidy induced meiosis!

I also want to thank Sigbjørn Lien for valuable input on the Atlantic salmon genome.

Thank you to Tormod Ådnøy for revising the Norwegian summary and for always showing interest and taking the time for discussions.

Cathrine

Table of content

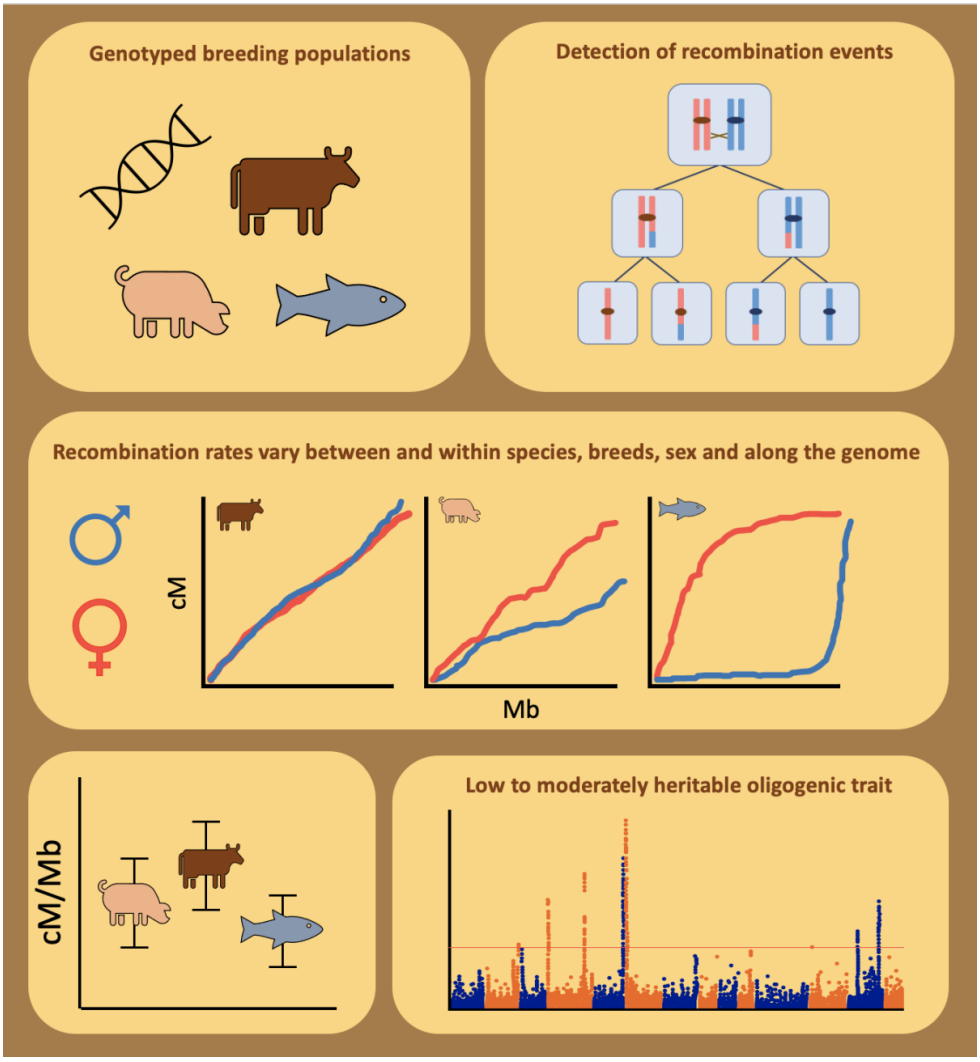
Summary

Meiotic recombination ensures proper segregation of homologues chromosomes in meiosis while also breaking down linkage disequilibrium and creating novel haplotypes by shuffling alleles of genes located on the same chromosome. Understanding how rates and distribution of meiotic recombination vary between populations, sex and individuals is of particular interest in breeding as it can break up unfavorable linkage and produce genetic variance to exploit in selection. In this study we use dense SNP marker data on full-sib families from Norwegian Red cattle, an Atlantic salmon breeding population and five different pig breeding populations to investigate the genetic variation in individual measures of recombination within and between domestic species. All three species exhibit marked sex differences, but with different direction and magnitude. Genome-wide rates are higher in females in pigs and Atlantic salmon, but higher in males in cattle. Heritability of genome-wide rates of recombination was low but significant in both sexes in all three species (ranging from h^2 = 0.04 in female cattle to 0.12 in male Atlantic salmon) which is comparable to estimates in other mammalian species. We detected five regions associated with variation in individual recombination rates in pig with candidate genes associated with meiosis, namely PRDM7, MEI1, RNF212, SYCP2 and MSH4. In cattle, three loci were significantly associated with the trait with four candidate genes that have all been associated with individual recombination rates previously: CEP55, NEK9, MLH3 and RNF212b. The sex specific patterns of recombination in pigs and Atlantic salmon lead to marked differences in the amount of intrachromosomal shuffling of alleles in maternal and paternal gametes, however it is not clear how this affects the overall genetic shuffling in a breeding population across generations. The genetic variation, and oligogenic architecture suggests a potential for genetic change in overall rates and distribution of recombination in these populations.
In summary, the findings in this study contribute to the understanding of the genetic mechanisms underlying recombination rate variation and how this relates to variation in genetic shuffling in a breeding population as well as in natural populations

Samandrag

Rekombinasjon sikrar korrekt fordeling av dei homologe kromosoma i meiosen og samstundes brytar opp koplingsulikevekt og produserer nye haplotypar ved å stokka om på allel frå gen på same kromosom. Forståing av korleis rekombinasjonsratar og fordelinga av overkryssingar langs genomet varierer mellom populasjonar, kjønn og individ og er av spesiell interesse for avl fordi det kan bryta opp uønskt kopling av gen i same kromosom og slik produsera genetisk variasjon som kan nyttast i seleksjon. I denne studien bruker vi SNP-genotype-data på fullsøskenfamiliar frå Norsk Rødt Fe, ein avlspopulasjon av atlanterhavslaks, og fem forskjellige rasar av gris, for å studera genetisk variasjon i individuelle rekombinasjonsratar innan og mellom domestiserte artar. Det er tydelege forskjellar mellom kjønn i alle dei tre artane, men i ulik grad og i forskjellig retning. Hos gris og laks er total rekombinasjonsrate høgare for hokjønn, men hos storfe er rekombinasjonsraten høgare for hannar. Arvegraden for tal overkryssingar er låg, men signifikant, hos alle dei tre artane og for begge kjønn (varierer frå $h^2 = 0.04$ for storfe hokjønn til 0.12 for hannlaks). Dette er liknande nivå som for estimat hos andre pattedyr. Vi observerte fem loci med assosiasjon til variasjon i individuelle rekombinasjonsratar hos gris, med følgande kandidatgen: PRDM7, MEI1, RNF212, SYCP2 og MSH4 er. Hos storfe er tre loci signifikant assosierte med variasjon i rekombinasjonsrater og det er fire kandidatgen som alle har vore assosierte med individuelle rekombinasjonsratar i tidlegare studiar: CEP55, NEK9, MLH3 og RNF212b. Dei kjønns-spesifikke fordelingane av rekombinasjonen hos gris og laks fører til tydelege forskjellar i stokking av allel i maternale og paternale gametar, men det er ikkje klart korleis dette verkar på den generelle genetiske stokkinga av allel på tvers av generasjonar i ein avlspopulasjon. Den genetiske variasjonen og antydninga av oligogen determinisme tyder på at det finst eit potensiale for genetisk endring i rekombinasjonsrate og fordeling av overkryssingar i desse populasjonane. Oppsummert bidrar resultata i denne studien til forståing av dei genetiske mekanismane som verkar på variasjonen i rekombinasjonsratar og korleis dette relaterer til variasjon i genetisk omstokking i ein avlspopulasjon eller i ein naturleg populasjon.

Graphical Abstract



Glossary and clarification of terminology

Achiasmy
When the male or female of a species completely lack meiotic recombination

Aneuploidy
wrong number of chromosomes in a gamete or individual compared to the karyotype of the species

Crossover interference
When a crossover in one location reduces the chance of a crossover occurring in close proximity, resulting in widely spaced crossovers along chromosomes.

Heterochiasmy
When the two sexes of a species exhibit differential rates of recombination

Heterogametic
The heterogametic sex is the sex in which the two sex chromosomes are different in a species with genetic sex determination

Homeologue
A pair of distinct chromosomes that have emerged from a genome duplication and exhibit high sequence similarity

Homologue
A pair of chromosomes that share the same physical structure, i.e the maternally and paternally inherited chromosome in a diploid cell

Recombination
Exchange of genetic information between maternal and paternal homologues chromosomes. A recombination can result in a crossover, where the rest of the chromatin following the crossover is exchanged, or a non-crossover where only a small DNA sequence is exchanged or copied. Throughout this thesis, recombination refers to the crossover recombination.

Abbreviations

ACC – Autosomal crossover count
cM – centimorgan
CO – crossover
GWA – Genome-wide association
Mb - Megabases
WGD – whole genome duplication

List of papers

The following papers are included in the thesis and referred to in the text by their roman numbers.

I. **Brekke, C.,** Berg, P., Gjuvsland, A.B., Johnston, S.E. (2022). Recombination rates vary between breeds, sex and individuals in the domestic pig and is associated with *RNF212*, *SYCP2*, *PRDM7, MEI1* and *MSH4. Submitted manuscript*

II. **Brekke, C.,** Johnston, S.E., Gjuvsland, A.B., Berg, P. (2022). Individual recombination rates are heritable and associated with RNF212b, NEK9, MLH3 and CEP55 in Norwegian Red cattle. *Manuscript*

III. **Brekke, C.,** Johnston, S.E., Knutsen, T., Berg, P. (2022). Genetic variation in genome wide recombination rates and genetic shuffling in an Atlantic salmon breeding population. *Manuscript.*

IV. **Brekke, C.,** Johnston, S.E., Gjuvsland, A.B., Berg, P. (2022). Variation in patterns of recombination result in genetic variation in intrachromosomal shuffling in the domestic pig. *Submitted manuscript*

1. General introduction

1.1 Meiotic recombination

In sexually reproducing eukaryotic systems genetic material is transmitted to the next generation by producing haploid germ cells that are created through the meiotic process. During early prophase 1 of meiosis, the maternal and paternal **homologues** chromosomes align. Homology search is enabled by double-strand breaks along the DNA (Lenormand *et al.*, 2016). An important part of this process is meiotic **recombination**, where double strand breaks in the homologues DNA strands cross over and genetic material between the homologues chromosomes is exchanged. This process ensures that chromosomes align at the correct position (Fledel-Alon *et al.*, 2009), but because this also leads to a novel combination of the parental alleles, meiotic recombination is an important evolutionary force creating haplotypic diversity in each generation. Meiotic recombination has shown to be under tight regulation both in number and distribution. Most species studied have one obligate crossover per chromosome (Stapley *et al.*, 2017). Because of the vital role in correct alignment, recombination ensures proper segregation of chromosomes, and a lack of this obligate crossover can lead to aneuploidy, which is usually lethal (Sherman *et al.*, 1991; Hassold *et al.*, 1995; Koehler *et al.*, 1996). Exactly how the number of crossovers and their distribution is controlled remains to be fully understood, but some principles have been established. Most species studied display a level of crossover interference ensuring that when a crossover happens in one region, a new crossover cannot occur in close proximity (Lenormand *et al.*, 2016). Many species exhibit recombination hotspots along the genome, for example in human more than 80% of crossovers happen in regions covering only 10-20 % of the genome (Coop and Przeworski, 2007). The gene PRDM9 has been identified as a fast-evolving gene that drives this variation in hotspot usage across taxa (Baudat *et al.*, 2010). However, some species, like Drosophila and dogs do not have defined hotspots and lack the PRDM9 gene completely ((Muñoz-Fuentes, Rienzo and Vilà, 2011; Smukowski Heil *et al.*, 2015).

1.2 Why relevant for breeding

Genetic variation is a prerequisite for evolutionary change of a population and genetic gain in a breeding population. In each generation, individuals with combinations of alleles favorable for the traits of interest in a population will be selected to produce offspring in the next generation. Most of the reshuffling of alleles from one generation to the next is due to independent assortment of loci located on different chromosomes, Mendels second law. However, if two different genes affecting traits of interest are linked on the same chromosome, reshuffling of these alleles require meiotic recombination. The number and location of recombination events affect how many pairs of loci are shuffled during meiosis (Veller, Kleckner and Nowak, 2019). Despite this, the effect of variation in rates and distribution of crossovers on the amount of genetic shuffling in a population has received little attention in breeding. And it would be of interest to better understand the impact of recombination for the production and maintenance of genetic variation in a breeding population.

1.3 Status of knowledge genetic variation in recombination

1.3.1 Difference between species and populations

Genome-wide rates of recombination vary within and between species, sex and individuals in virtually all species studied to date (Lenormand and Dutheil, 2005a; Smukowski and Noor, 2011; Ritz, Noor and Singh, 2017; Zelkowski *et al.*, 2019)**.** There are a few examples of species with extreme rates of recombination compared to other species, like social insects such as honeybees and bumblebees with rates as high as ~ 26 and ~9 cM per Mb, respectively (Kawakami *et al.*, 2019). However, across most species studied to date there seems to be a shared magnitude of variation in crossover count (Ritz, Noor and Singh, 2017). A general rule of thumb that holds across an impressive number of species is that one recombination event occurs per chromosome arm (Coop and Przeworski, 2007). However, even with similar rates of recombination, two species may display very different distributions of crossovers, where in some species recombination rates are relatively evenly distributed along the genome other species may display elevated rates in telomeric or centromeric regions (Martinez-Perez and Colaiácovo, 2009; Stapley *et al.*, 2017).

1.3.2 Difference between sex

Sex differences in rates of recombination has been of interest for decades (Dunn and Bennett, 1967; Burt, Bell and Harvey, 1991; Barton and Charlesworth, 1998; Lenormand and Dutheil, 2005). The extreme version is species where only one of the sexes recombine, known as **achiasmy**, forexample in Drosophila (Brooks and Marks, 1986). But also, in species where both sexes recombine, there is usually a level of variation in recombination rate or landscape between males and females, referred to as **heterochiasmy** (Mank, 2009; Stapley *et al.*, 2017). Despite the large interest in this phenomenon, it remains to be fully understood. One longstanding theory is that because recombination between two different sex chromosomes can be detrimental, lower recombination rates have evolved in the **heterogametic** sex. This theory holds for most achiasmate species studied, but not in heterochiasmate species, where there are many examples of species where the heterogametic sex display higher recombination rates (Sandor *et al.*, 2012; Ma *et al.*, 2015; Johnston *et al.*, 2016; Petit *et al.*, 2017). Even though the mechanisms and evolutionary forces leading to heterochiasmy remain to be fully understood, sex is generally established as the main driver of variation in recombination rates within a population, even in species that don't have sex chromosomes, like the saltwater crocodile (Isberg *et al.*, 2006) and salmonid fish (Sakamoto *et al.*, 2000; Lien *et al.*, 2011).

1.3.3 Individual rates of recombination

Genetic variation in individual recombination rates have been a topic of great interest the last decades. In mammals in particular great insights into the genetic architecture of the trait has been achieved (Kong *et al.*, 2014; Ma *et al.*, 2015; Johnston *et al.*, 2016; Petit *et al.*, 2017; Wang *et al.*, 2017; Johnston, Huisman and Pemberton, 2018; Johnston, Stoffel and Pemberton, 2020; Johnsson *et al.*, 2021)**.** Phenotypic variation explained by genetic variation is usually around 10-15% but estimates as low as 5% in some pig breeds (Johnsson *et al.*, 2021) and as high as 46 % in house mice populations (Booker, Ness and Keightley, 2017) have been presented. Several genes are reported to have an effect on the trait, and some of these genes, like RNF212, RNF212B and REC8 are observed in a variety of species like cattle(Sandor *et al.*, 2012), soay sheep (Johnston *et al.*, 2016), domestic sheep (Petit *et al.*, 2017), Red deer (Johnston, Huisman and Pemberton, 2018), and human (Kong *et al.*, 2014). Overall, results point to an oligogenic control of the trait, and even though some genetic architecture seems to be conserved between distantly related species and even across taxa, some genes that have a high effect on the trait in some species are completely lacking in

others, for example drosophila lacks RNF212 and Rec8 entirely, but still exhibit genetic variation in the trait, associated with other genes (Hunter *et al.*, 2016)

1.4 Populations in this study

In this study, recombination rate variation is studied in commercial breeding populations, one cattle breed, Norwegian red cattle from the Norwegian breeding company Geno SA, five pig breeds; Norwegian landrace, Norwegian Duroc, Large white, Piétrain and a synthetic breed from the Dutch/Norwegian pig breeding company Topigs Norsvin SA and Atlantic salmon population from the Norwegian breeding company Aquagen. Cattle, pigs and Atlantic salmon are the three most important species economically for breeding in Norway. The breeding work on these three species have therefore included extensive genotyping for several years. A high number of genotyped individuals with accurate pedigrees make these populations well suited for studying individual recombination rates.

1.4.1 Norwegian red cattle

The breeding work on the Norwegian red cattle dates back to 1935 and the basis of the population is a mix of imported Ayrshire from Scottland and Swedish red as well as many of the different landraces in Norway; Trønderfe, Telemarksfe, Dølafe, Rødkolle and Sør og Vestlandsfe. The Breeding work on Norwegian red cattle implemented genomic selection in 2016 and have since genotyped around 30 000 individuals yearly.

1.4.2 Norwegian and dutch pig breeds

The breeding work of the Norwegian pig breeding company Norsvin started in 1958. The Norwegian Landrace have been a closed breeding population since the beginning with no introgression from other breeds. In 2014 Norsvin merged with the Dutch pig breeding company Topigs and the company Topigs Norsvin today has active breeding on the five breeds in this study Norwegian Landrace, Norwegian Duroc, Piétrain, Large white and a synthetic mixed breed. Genomic selection was implemented in 2012 and individuals from all five breeds are genotyped every year making it possible to study genetic variation within and across breeds in these datasets.

1.4.3 Atlantic Salmon

The breeding work on Atlantic salmon in Aquagen dates back to 1971. Atlantic Salmon from 40 different Norwegian rivers and 1 Swedish river made up the basis of the breeding population at that point. After implementation of genomic prediction in the breeding work in 2016, thousands of individuals are now genotyped per year. Both male and female breeding individuals can have several hundreds of genotyped offspring resulting in a high number of repeated measures on individual crossover counts.

1.5 Objectives and aims

In this thesis, breed and sex specific rates and distribution of crossovers was estimated in five domestic pig breeds, Norwegian red cattle and an Atlantic salmon breeding population. The aim was to:

- Determine whether individual recombination rates were heritable in these breeding populations and identify potential loci associated with the observed variation in recombination rates.
- Compare estimates of variation in recombination rates within and between sex, breeds and species in this study to contribute to the understanding of how patterns of recombination may evolve across species and populations.

2. Material and Methods

2.1 Summary of datasets

In this project genotyped breeding populations with pedigree information are used to detect recombination events. Data for paper I. and IV. is on five different pig breeds provided by the Dutch and Norwegian pig breeding company Topigs Norsvin. The dataset in paper II. Is on the Norwegian red cattle breed provided by Geno SA. The data on Atlantic Salmon in paper III. is from the breeding company Aquagen SA. The datasets are described in detail in the respective papers. An overview of the datasets is provided in table 1.

					FIDs	
Species	**Individuals**	**Families**	**Breeds**	**markers**	**Male**	**Female**
Pig	257 295	26 048	5	50 705	1204	15 176
Cattle	110 555	19603	1	35 880	603	14 815
Atlantic salmon	128 363	1952	1	35715	621	416

Table 1. Total number of genotyped **individuals**, number of three generation full sib **families, SNP** is number of markers with genotypes, **FIDs** is number of unique males and females where recombination rates could be measured.

2.2 Methods

2.2.1 Pipeline

The following pipeline was developed to quantify individual rates and patterns of recombination which made the basis of all analysis in the project. The methods are described in detail in the respective papers, but a general overview is given here and illustrated in Figure 1. The pedigrees were sorted into three generation full sib families where each unique sire dam mating pair in the pedigree made the basis of a family (Figure 1A). Number of full sib families are presented in table 1. Sex and breed specific linkage maps were constructed with the software Lepmap3 (Rastas, 2017) to obtain population level and genome wide patterns and rates of recombination (Figure 2A). Number and location of individual recombination events were obtained from the phased offspring gametes (Figure 2C) and assigned as an observation in the parent where meiosis took place. Parents are hereafter referred to as the focal individuals or FIDs. Resulting in a dataset with related individuals

with repeated observations on measures of recombination. Further, variance components for the traits were estimated with DMU (Madsen *et al.*, 2014) and genome-wide association analysis was carried out with the GCTA software (Yang *et al.*, 2011) (Figure 1D). Development of scripts for data manipulation was done with R (R Core Team, 2020) and Bash (Ramey, 2011). Plots and figures were made with base R or ggplot2 (Wickham H, 2016).

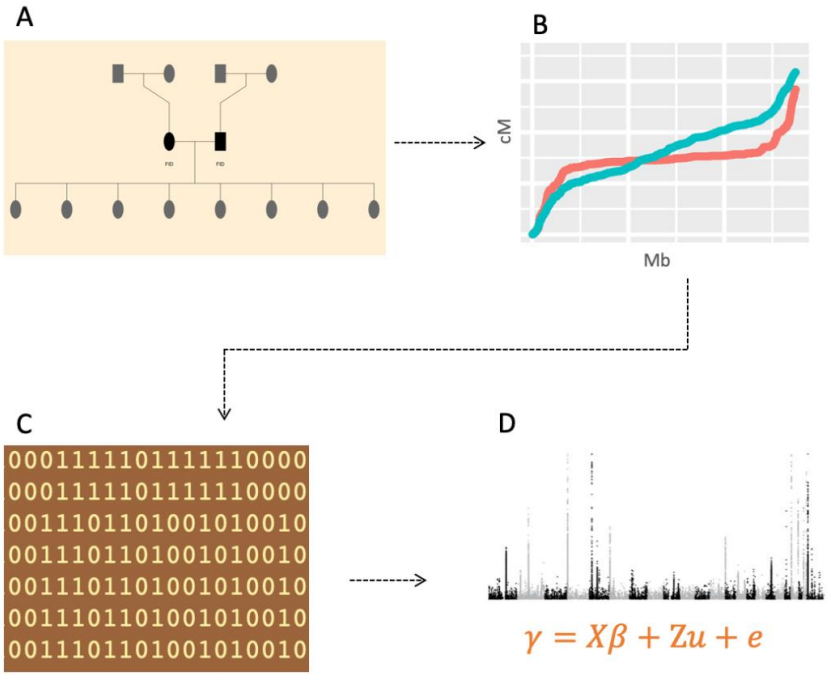


Figure 1. Pipeline developed for analysis of genetic variation in individual measures of recombination. 1A is the three-generation full-sib family structure the pedigrees were ordered into. 1B illustrates sex sepsific linkage map construction. 1C shows phase information of gametes, from where number and position of individual crossovers were detected. And 1D illustrates genetic analysis performed on the datasets obtained with the pipeline.

3. Papers

Paper I.

Recombination rates vary between breeds, sex and individuals in the domestic pig and is associated with *RNF212*, *SYCP2*, *PRDM7*, *MEI1* and *MSH4*.

Brekke, C., Berg, P., Gjuvsland, A.B., Johnston, S.E. (2022).

Submitted manuscript

Recombination rates vary between breeds, sex and individuals in the domestic pig and is associated with *RNF212*, *SYCP2*, *PRDM7*, *MEI1* and *MSH4*

Cathrine Brekke[1*], Peer Berg[1], Arne Gjuvsland[2], Susan E. Johnston[3]

[1] Norwegian University of Life Sciences, Department of Animal and Aquacultural sciences, Oluf Thesens vei 6, 1433 Ås, Norway

[2] Norsvin, Oluf Thesens vei 6, 1433 Ås, Norway

[3] Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh, Edinburgh, EH9 3FL, United Kingdom

E-mail addresses:

CB: cathrine.brekke@nmbu.no

PB: peer.berg@nmbu.no

AG: arne.gjuvsland@norsvin.no

SEJ: susan.johnston@ed.ac.uk

ABSTRACT

Background: Recombination is a fundamental part of mammalian meiosis that leads to the exchange of large segments of DNA between homologous chromosomes and is therefore an important driver of genetic diversity in populations. In breeding populations, understanding recombination is of particular interest because it can break up undesired linkage and produce novel combinations of alleles to exploit in selection. In this study, we use dense SNP marker data and pedigree information to look at individual and sex-specific variation and genetic architectures of recombination rates within and between five pig breeds.

Results: In agreement with previous studies recombination rates were higher in females than males for all breeds and all chromosomes except 1 and 13, where male rates were slightly higher. There was variation in total recombination rate between the breeds, but the pattern of recombination along chromosomes was well conserved across breeds for the same sex. The autosomal linkage maps spanned a total of 1731 to 1887 cM for males and 2231 to 2515 cM for females. The heritability for individual autosomal crossover count ranged from $h^2 = 0.04$ to 0.07 in males and $h^2 = 0.08$ to 0.11 in females. We found 14 regions associated with individual autosomal crossover count: two peaks had novel candidate genes, *PRDM7* and *MEI1*, that are involved in meiosis but have not been previously associated with individual recombination rate; and four peaks were close or within candidate genes that have been associated with individual recombination rates in pigs and other mammals previously, namely *RNF212*, *SYCP2* and *MSH4*.

Conclusions: This study shows that genetic variation in autosomal recombination rate persists in domesticated species under strong selection, with differences between closely related breeds and marked differences between the sexes. Our findings support results from

other studies that recombination rates have an oligogenic and relatively conserved genetic architecture in mammals.

Background

Meiotic recombination is the event in meiosis where double strand breaks are resolved as crossovers, resulting in recombined homologous chromosomes. Recombination therefore leads to haplotypic diversity by breaking up linkage within chromosomes and creating novel combinations of alleles for selection to act upon. Recombination also has an important function in the proper segregation of homologous chromosomes in meiosis, and its absence can often lead to nondisjunction in meiosis and aneuploidy in the resulting gametes (Sherman *et al.*, 1991; Hassold *et al.*, 1995; Koehler *et al.*, 1996; Fledel-Alon *et al.*, 2009). Hence, most species have at least one obligate crossover per chromosome pair (Stapley *et al.*, 2017). However, recombination can also break up beneficial allele combinations previously favoured by selection (Barton and Charlesworth, 1998) and double strand break formation can increase the risk of mutations and chromosomal rearrangements (Arbeithuber *et al.*, 2015; Halldorsson *et al.*, 2019). These benefits and costs were thought to tightly regulate the rate of recombination (Coop and Przeworski, 2007), yet recombination rates have been found to vary to a large degree across a diverse range of taxa (Ritz, Noor and Singh, 2017; Stapley *et al.*, 2017).

In the last decade, studies of variation in recombination rates have been conducted in a number of mammal populations, including model species such as mice (Dumont, Broman and Payseur, 2009; Wang *et al.*, 2017), domestic species such as pigs, cattle and sheep (Ma *et al.*, 2015; Petit *et al.*, 2017; Johnsson *et al.*, 2021), and natural populations such as humans, Soay sheep and Red deer (Halldorsson *et al.*, 2019)(Johnston *et al.*, 2016; Johnston, Huisman

and Pemberton, 2018)(Petit *et al.*, 2017; Johnsson *et al.*, 2021). Recombination rates often have substantial genetic variation in most species studied, with heritabilities (h^2) ranging from 5% in pigs (Johnsson *et al.*, 2021) to 46% in house mice (Dumont, Broman and Payseur, 2009). In addition, most mammals are heterochiasmate (i.e. both sexes recombine, but at different rates), but the direction and degree can vary even between closely related species (Burt and Bell, 1987). Some loci have repeatedly been found to be associated with individual recombination rates in mammals, including *RNF212*, *RNF212B* and *REC8* (Kong *et al.*, 2008; Sandor *et al.*, 2012; Johnston *et al.*, 2016; Petit *et al.*, 2017; Johnston, Huisman and Pemberton, 2018; Halldorsson *et al.*, 2019), suggesting that some of the genetic architecture of the trait is well conserved across species. However, there are also novel candidate loci that may be specific for the species studied or that for other reasons have not been detected in previous studies (Kadri *et al.*, 2016; Halldorsson *et al.*, 2019; Johnsson *et al.*, 2021). Hence, there is still significant interest in understanding the genetic mechanisms that drive changes in recombination rate within populations.

In breeding populations, understanding recombination is of particular interest because it can break up undesired linkage and produce novel combinations of alleles to exploit in selection. Higher recombination rates may help quantitative traits respond to selection faster and potentially to a greater degree (Battagin *et al.*, 2016) as they can increase additive genetic variance for selection on fitness and production-related traits (Charlesworth and Barton, 1996). This potential advantage has led to long-standing theory that domestication has indirectly selected for increased recombination rates in domestic mammals (Burt and Bell, 1987), although this view has been challenged by more recent studies where domesticated species are found to have lower recombination rates than their wild counterparts (Munoz-Fuentes *et al.*, 2015).

In this study, we use genomic data from more than 50,000 SNP loci and extensive pedigrees in five domestic pig breeds (*Sus scrofa*) to study genetic architecture and variation in individual autosomal crossover count (ACC) in more than 250,000 pigs. Our objectives were to: (a) construct high density linkage maps to characterise sex-specific recombination landscapes; (b) characterise the genetic architecture of ACC by determining its heritability and identifying individual loci associated with variation; and (c) examine differences in ACC and its genetic architecture between different breeds and sexes.

MATERIALS AND METHODS

Breeds

This study focused on five purebred commercial breeding populations with pedigree and genotype data: two sow breeds, Landrace (LR) and Large White (LW); and three boar breeds, Duroc (DU), Synthetic (SY) and Pietrain (PI).

Genotype data

Genotypes were available from two different medium density SNP chips: Illumina GeneSeek custom 80K SNP chip and Illumina GeneSeek custom 50K SNP chip. The physical positions of the SNPs were determined based on the Sscrofa11.1 reference genome. The two SNP arrays have 50 705 SNPs in common. The data was filtered to remove loci with minor allele frequencies < 0.01, call rate < 0.95 strong deviation from Hardy Weinberg equilibrium (χ^2> 600). The sex chromosomes were not included in the study. An overview of the number of markers and animals from the five populations can be found in Table 1. This genotype set will be referred to as the 50K set. A set of imputed genotypes to 660K (Axiom Porcine Genotyping Array) was available for all breeds, this dataset will be referred to as the 660K set.

Table 1. Genotype data information

Line	Nr of markers	Nr of animals	Males	Females
LR	50705	70943	11685	59258
DU	50705	17137	8397	8740
LW	50705	95613	32683	62930
PI	50705	22784	15009	7775
SY	50705	50818	30198	20620

Number of animals and SNP markers in the datasets after filtering for minor allele frequencies < 0.01, call rate < 0.95 strong deviation from Hardy Weinberg (χ^2> 600), markers on sex chromosomes and unmapped markers.

Full-sib family pedigrees

For each breed, we sub-divided pedigrees into three-generation full-sib families, where each unique dam and sire mating pair combination (hereafter referred to as focal individuals, or FID) was included with their parents and offspring. This allows for phasing of the FID and offspring genomes and determining the recombination that occurred in the gamete transmitted from an FID to its offspring. An FID can be in several different full-sib families (i.e., when mating with a different individual), but the above design means that each individual meiosis is only counted once. An example of a three-generation full-sib family is illustrated in Figure 1. Only families with at least one offspring, two FIDs (parents) and all four grandparents genotyped were included. An overview of the number of families and unique FIDs can be found in Table 2.

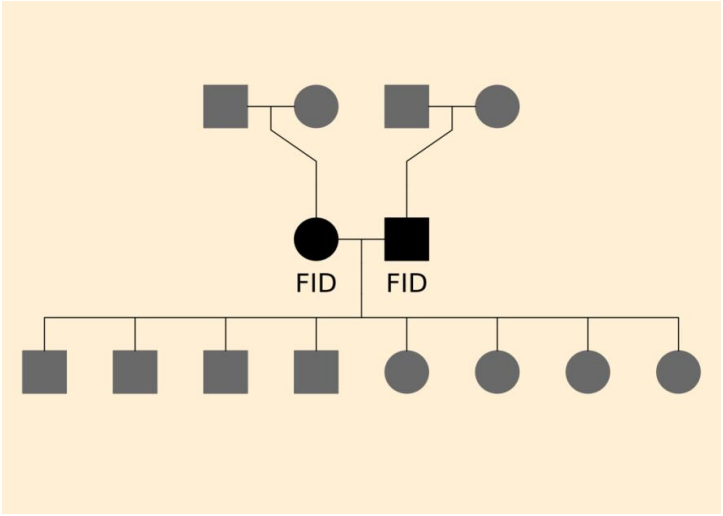


Figure 1 Illustration of the full sib family structures.
The focal individuals (FIDs) in the study are the parents, in black, and the recombination events studied are the ones transmitted in gametes from the FIDs to the offspring.

Table 2. Full-sib family datasets

Line	N_{fam}	N_{obs}	$N_{offspring/family}$	Sires	Dams
LR	7295	74534	1-24	319	4808
DU	5101	18365	1-18	192	1687
LW	6845	82196	1-24	273	4695
PI	2370	24198	1-27	196	1353
SY	4437	51245	1-20	224	2633

N_{fam} is the number of families in each line. N_{obs} is the total number of observations within each line, i.e. total number of meioses. $N_{offspring/family}$ is the range in number of offspring in the families. Sires and dams are the number of unique male and female FIDs in each line.

Linkage mapping

The markers were physically mapped using positions on the Sscrofa11.1 reference genome. This was done by extracting flanking sequences for each marker from the chip manifest files and aligning them to the 11.1 reference genome using the bwa software (Li and Durbin, 2009). We then constructed the population specific linkage maps with LepMap3 (Rastas, 2017) using the linkage groups and marker ordering from the physical mapping. The *filtering2* module was run to filter markers based on segregation distortion, with *dataTolerance* = 0.01 as suggested for multi-family datasets. The *Seperatechromosomes2* module was used as a filtering step rather than to identify the linkage groups from scratch; each set of SNP markers in a previously known linkage group was run through the module, and markers that were not assigned (LOD score < 5) to the main linkage group were excluded. The number of markers in the final linkage maps can be found in Supplementary Table 3. The *Ordermarkers2* module was run with the option to evaluate the given marker order, i.e. to calculate the centimorgan (cM) positions for the markers based on the pre-ordering of the markers using the Haldane mapping function option.

Individual recombination rates

Individual recombination rates were measured as autosomal crossover counts (ACC). Crossovers were counted from the gamete phase from the output of the *orderMarkers2* module from offspring, and assigned to each parent, i.e., FIDs in which the meiosis took place.

Genetic variation

We estimated variance components for individual ACC with a repeatability model using the restricted maximum likelihood (REML) method and average information (AI) algorithm in DMU v 6 (Madsen *et al.*, 2014). The following model was used:

$$\mathbf{Y = sex + b_1*age + id1 + id2 + b_2*het + e}$$

where Y is the ACC, **sex** is the fixed effect of sex, $\mathbf{b_1}$ is the fixed regression of age of the FID when the offspring is born (from ages 1 to 4), **id1** is the random additive genetic effect of the FID, **id2** is the random effect of the FID permanent environment (i.e. environmental effects that are constant across repeated measures on an FID), **het** is the method-of-moments F coefficient estimates (i.e. observed homozygosity count – expected homozygosity count) / total observations – expected homozygosity count) calculated with the *--het* function in PLINK1.9 (Chang *et al.*, 2015), $\mathbf{b_2}$ is the regression of ACC on **het** of the FID, and **e** is the residual effect. The narrow-sense heritability (h^2) was defined as the proportion of phenotypic variance explained by the additive genetic effect and was estimated separately for each breed and sex.

GWAS

Genome wide associations with mean ACC was analysed with the *fastGWA* module implemented in GCTA (Yang *et al.*, 2011) using the 660K datasets. This is a mixed models-based tool that uses a sparse genomic relatedness matrix (GRM) to correct for relatedness and principal components to control for population stratification. The sparse GRM was calculated with the *--make-bK-sparse* option using a cut-off value of 0.05 based on the full genomic dataset, i.e., off-diagonal elements below 0.05 will be set to 0 (Yang *et al.*, 2011).

All analysis and data manipulation where no other software is mentioned was carried out in R 3.6.3 (R Core Team, 2020).

Results

Linkage maps

The sex-specific autosomal linkage maps spanned a total length of 1731.05 (PI) to 1887.14 (LW) cM for males and 2231.32 (DU) to 2515.40 (LW) cM for females. In all breeds, the female recombination rates are higher on all chromosomes except chromosome 1 and 13, where the male rates are slightly higher. The LW breed has the highest recombination rate of all five breeds for both sexes. The genetic length of each chromosome for all breeds are presented in Table 3. The relationship between physical length (Mb) and genetic length (cM) of the chromosomes are close to linear in males, but clearly non-linear in females, the relationship is plotted with robust locally weighted regression in Figure 2 (Cleveland, 1979). The recombination rate is elevated towards the telomeres in both sexes in all chromosomes including the acrocentric chromosomes 13-18. The total recombination rate and pattern along the chromosomes are more similar between the same sex across breeds (Figure 3 and 4) than between sexes from the same breed (see example comparison of male and female maps for the LW breed, Figure 5).

Table 3. Summary of linkage map lengths (in cM) by line, sex and chromosome.

			Genetic length (cM)									
			LR		DU		LW		PI		SY	
Chr	Mb	N_{SNPs}	M	F	M	F	M	F	M	F	M	F
1	274.3	4858	145.2	122.9	156.8	128.1	158.9	142.0	145.8	129.9	159.3	134.8
2	151.9	3306	109.5	129.9	108.8	131.3	117.8	149.0	107.9	129.5	115.8	138.5
3	132.9	2912	111.6	133.5	112.2	131.5	121.0	154.7	113.4	130.4	120.4	140.2
4	130.9	3002	104.2	132.0	102.6	133.4	110.2	151.5	103.3	135.9	108.7	140.4
5	104.5	2290	98.1	140.8	98.4	136.0	103.6	156.8	96.8	144.2	105.1	148.5
6	170.8	3440	119.6	149.3	125.9	156.8	130.3	179.4	120.8	149.8	127.7	162.5
7	121.8	2751	113.4	137.2	112.3	137.8	117.6	158.6	107.8	140.0	115.6	146.6
8	139.0	2924	102.1	122.2	103.1	124.3	105.4	137.6	104.0	126.8	106.3	132.6
9	139.5	3168	104.7	142.0	103.6	143.5	107.3	163.7	104.7	147.8	108.8	156.5
10	69.4	1510	93.9	121.3	91.8	124.3	98.0	138.3	88.3	131.3	94.7	131.4
11	79.2	1846	69.7	114.4	67.8	110.0	79.3	122.6	63.9	120.6	78.3	117.8
12	61.6	1296	78.3	122.0	77.3	118.0	86.5	133.7	71.3	123.6	82.4	127.2
13	208.3	3669	116.9	109.5	126.0	116.9	127.5	123.4	117.6	115.5	126.1	117.6
14	141.8	3284	105.3	124.7	109.8	126.8	114.8	151.9	105.8	124.4	112.7	137.0
15	140.7	2916	98.9	113.0	101.7	115.6	104.9	124.8	99.8	116.4	106.0	119.8
16	79.9	1829	67.0	105.5	69.9	103.6	77.0	111.3	63.2	110.5	74.4	108.4
17	63.5	1399	61.4	106.0	71.9	106.3	69.6	115.7	60.5	106.4	69.3	113.2
18	56.0	1257	54.5	89.5	54.6	87.4	57.6	100.3	56.1	94.7	59.4	95.6
Sum	2266.1	47657	1754.4	2215.7	1794.4	2231.3	1887.1	2515.4	1731.1	2277.8	1871.2	2368.7

Mb is the physical length, N_{SNPS} is number of SNPs in each linkage map after filtering in LepMap3. Sum is the total autosomal genetic length in cM. M is for male and F is for female.

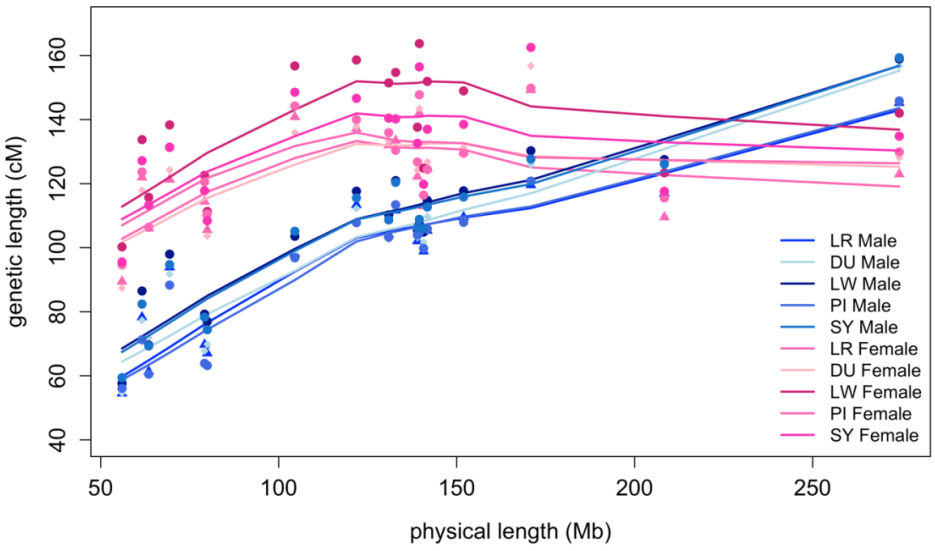


Figure 2 Relationship between the physical (Mb) and genetic (cM) length of the chromosomes. The relationship between the genetic length in cM (x-axis) and physical length in Mb (y-axis is plotted for each chromosome per breed and sex. The relationship is plotted with robust locally weighted regression using the Lowess smoother in R.

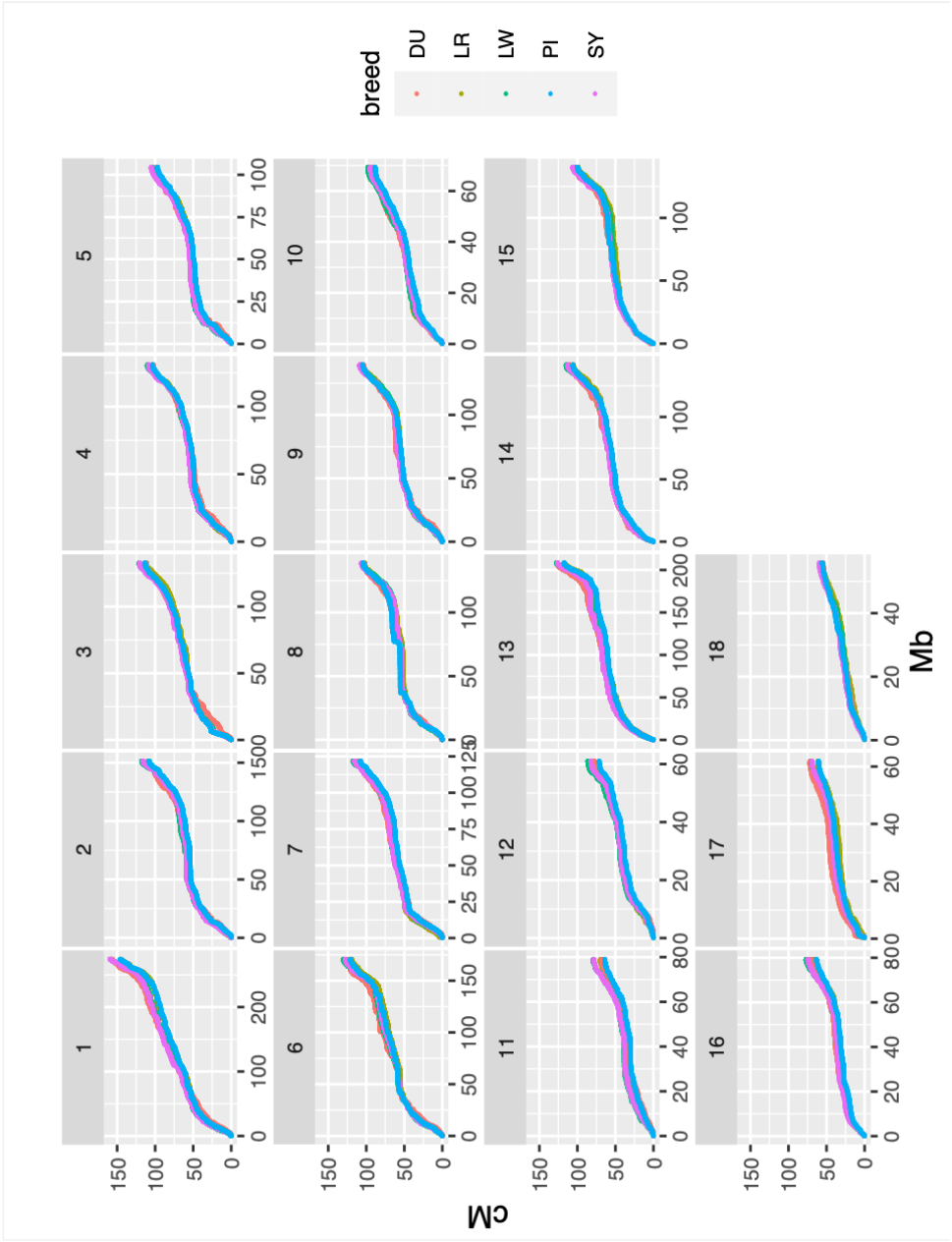


Figure 3 Male autosomal linkage maps for all five breeds.
Genetic positions of the SNP markers in cM are plotted against the physical positions of the SNP markers in Mb. Chromosome numbers are given in the facet headers.

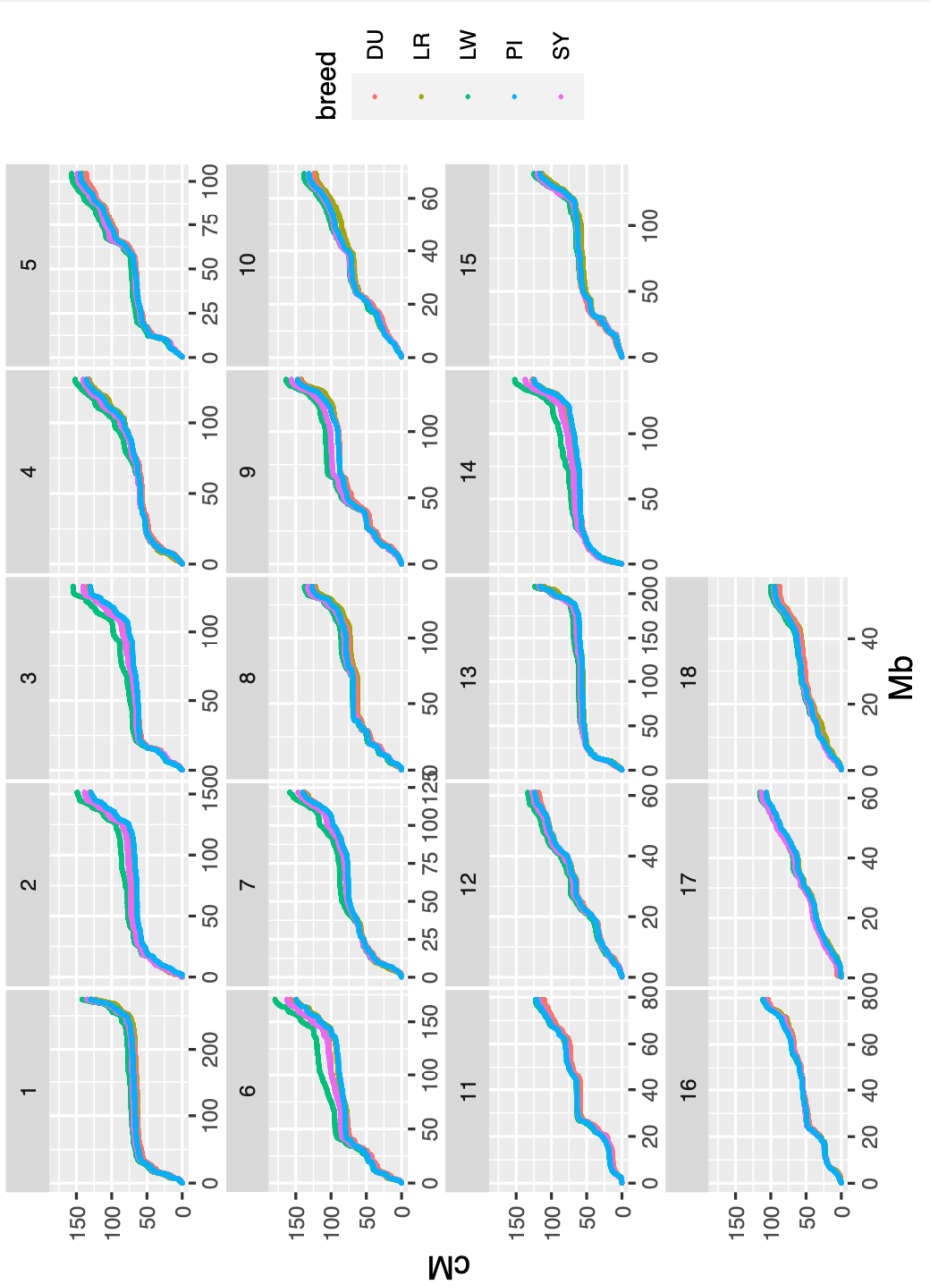


Figure 4 Female autosomal linkage maps for all five lines.
Genetic positions of the SNP markers in cM are plotted against the physical positions of the SNP markers in Mb. Chromosome numbers are given in the facet headers.

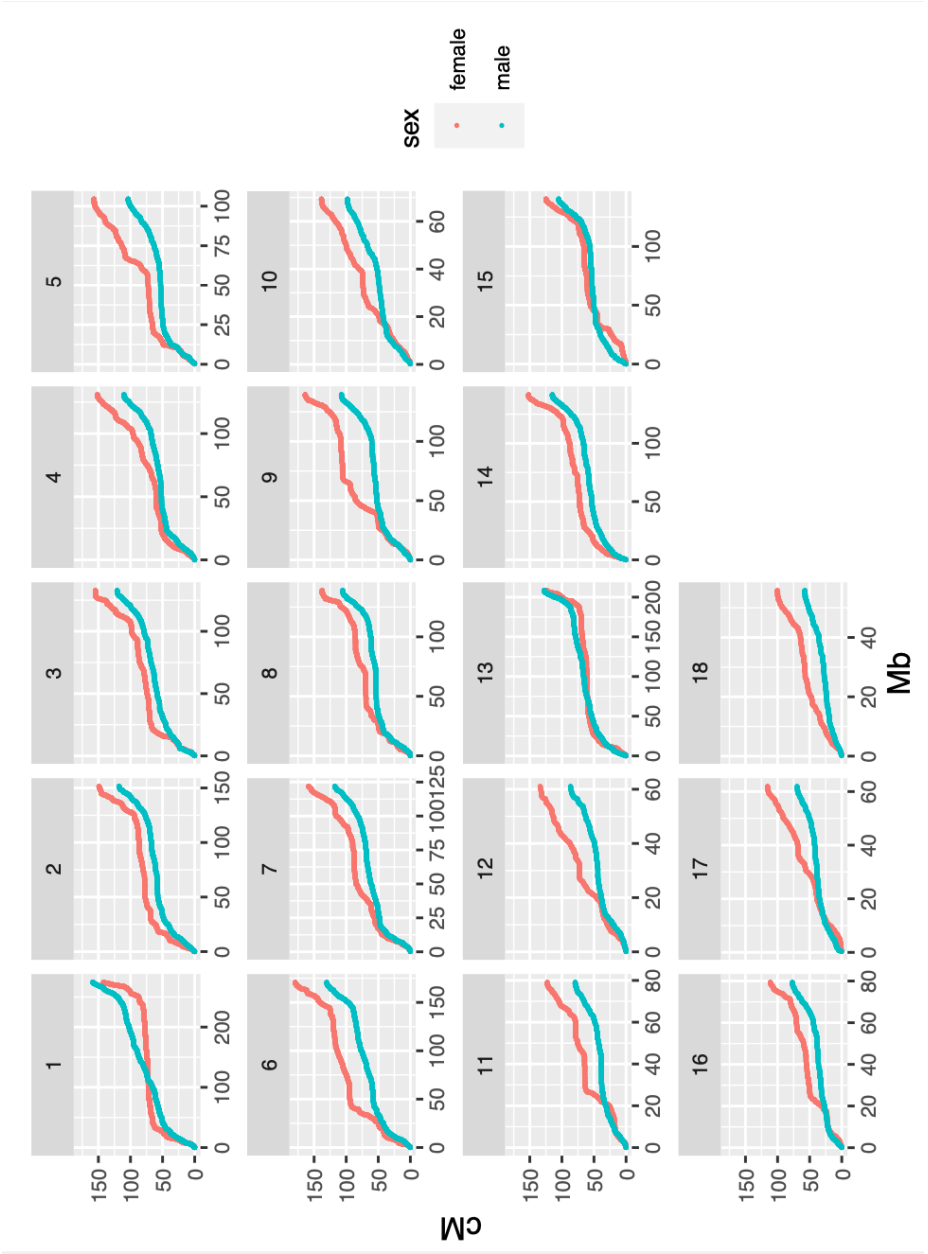


Figure 5 Comparison plot of the LW males (blue) and females (red).
Genetic positions of the SNP markers in cM are plotted against the physical positionsof the SNP markers in Mb. Chromosome numbers are given in the facet headers.

Genetic variation in individual recombination rates

The average ACC per gamete ranged from 16.3 (PI) to 18.2 (LW) in males and 21.3 (LR) to 24.4 (LW) in females. The trait was normally distributed with a higher standard deviation in females than in males (Figure 6). Observations (gametes) with ACC of <6 or >50 were excluded. The distributions of ACCs are plotted in Figure S1. The h^2 ranged from 0.04 (SE = 0.01) (SY) to 0.07 (SE = 0.02) (DU) in males and 0.08 (SE = 0.01) (DU and PI) to 0.11 (SE = 0.01) (LR and LW) in females. For all breeds, the heritability is higher in females than in males and the phenotypic variance is substantially higher in females than in males for all breeds except Landrace. Most of the phenotypic variance is explained by the error term in all breeds and both sexes. Results from analysis on genetic variation in individual crossover counts can be found in table 4. Inbreeding had an effect on ACC, where higher inbreeding coefficients were associated with reduced crossover counts (Table S1).

Table 4. Results from variance component estimation of ACC and information on the dataset in the analysis.

Line	Sex	N_{FIDs}	N_{obs}	Mean (SD)	h^2 (SE)	V_p	V_e
LR	Female	4808	37148	21.3 (4.2)	0.11 (0.01)	18.53	16.40
LR	Male	319	37386	16.8 (3.5)	0.05 (0.05)	20.34	10.92
LW	Female	4695	41092	24.4 (4.6)	0.11 (0.01)	21.41	19.08
LW	Male	273	41104	18.2 (3.4)	0.06 (0.01)	11.44	10.81
DU	Female	1687	9268	21.4 (4.3)	0.08 (0.01)	18.35	16.55
DU	Male	192	9097	17.4 (3.3)	0.07 (0.02)	10.69	9.89
SY	Female	2633	25623	22.7 (4.4)	0.10 (0.01)	19.73	17.45
SY	Male	224	25622	17.6 (3.4)	0.04 (0.01)	11.38	10.83
PI	Female	1353	12101	21.8 (4.4)	0.08 (0.01)	19.08	16.89
PI	Male	196	12097	16.3 (3.3)	0.06 (0.02)	10.89	10.19

N_{FIDs} are the total number of FIDs (with repeated measures), N_{obs} is the total number of observations (meioses) in each sex and line. Mean is the mean ACC with standard deviations in parenthesis, h^2 is the heritability estimate with standard errors in parenthesis. V_p and V_e is the phenotypic variance and error variance, respectively.

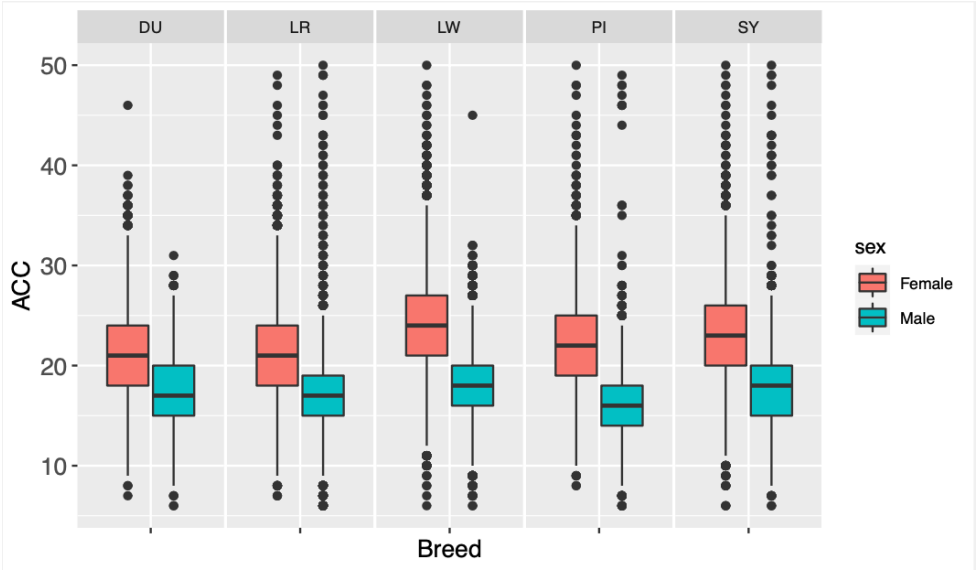


Figure 6 Distribution of autosomal crossover count.
The blue boxes are male count and red boxes are female counts for each line. The midline is the median and the box is from the 25th percentile to the 75th percentile.

GWAS

We found 14 loci significantly associated with mean ACC in females and 1 locus in males (Table 5). One region on chromosome 8 was in common to both males and females and showed the strongest association in both cases; this region is significantly associated with ACC in females in all breeds and in males in 4 out of 5 breeds, with P values for top SNPs ranging from 2.87×10^{-23} to 7.35×10^{-79} in females and 2.89×10^{-8} to 5.94×10^{-10} in males. The top SNPs are between 0.15 (LR female) and 1.13 (PI female) Mb away from the gene *RNF212* which has been found to be associated with individual recombination rates in several other mammals including humans (Kong *et al.*, 2008), cattle (Kadri *et al.*, 2016), Soay sheep (Johnston *et al.*, 2016) and pigs (Johnsson *et al.*, 2021). On chromosome 17, there is a peak for females in breeds LR, LW and SY, where the top SNPs in LR and SY are inside the gene *SYCP2* (Synaptonemal complex protein 2). The top SNP in the peak on chromosome 17 for LW females is 0.03 Mb from the *SYCP2* gene. On chromosome 6, there are two peaks in LR

females. On the first peak, the most significant SNP ($P = 1.36 \times 10^{-14}$) is located 2 kb away from the gene *PRDM7* (PR domain-containing protein 7). The second region on chromosome 6 also has a significant association in SY breed females, where the top SNPs are 0.03 (LR) and 0.29 (SY) Mb away from the gene *MSH4* (MutS homolog 4). LW females also have a significant peak on chromosome 6, but with no clear candidate gene in close proximity. On chromosome 5, there is a peak in LR females with the top SNP ($P = 1.54 \times 10^{-8}$) located inside the gene MEI1 (Meiotic Double-Stranded Break Formation Protein 1). Three breeds have a significant region on chromosome 7; the LR, LW and DU females. There are no candidate genes in immediate proximity to any of the top SNPs, but it should be noted that the genes *REC8* & *RNF212B*, which has been associated with individual recombination rates in several other mammal species are located on chromosome 7 (Sandor *et al.*, 2012; Petit *et al.*, 2017; Johnston, Huisman and Pemberton, 2018; Johnston, Stoffel and Pemberton, 2020). The analysis on LR females further shows significant associations in five more regions, on chromosome 4, 9, 12 and two on chromosome 15. However, we were not able to find any likely candidate genes involved in meiosis or recombination near any of these five peaks. Base pair positions and results from the GWA analysis for the top SNPs for all peaks and breeds can be found in Table 5 and the results are plotted for all lines and sex in Figure 7.

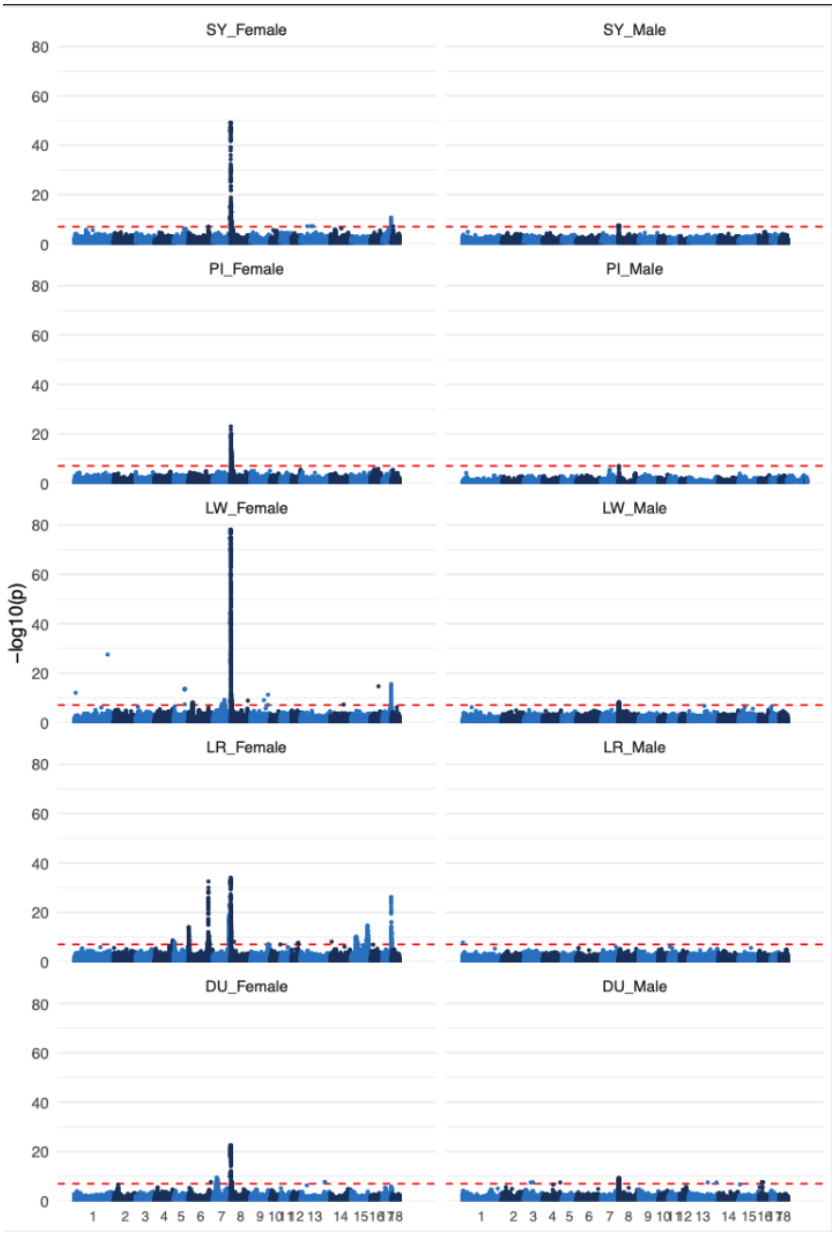


Figure 7 Genome-wide associations between markers and autosomal crossover count. Genome-wide associations between markers and autosomal crossover count for all breeds and sex separately. The dotted line is the statistical significance threshold = 0.05/Number of markers per analysis. The Y axis is the negative logarithm of the p-value and the x axis is the physical positions of the markers with alternating colors from autosome 1-18.

Table 5. Top SNPs in genome-wide association with ACC in all lines and both sex.

Chr	Line	Sex	SNP position	AF	BETA	SE	P	PVE	Candidate gene
4	LR	F	124741468	0.20	0.41	0.07	2.70E-09	0.01	
5	LR	F	6863929	0.45	0.32	0.06	1.54E-08	0.01	MEI1
6	LR	F	61205	0.18	0.55	0.07	1.36E-14	0.01	PRDM7
	LW	F	28051609	0.20	0.48	0.08	7.14E-09	0.01	
	LR	F	137566317	0.18	0.86	0.07	3.02E-33	0.03	MSH4
	SY	F	137825063	0.48	0.42	0.08	8.51E-08	0.01	
7	LW	F	78531203	0.39	0.43	0.07	6.83E-10	0.01	RNF212B/ Rec8*
	LR	F	113235081	0.41	0.56	0.06	9.39E-24	0.02	
	DU	F	23426839	0.34	-0.72	0.12	3.19E-10	0.02	
8	LR	F	562791	0.06	1.47	0.12	8.51E-35	0.03	RNF212
	LW	F	164462	0.29	-1.37	0.07	7.35E-79	0.07	
	LW	M	1197996	0.36	-0.72	0.12	4.76E-09	0.12	
	SY	F	164462	0.63	1.16	0.08	7.16E-50	0.08	
	SY	M	144871	0.65	0.60	0.11	2.89E-08	0.13	
	DU	F	178951	0.18	-1.41	0.14	2.87E-23	0.06	
	DU	M	794886	0.20	-1.03	0.17	5.94E-10	0.17	
	PI	F	1543451	0.18	1.34	0.13	9.50E-24	0.07	
	PI	M	1324636	0.24	0.87	0.15	3.11E-09	0.15	
9	LR	F	123194749	0.43	0.30	0.06	6.95E-08	0.01	-
12	LR	F	43697634	0.39	0.32	0.06	2.35E-08	0.01	-
15	LR	F	115153652	0.20	-0.55	0.07	2.03E-15	0.01	-
	LR	F	35743431	0.35	-0.37	0.06	1.02E-10	0.01	-
17	LR	F	59924616	0.27	-0.67	0.06	6.05E-27	0.02	
	LW	F	59845939	0.50	-0.55	0.07	2.58E-16	0.01	SYCP2
	SY	F	59911489	0.36	-0.53	0.08	2.32E-11	0.02	

SNP positions in base pairs, AF is the allele frequency of the effect allele. BETA is the additive effect size (i.e. the slope) of allele, SE is the standard error of BETA, P is the P value, and PVE is the proportion of phenotypic variance explained by the SNP.

*relatively distant from the peaks. Rec 8 position: 75118744-75139630. RNF212B position: 75796882-75829620

Discussion

In this study, we confirm previous findings by Tortereau et al. (Tortereau *et al.*, 2012) and Johnsson et al. (Johnsson *et al.*, 2021) that recombination rates in the domestic pig vary between sex, along the genome, between and within breeds and that there is a heritable component to the variation between individuals. Sex explains the majority of variation in genome-wide patterns and rates of recombination. We find that a locus corresponding to *RNF212* strongly influences recombination rate in all breeds and in both sexes. We also identified ~13 other loci associated with the trait in one or a few of the breeds, or only within one sex. Some of these loci corresponded to genes with known functions in meiosis, including *SYCP2*, *PRMD7* and *MSH4*. In the following sections, we discuss the results in more detail, provide context to how they may aid in better understanding the genetic mechanisms of recombination rate, and the implications of our findings in commercial breeding populations.

Genome-wide rates vary between breeds, but not recombination landscapes.

The genome-wide rates in our study are more in line with Tortereau et al.(Tortereau *et al.*, 2012) than Johnsson et al. (Johnsson *et al.*, 2021), where estimates are slightly higher. In the latter study, potential overestimation of total genetic lengths is reported. The pattern of recombination rate along the chromosomes are well conserved across breeds (Figures 3 & 4). It is possible that similarity in patterns could be an artefact of differences in marker densities or due to biological explanations such as runs of homozygosity in some areas, making it difficult to detect recombination events. However, there is a substantial difference in recombination patterns between the sexes that would be subject to the same artifacts. This suggests that the observed patterns accurately reflect the variation in recombination rates along the genome. The genome-wide rate of recombination rate also differed between breeds

in agreement with previous studies (Tortereau *et al.*, 2012; Mary *et al.*, 2014; Johnsson *et al.*, 2021). This difference is mainly driven by the largest chromosomes, where it is more common to have more than the obligate crossover due to reduced effects of crossover interference(Otto and Payseur, 2019). The difference in overall rate between the breeds is larger in females. In the animal models there is an effect of inbreeding, where higher inbreeding coefficients were associated with reduced crossover counts (Table S1). The most likely reason for this observation is that higher levels of recent inbreeding can lead to longer runs of homozygosity, meaning that double crossovers occurring within a run cannot be detected despite the effects of crossover interference (Johnston *et al.*, 2016).

Recombination landscapes and rates differ between the sexes.

There is a substantial difference in recombination rates between the sex in all five breeds in this study with females showing 1.28 times more recombination than in males. This direction of heterochiasmy is common in mammalian recombination, and whilst there has been great interest in determining the mechanisms underpinning to this sexual dimorphism, there are few compelling explanations (Brandvain and Coop, 2012) as the direction and degree can vary even between closely related species. There is some evidence of a molecular basis to heterochiasmy. In humans, there is evidence for selection against non-recombinant chromatids in meiosis II (Ottolini *et al.*, 2015). In addition, there can be sex differences in the packing of the chromosomes in the early prophase of meiosis, with differences in synaptonemal complex length between the sexes correlating with the recombination rate in humans (Tease and Hultén, 2004), bovid species (Ruiz-Herrera *et al.*, 2017) and even *Arabidopsis* (Capilla-Pérez *et al.*, 2020). There may also be potential evolutionary drivers of heterochiasmy; such as differences in haploid selection between the sexes (Lenormand and Dutheil, 2005), sexual dimorphism and sperm competition (Mank, 2009), and the effects of

meiotic drive (Brandvain and Coop, 2012), but evidence for these hypotheses remains limited. It is not unlikely that the some of these mechanisms, particularly related to the synaptonemal complex, explain the sex difference in the domestic pig. In our study, it may be that some differences are driven by differences in the number of male and female FIDs due to the breeding structure of the pigs; one male has on average around 150 offspring in the dataset, whereas a female has only around 7. However, the total number of meiosis investigated in this study is the same between the sexes; difference in rates between sexes are still found in studies where the number of males and females are almost the same e.g. in humans (Halldorsson *et al.*, 2016).

There is genetic variation in individual crossover count

The heritability estimates of the trait are low, but significant, in agreement with Johnsson et al. and Lozada-Soto et al. [15, 42]. Heritability is lower in males than in females, which is consistent with what has been found in many other mammal studies (Fledel-Alon *et al.*, 2011; Johnston *et al.*, 2016; Johnston, Huisman and Pemberton, 2018; Johnsson *et al.*, 2021). The standard errors for the heritability estimates are higher for the males in general, which may be due to the smaller number of unique male FIDs and lower heritability values leading to more uncertainty in the estimates (Table 4). Most of the phenotypic variance is explained by the error terms (Table 4). One explanation may be that each recombination occurring in the meiotic division has a 50:50 chance of segregating into a particular gamete, leading to variance in the number of recombinations in the sampled gamete. In addition, we can only measure recombination in gametes that result in live animals, which excludes all other products of meiosis. Therefore, as the analysis is biased towards a sample of successful gametes, we may be underestimating the genetic variance. A study of recombination in all products of the female meiosis in human showed selection against non-recombinant

chromatids in meiosis II (Ottolini *et al.*, 2015). It would be of great interest to study more of the products from meiosis in these pig breeds to look for signs of selection between gametes.

Genome-wide associations with autosomal crossover count

The strongest association with individual ACC is on a locus close to the gene *RNF212*, which is a gene that have been found to be associated with individual recombination rates in several studies in mammals previously, including other pig breeds (Kong *et al.*, 2008, 2014; Sandor *et al.*, 2012; Kadri *et al.*, 2016; Johnston, Huisman and Pemberton, 2018; Johnsson *et al.*, 2021). This region is the only one that shows association with individual ACC in all five breeds. A study by Reynolds et al (Reynolds *et al.*, 2013) established that *RNF212* has a dosage sensitive effect on recombination rates, and that *RNF212* knockout mice were sterile, implying a critical role in chromosome segregation and fertility. In the region on chromosome 8 close to *RNF212*, LR and PI have a low frequency of the allele for increased crossover counts (Table 5), these are also the two breeds with the lowest mean ACC (Table 4.) However, the higher rates in LW does not seem to be explained by the allele frequencies in this region. It should also be noted that none of the top SNPs in this region for the five breeds are in immediate proximity to the gene, meaning that the linkage phase with the causal variant may differ between the breeds.

LR and SY females have a top SNP inside the gene *SYCP2* and LW females a top SNP close by (0.03 Mb). *SYCP2* is a gene coding for a protein which is part of the axial elements that the chromatids are organized along in the early prophase of meiosis. Together with SYCP3, the SYCP2 proteins become the lateral elements of the synaptonemal complex during synapsis (Fraune *et al.*, 2012). The effect of this gene on ACC may be linked to the correlation between synaptonemal complex length and genetic length discussed earlier.

Halldorsson et al (Halldorsson *et al.*, 2019) found an association for the same trait in humans with a gene encoding one of the other synaptonemal complex proteins, namely *SYCE1*, and another study in pigs by Johnsson et al (Johnsson *et al.*, 2021) also found a peak very close to the lead SNPs on chromosome 17 in this study.

The first peak on chromosome 6 in LR females is potentially associated with a gene that is annotated as *PRDM7* on the NCBI annotation release 106, although in the study by johnsson et al (Johnsson *et al.*, 2021) this gene is identified as *PRDM9*. *PRDM9* is a gene that has been associated with recombination hotspot positioning in mammals but has also been associated with genome wide rates in for example cattle, possibly as a result of differences in hotspot abundance corresponding to different *PRDM9* alleles (Ma *et al.*, 2015). Blazer et al. (Blazer *et al.*, 2016) reports high homology between *PRDM9* and *PRDM7*. The gene *MSH4*, coding for a meiosis specific protein essential for reciprocal recombination (Paquis-Flucklinger *et al.*, 1997) is close to a significant peak for LR and SY females. This gene has been associated with individual recombination rates in humans (Halldorsson *et al.*, 2019) and pigs (Johnsson *et al.*, 2021). The *MEI1* gene encodes a double stranded break formation protein in meiosis (Kumar and de Massy, 2010), but to our knowledge has not been previously reported as associated with individual recombination rates.

Several of the genome wide hits reported are only present in one or two breeds or only in females. This suggests that there is sexual dimorphism in the genetic architecture of ACC in pigs, but it might also be explained by the different number of focal individuals for males and females. The difference between breeds may also be due to differences in allele frequencies; LR has been a closed breeding line since 1958 and may very well differ in allele frequencies in some of these regions either due to drift or as a consequence of selection. This may explain

all the GWAS peaks that are unique to LR. In a study by Lozada-Soto et al. (Lozada-Soto *et al.*, 2021) genome-wide associations with individual recombination rates were also investigated, but with none of the regions identified overlapping with the findings in our study. However, the study by Johnsson et al. (Johnsson *et al.*, 2021) agrees well with our findings, with overlapping significant regions on chromosome 4, 6, 8 and 17.

What is the impact of recombination rate variation in the domestic pig?

The difference in recombination rate is relatively large between some of these breeds considering that they are closely related in an evolutionary perspective; combined the heritable variation and oligogenic architecture may suggest that recombination rates have the potential to change rapidly. As there is genetic variation in individual recombination rates in the pig breeds in this study, there might be a potential to increase genetic variance by increasing recombination rates. However, for there to be a significant increase in production trait values in response to selection, recombination rates may have to be 10 or 20 times multiplied, as shown by Battagin et al. (Battagin *et al.*, 2016). It is unlikely that a breeding program would wish to put a strong emphasis on recombination at the cost of other traits. There may also be costs associated with increased recombination that will outweigh the benefits; for instance, high recombination rates have been associated with cancer in human (Mao *et al.*, 2009). In addition, and as previously mentioned, recombination can also break up beneficial allele combinations previously built up by selection(Barton and Charlesworth, 1998). A review of recombination rate variation across a broad selection of taxa by Ritz et al. (Ritz, Noor and Singh, 2017) found that the magnitude of variation in recombination rates is equivalent across taxa and that mechanisms, such as crossover interference, lead to an upper limit that is universal across most species studied, suggesting that large numbers of

recombinations are not beneficial despite the potential for increased genetic variance, possibly due to biological consequences that remain to be fully understood.

Conclusions

In this study, we find that both patterns and overall rates of recombination differ between the sexes in five domestic pig breeding populations. The overall rate varies between the breeds, but the patterns of recombination rate along the chromosomes is well conserved across breeds. We show that there is a heritable component to the genetic diversity, but that the trait has a large error variance. Studies including more products of the meiosis would be of interest to potentially explain more of the phenotypic variance in these populations as well as in other populations. We find genes associated with individual recombination rates found in several other studies suggesting that some of the genetic architecture of the trait is well conserved across a large number of species, but we also find associations with genes that have not yet been reported. Whether the difference in overall rates and genetic architecture of recombination rates between the breeds is due to indirect selection or merely random drift is not answered and remains one of great interest to study in these breeds. Further studies should also look at the mechanisms leading to the substantial difference in recombination rates between sex as well as how individual rates may relate to reproductive traits. This study provides an example of the state of recombination rates in a domesticated species under strong selection, as well as how it may differ between relatively closely related breeds, contributing to the understanding of variation in recombination rate in mammals in general.

List of abbreviations

ACC = Autosomal crossover count

SNP = single nucleotide polymorphism

FID = focal individual

LR = Landrace

LW = Large White

DU = Duroc

SY = Synthetic

PI = Piétrain

Declarations

Ethics approval and consent to participate

All samples included in the study are derived from routine work at Norsvin and Topigs Norsvin.

Consent for publication

Not applicable

Availability of data and materials

The data that support the findings of this study are available from Norsvin and Topigs Norsvin but restrictions apply to the availability of these data, which were used under license for the current study, and thus are not publicly available. However, data are available from the authors upon reasonable request and with permission of Norsvin and Topigs Norsvin.

Competing interests

AG is employed in Topigs Norsvin. All authors declare that the results are presented in full and as such present no conflict of interest. The other authors declare to have no competing interests for this study.

Funding

SEJ is supported by a Royal Society University Research Fellowship (UF150448) funded by the The Royal Society. CB is funded by a phd grant from the Norwegian University of life sciences.

Authors' contributions

PB conceived the study. SEJ designed the pipeline. CB conducted the analysis and wrote the paper. AG filtered and imputed data. All authors provided critical feedback and helped shape the research, analysis and manuscript.

Acknowledgements

We want to thank Norsvin and Topigs Norsvin for providing the data for this study. Pasi Rastas answered several questions regarding the Lepmap3 software used for linkage mapping. Most of the analysis in this study have been executed on the Orion cluster at NMBU (https://orion.nmbu.no). We also thank John McAuley and members of the Johnston lab for discussions.

Author's information (optional)

References

Arbeithuber, B. *et al.* (2015) “Crossovers are associated with mutation and biased gene conversion at recombination hotspots,” *Proceedings of the National Academy of Sciences of the United States of America*, 112(7), pp. 2109–2114. doi:10.1073/pnas.1416622112.

Barton, N.H. and Charlesworth, B. (1998) *Why Sex and Recombination?* doi:10.1126/science.281.5385.1986.

Battagin, M. *et al.* (2016) “Effect of manipulating recombination rates on response to selection in livestock breeding programs,” *Genetics Selection Evolution*, 48(1). doi:10.1186/s12711-016-0221-1.

Blazer, L.L. *et al.* (2016) “PR domain-containing protein 7 (PRDM7) is a histone 3 lysine 4 trimethyltransferase,” *Journal of Biological Chemistry*, 291(26), pp. 13509–13519. doi:10.1074/jbc.M116.721472.

Brandvain, Y. and Coop, G. (2012) “Scrambling eggs: Meiotic drive and the evolution of female recombination rates,” *Genetics*, 190(2), pp. 709–723. doi:10.1534/genetics.111.136721.

Burt, A. and Bell, G. (1987) “Mammalian chiasma frequencies as a test of two theories of recombination,” *Nature*, 326, pp. 803–805.

Capilla-Pérez, L. *et al.* (2020) “The synaptonemal complex imposes crossover interference and heterochiasmy in Arabidopsis,” *PNAS* [Preprint]. doi:10.1073/pnas.2023613118/-/DCSupplemental.

Chang, C.C. *et al.* (2015) “Second-generation PLINK: Rising to the challenge of larger and richer datasets,” *GigaScience*, 4(1). doi:10.1186/s13742-015-0047-8.

Charlesworth, B. and Barton, N.H. (1996) “Recombination load associated with selection for increased recombination,” *Genetical Research*, 67(1), pp. 27–41. doi:10.1017/s0016672300033450.

Cleveland, W.S. (1979) *Robust Locally Weighted Regression and Smoothing Scatterplots*, *Source: Journal of the American Statistical Association*.

Coop, G. and Przeworski, M. (2007) “An evolutionary view of human recombination,” *Nature Reviews Genetics*, pp. 23–34. doi:10.1038/nrg1947.

Dumont, B.L., Broman, K.W. and Payseur, B.A. (2009) “Variation in genomic recombination rates among heterogeneous stock mice,” *Genetics*, 182(4), pp. 1345–1349. doi:10.1534/genetics.109.105114.

Fledel-Alon, A. *et al.* (2009) “Broad-scale recombination patterns underlying proper disjunction in humans,” *PLoS Genetics*, 5(9). doi:10.1371/journal.pgen.1000658.

Fledel-Alon, A. *et al.* (2011) “Variation in human recombination rates and its genetic determinants,” *PLoS ONE*, 6(6). doi:10.1371/journal.pone.0020321.

Fraune, J. *et al.* (2012) “The mammalian synaptonemal complex: Protein components, assembly and role in meiotic recombination,” *Experimental Cell Research*. Academic Press Inc., pp. 1340–1346. doi:10.1016/j.yexcr.2012.02.018.

Halldorsson, B. v. *et al.* (2016) “The rate of meiotic gene conversion varies by sex and age,” *Nature Genetics*, 48(11), pp. 1377–1384. doi:10.1038/ng.3669.

Halldorsson, B. v. *et al.* (2019) “Human genetics: Characterizing mutagenic effects of recombination through a sequence-level genetic map,” *Science*, 363(6425). doi:10.1126/science.aau1043.

Hassold, T. *et al.* (1995) *Recombination and Maternal Age-Dependent Nondisjunction: Molecular Studies of Trisomy 16*, *Am. J. Hum. Genet*.

Johnsson, M. *et al.* (2021) “Genetic variation in recombination rate in the pig,” *Genetics Selection Evolution*, 53(1). doi:10.1186/s12711-021-00643-0.
Johnston, S., Stoffel, M. and Pemberton, J. (2020) “Variants at RNF212 and RNF212B are associated with recombination rate variation in Soay sheep (Ovis aries).” doi:10.1101/2020.07.26.217802.
Johnston, S.E. *et al.* (2016) “Conserved genetic architecture underlying individual recombination rate variation in a wild population of soay sheep (Ovis aries),” *Genetics*, 203(1), pp. 583–598. doi:10.1534/genetics.115.185553.
Johnston, S.E., Huisman, J. and Pemberton, J.M. (2018) “A genomic region containing REC8 and RNF212B is associated with individual recombination rate variation in a wild population of red deer (Cervus elaphus),” *G3: Genes, Genomes, Genetics*, 8(7), pp. 2265–2276. doi:10.1534/g3.118.200063.
Kadri, N.K. *et al.* (2016) “Coding and noncoding variants in HFM1, MLH3, MSH4, MSH5, RNF212, and RNF212B affect recombination rate in cattle,” *Genome Research*, 26(10), pp. 1323–1332. doi:10.1101/gr.204214.116.
Koehler, K.E. *et al.* (1996) *Recombination and nondisjunction in humans and flies*, *Human Molecular Genetics*. Oxford University Press. Available at: https://academic.oup.com/hmg/article/5/Supplement_1/1495/663563.
Kong, A. *et al.* (2008) “Sequence Variants in the RNF212 Gene Associate with Genome-Wide Recombination Rate,” *Science*, 319(5868), pp. 1398–1401. doi:10.1126/science.1152422.
Kong, A. *et al.* (2014) “Common and low-frequency variants associated with genome-wide recombination rate,” *Nature Genetics*, 46(1), pp. 11–16. doi:10.1038/ng.2833.
Kumar, R. and de Massy, B. (2010) “Initiation of meiotic recombination in mammals,” *Genes*, pp. 521–549. doi:10.3390/genes1030521.
Lenormand, T. and Dutheil, J. (2005) “Recombination difference between sexes: A role for haploid selection,” in *PLoS Biology*, pp. 0396–0403. doi:10.1371/journal.pbio.0030063.
Li, H. and Durbin, R. (2009) “Fast and accurate short read alignment with Burrows-Wheeler transform,” *Bioinformatics*, 25(14), pp. 1754–1760. doi:10.1093/bioinformatics/btp324.
Lozada-Soto, E.A. *et al.* (2021) “Evidence for recombination variability in purebred swine populations,” *Journal of Animal Breeding and Genetics*, 138(2), pp. 259–273. doi:10.1111/jbg.12510.
Ma, L. *et al.* (2015) “Cattle Sex-Specific Recombination and Genetic Control from a Large Pedigree Analysis,” *PLoS Genetics*, 11(11). doi:10.1371/journal.pgen.1005387.
Madsen, P. *et al.* (2014) “DMU - A Package for Analyzing Multivariate Mixed Models in quantitative Genetics and Genomics,” in *Journal of Applied Statistics*. Routledge, pp. 1286–1306.
Mank, J.E. (2009) “The evolution of heterochiasmy: The role of sexual selection and sperm competition in determining sex-specific recombination rates in eutherian mammals,” *Genetics Research*, 91(5), pp. 355–363. doi:10.1017/S0016672309990255.
Mao, Z. *et al.* (2009) “DNA repair by homologous recombination, but not by nonhomologous end joining, is elevated in breast cancer cells,” *Neoplasia*, 11(7), pp. 683–691. doi:10.1593/neo.09312.
Mary, N. *et al.* (2014) “Meiotic recombination analyses of individual chromosomes in male domestic pigs (Sus scrofa domestica),” *PLoS ONE*, 9(6). doi:10.1371/journal.pone.0099123.
Munoz-Fuentes, V. *et al.* (2015) “Strong artificial selection in domestic mammals did not result in an increased recombination rate,” *Molecular Biology and Evolution*, 32(2), pp. 510–523. doi:10.1093/molbev/msu322.
Otto, S.P. and Payseur, B.A. (2019) “Annual Review of Genetics Crossover Interference: Shedding Light on the Evolution of Recombination.” doi:10.1146/annurev-genet-040119.

Ottolini, C.S. *et al.* (2015) "Genome-wide maps of recombination and chromosome segregation in human oocytes and embryos show selection for maternal recombination rates," *Nature Genetics*, 47(7), pp. 727–735. doi:10.1038/ng.3306.
Paquis-Flucklinger, V. *et al.* (1997) *Cloning and Expression Analysis of a Meiosis-Specific MutS Homolog: The Human MSH4 Gene*, *GENOMICS*.
Petit, M. *et al.* (2017) "Variation in recombination rate and its genetic determinism in sheep populations," *Genetics*, 207(2), pp. 767–784. doi:10.1534/genetics.117.300123.
R Core Team (2020) "R: A Language and Environment for Statistical Computing." Vienna.
Rastas, P. (2017) "Lep-MAP3: Robust linkage mapping even for low-coverage whole genome sequencing data," *Bioinformatics*, 33(23), pp. 3726–3732. doi:10.1093/bioinformatics/btx494.
Reynolds, A. *et al.* (2013) "RNF212 is a dosage-sensitive regulator of crossing-over during mammalian meiosis," *Nature Genetics*, 45(3), pp. 269–278. doi:10.1038/ng.2541.
Ritz, K.R., Noor, M.A.F. and Singh, N.D. (2017) "Variation in Recombination Rate: Adaptive or Not?," *Trends in Genetics*. Elsevier Ltd, pp. 364–374. doi:10.1016/j.tig.2017.03.003.
Ruiz-Herrera, A. *et al.* (2017) "Recombination correlates with synaptonemal complex length and chromatin loop size in bovids—insights into mammalian meiotic chromosomal organization," *Chromosoma*, 126(5), pp. 615–631. doi:10.1007/s00412-016-0624-3.
Sandor, C. *et al.* (2012) "Genetic variants in REC8, RNF212, and PRDM9 influence male recombination in cattle," *PLoS Genetics*, 8(7). doi:10.1371/journal.pgen.1002854.
Sherman, S.L. *et al.* (1991) *Trisomy 21: Association between Reduced Recombination and Nondisjunction*, *Am. J. Hum. Genet*.
Stapley, J. *et al.* (2017) "Variation in recombination frequency and distribution across eukaryotes: Patterns and processes," *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1736). doi:10.1098/rstb.2016.0455.
Tease, C. and Hultén, M.A. (2004) "Inter-sex variation in synaptonemal complex lengths largely determine the different recombination rates in male and female germ cells," *Cytogenetic and Genome Research*, 107(3–4), pp. 208–215. doi:10.1159/000080599.
Tortereau, F. *et al.* (2012) "A high density recombination map of the pig reveals a correlation between sex-specific recombination and GC content," *BMC Genomics*, 13(1). doi:10.1186/1471-2164-13-586.
Wang, R.J. *et al.* (2017) "Recombination rate variation in mice from an isolated island," *Molecular Ecology*, 26(2), pp. 457–470. doi:10.1111/mec.13932.
Yang, J. *et al.* (2011) "GCTA: A tool for genome-wide complex trait analysis," *American Journal of Human Genetics*, 88(1), pp. 76–82. doi:10.1016/j.ajhg.2010.11.011.

Supplementary

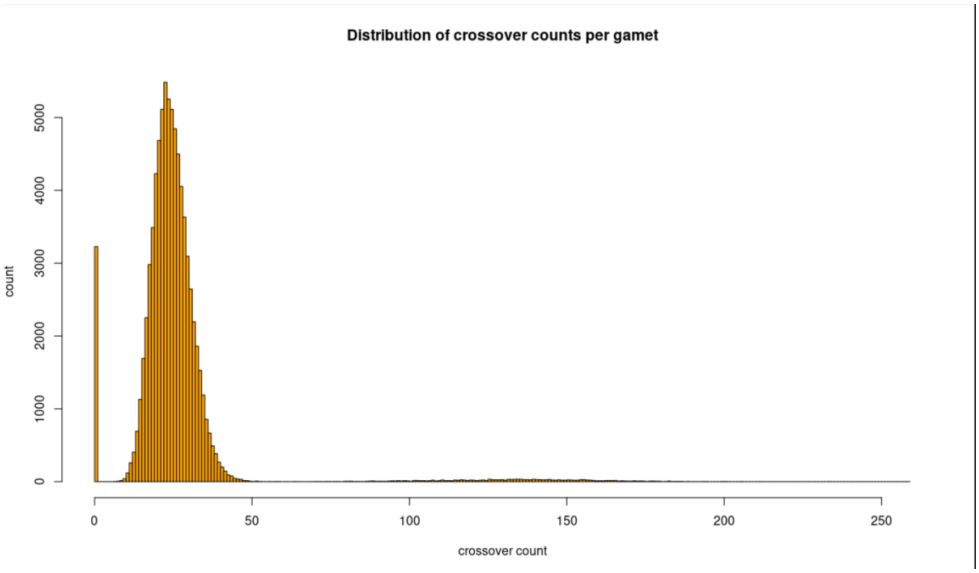


Figure S1 Distribution of crossover counts per gamete before filtering (LR)
The distribution of crossover count per gamete plotted for the LR breed as an example. The x-axis is the autosomal crossover count for a gamete and the y-axis is the count of how many gametes had a particular crossover count.

Table S2 Effect of inbreeding on ACC

Line	Sex	N FIDs	mean F	E_BLUE F (SE)
LR	Female	4808	0.01	-2.48 (0.66)
LR	Male	319	0.04	2.65 (4.47)
LW	Female	4695	0.29	-5.53 (1.05)
LW	Male	273	0.31	-3-34 (1.88)
DU	Female	1687	0.05	-3.41 (0.95)
DU	Male	192	0.07	-1.87 (1.02)
SY	Female	2633	-0.03	-3.56 (0.83)
SY	Male	224	0.03	-0.64 (1.12)
PI	Female	1353	-0.04	-4.54 (1.18)
PI	Male	196	0.01	-3.86 (1.33)

N_{FIDs} are the total number of unique FIDs, mean F is the mean inbreeding coefficient, and E_BLUE F is the estimated effect of inbreeding with standard errors in parenthesis.

Paper II.

Individual recombination rates are heritable and associated with RNF212b, NEK9, MLH3 and CEP55 in Norwegian Red cattle

Brekke, C., Johnston, S.E., Gjuvsland, A.B., Berg, P. (2022).
Manuscript

Individual recombination rates are heritable and associated with RNF212B, NEK9, MLH3 and CEP55 in Norwegian Red cattle

Cathrine Brekke[1*], Susan E. Johnston[2], Arne B. Gjuvsland[3], Peer Berg[1]

[1] Norwegian University of Life Sciences, Department of Animal and Aquacultural sciences, Oluf Thesens vei 6, 1433 Ås, Norway

[2] Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh, Edinburgh, EH9 3FL, United Kingdom

[3] Geno SA, Storhamargata 44, 2317 Hamar, Norway

Abstract

Meiotic recombination is an important evolutionary mechanism that breaks up linkage between loci and create novel haplotypes for selection to act upon. Understanding the genetic control of variation in recombination rates is therefore of great interest in both natural populations and domestic breeding population. In this study we use pedigree information and medium density (~50K) genotyped data in a large cattle (*Bos Taurus*) breeding population in Norway, Norwegian Red cattle, to investigate recombination rate variation between sex and individuals. Sex specific linkage mapping show higher rates in males than in females (total genetic length of autosome = 2492.9 cM in males and 2308.9 cM in females), However, distribution of recombination along the genome show little variation between males and females compared to other species. The heritability of autosomal crossover count was low but significant in both sex (h^2 = 0.04 and 0.09 in males and females, respectively). We identify three loci associated with variation in individual crossover counts in female, two are close to the candidate genes RNF212B and CEP55 and one is close to both MLH3 and NEK9. All

four genes have been associated with recombination rates in other cattle breeds, and RNF212B in several other mammals. Our study contributes to the understanding of how recombination rates are controlled and how they may vary between closely related breeds as well as between species.

Introduction

During eukaryotic meiosis, homologous chromosomes line up and exchange segments of DNA in a process known as recombination. The process is well conserved across taxa as it has a vital role in assuring homologous chromosome pairing during prophase I, meaning that the gametes resulting from the meiosis have the correct number of chromosomes (Fledel-Alon et al., 2011; Lenormand et al., 2016); human studies have shown that a lack of recombination can lead to aneuploidy, i.e. the incorrect number of chromosomes in gametes (Fledel-Alon et al., 2009; Sherman et al., 1991). From an evolutionary perspective, recombination is also an important mechanism as it breaks up linkage between alleles from genes located on the same chromosomes and creates new haplotypes for selection to act upon (Barton & Charlesworth, 1998; Sved & Hill, 2018). However, recombination may also break up beneficial linkage built up over many generations; there are also other costs associated with recombination, such as increased mutation rate associated with double strand break repair (Arbeithuber et al., 2015). These costs and benefits are thought to be the reason why the number of crossovers per chromosome seems to have a well conserved upper and lower limit across species (Ritz et al., 2017). Still, there is substantial variation in recombination rates between species, and between sexes and individuals within breeds and populations of the same species (Lenormand et al., 2016; Stapley et al., 2017). In addition, many species show large variation in rates of recombination along the chromosomes, including

recombination “hotspots/coldspots”, with patterns often differing between the sexes (Halldorsson et al., 2016; Johnsson et al., 2021; S. E. Johnston et al., 2016, 2017; Kong et al., 2008; Petit et al., 2017).

There has been an increasing interest in understanding variation in individual recombination rates over the last decade. Several studies have been conducted in model species, such as house mice (Booker et al., 2017; Dumont et al., 2009; Wang et al., 2017) and *Drosophila melanogaster* (Hunter et al., 2016; Samuk et al., 2020; Winbush & Singh, 2021), natural populations such as red deer (S. E. Johnston et al., 2018) and Soay sheep (S. E. Johnston et al., 2016) and domesticated species such as pigs (Johnsson et al., 2021), chickens (Z. Weng et al., 2019), sheep (Petit et al., 2017) and several cattle breeds (Kadri et al., 2016; Ma, O’Connell, et al., 2015; Sandor et al., 2012; Shen et al., 2018). There is a heritable component to individual recombination rates in all species studied (i.e. the proportion of phenotypic variance explained by additive genetic effects), ranging from around 5% in pigs (Johnsson et al., 2021) to as high as 46% in mice (Dumont et al., 2009). Studies on the genetic architecture of individual recombination rates has led to discovery of several meiotic genes associated with individual recombination rate variation. This includes *RNF212* and *Rec8* that have been found in various mammal species (S. Johnston et al., 2020; S. E. Johnston et al., 2018; Kadri et al., 2016; Kong et al., 2008; Sandor et al., 2012), but also genes such as *SYCP2*, *HEI10*, *MEIOB* and several others that are only found in one or a couple of species and populations (Halldorsson et al., 2019a; Johnsson et al., 2021; Petit et al., 2017). Most of the vertebrate species studied show striking sex differences in the amount and location of recombination in the genome, known as heterochiasmy, but the direction and magnitude can be remarkably variable across species(Lenormand & Dutheil, 2005; Mank, 2009). However, despite the potential for genome-wide rates to respond rapidly to selection,

the direction of heterochiasmy is conserved across relatively distantly related breeds and populations within species (see for example sheep; (S. E. Johnston et al., 2016; Petit et al., 2017), cattle (Shen et al., 2018) and pigs (Johnsson et al., 2021)). But sex spesific studies of populations and breeds within species remains limited.

Cattle (Bos Taurus) were domesticated around 10 000 years ago, and is believed to be one of the cornerstones of the Neolithic revolution (Götherström et al., 2005). Today, dairy and beef breeds are spread around the world and breeding programs with extensive pedigrees along with genotype information exist for a large number of individuals. Recombination rate variation has indeed been studied in several other Cattle breeding populations previously, like Holstein(Ma, O'Connell, et al., 2015) Holstein Freysian, Jersey ((Sandor et al., 2012), Brown Swiss, Ayrshire (Shen et al., 2018) and beef cattle like Angus and Limousin (Z. Q. Weng et al., 2014). Only some of the studies look at sex sepesific rates of recombination, but the ones that do all find higher genome wide rates in males and some indications for sex specific control of the trait ((Kadri et al., 2016; Ma, O'Connell, et al., 2015; Shen et al., 2018). Several genes are identified as genetic drivers of recombination rate variation in cattle, including PRDM9 (Ma, O'Connell, et al., 2015), which is usually associated with recombination hotspot usage in other species (Paigen & Petkov, 2018).

Recombination is one of the main sources of novel haplotypic variation, which is a prerequisite for selection response and genetic gain in animal and plant breeding. Therefore, it is of great interest to understand how and why recombination rates vary between individuals, sexes and along the genome. For example, if there is a heritable component to individual recombination rate, it presents the opportunity to select for higher rates, and thereby potentially help quantitative traits respond faster to selection (Battagin et al., 2016).

Click or tap here to enter text. More practically, having detailed information on how recombination rates vary within a breeding population may also be important for genomic prediction (Gao et al., 2018) and selection on quantitative trait loci (Lotterhos, 2019).

Here, we study the genetic architecture and variation in individual autosomal crossover count (ACC) using a large genomic dataset from more than 110 000 of Norwegian Red cattle *(Bos Taurus)*, a breed of dairy cattle that comprises the majority of all cattle in Norway. Our objectives were to: (1) create sex-specific high density linkage maps to understand variation in recombination landscapes; and (2) to determine the heritability of individual ACC and identify potential loci associated with the trait.

Material and Methods

Study population and Genotype data

In this study we used genotype data and pedigree information from Norwegian Red cattle, which is the most common cattle breed used for milk production in Norway. Genotype data was available from five different SNP arrays developed for cattle: Affymetrix 54K (customized chip), Illumina NRF v2 (customized chip), Illumina BovineSNP50 v1, Illumina BovineSNP50 v2.0 and Illumina BovineHD. There were 35,880 common SNPs genotyped for 110,555 individuals across these five arrays. The physical positions of the SNPs were determined based on the ARS-UCD1.2 reference genome. The data was filtered to remove markers with missing call rates exceeding 0.1 or a Hardy-Weinberg equilibrium exact test p-value$<10^{-6}$. This genotype dataset will be referred to as the 35K dataset. All individuals were also imputed to a set of 617,739 SNP markers on the BovineHD chip; this data-set was used for the genome wide association analysis only. This set will further be referred to as the 600K dataset. Only autosomal SNP markers were included in the study.

Quantification of individual crossovers

The pedigree was ordered into three generation full-sib families, comprised of each unique sire and dam mating combination with their offspring and parents (Figure 1). As recombination rates are estimated for meioses that occur in the sire and dam, they are hereafter referred to as the focal individuals or FIDs. This full-sib family format allowed for phasing of the gametes transmitted from the FIDs to the offspring and detection of crossovers occurring during meiosis in the FIDs. Whilst an individual can be present in several families, i.e., if mated to other individuals or as grandparent or offspring in a different family, this study design meant that each unique meiosis in an FID is only calculated and analysed once. The total number of full-sib families was 19 603. Due to the breeding structure of cattle, most

full-sib families only have one offspring, with a maximum of five offspring. In total, autosomal crossover counts were calculated for 603 unique bulls with 19 861 associated offspring and 14 815 unique cows with 19 824 associated offspring.

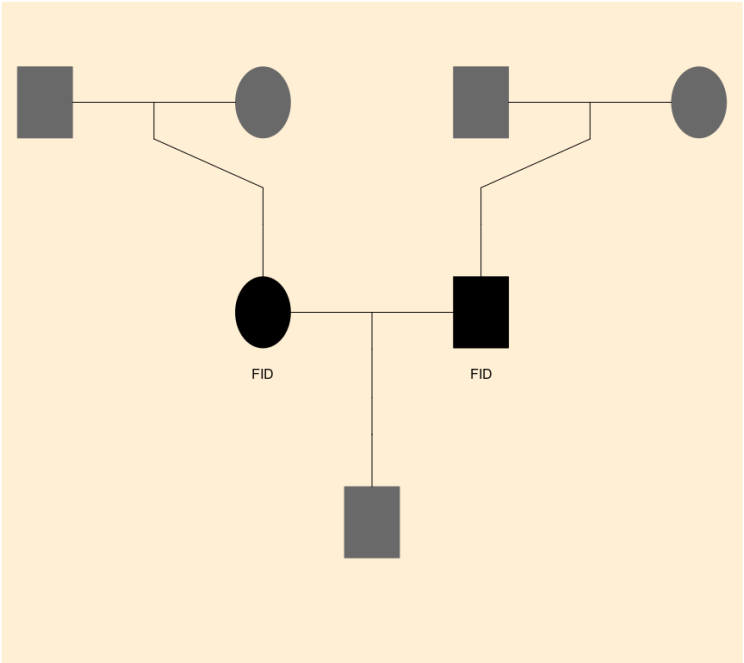


Figure 1. The three-generation full-sib family structure used to calculate autosomal crossover count (ACC). This structure allows quantification of autosomal crossovers in the gamete transmitted from the FIDs (black) to the offspring.

Linkage mapping & estimation of autosomal crossover counts.

Sex-specific linkage maps were created with Lepmap3 (Rastas, 2017). Marker order was fixed relative to their physical positions on the ARS-UCD1.2 reference genome and linkage maps were created for each chromosome separately. It should be noted that all cattle autosomes are acrocentric, with the centromere occurring at the beginning of the genomic

sequence for each chromosome. The *filtering2* module was run to filter markers based on segregation distortion, with the argument *dataTolerance* = 0.01 as suggested for multi-family datasets. The *Seperatechromosomes2* module is used to split markers into linkage groups *de novo*; here, this module was used as a quality control step to exclude any markers that were not assigned to their expected linkage group (LOD score < 5). The Haldane mapping function option was used to calculate the centimorgan (cM) positions of the SNP markers in the *Ordermarkers2* module. The number of markers in the final linkage maps can be found in Table 2. To quantify individual recombination rates, the number of crossovers per autosomal chromatid in each offspring was estimated using the output from the *Ordermarkers2* module. The crossovers were then summed across all autosomes in the offspring and defined as the autosomal crossover count (ACC), which was then assigned to the FID in which the crossover events occurred.

Heritability of individual recombination rates

Genetic variation for the trait autosomal crossover count (ACC) was estimated with a repeatability model in DMU v 6 (Madsen et al., 2014) as FIDs with several offspring either within one family or across families had multiple observations for the trait. We used the restricted maximum likelihood (REML) method with the average information (AI) algorithm and fitted the following model:

$$\mathbf{Y = sex + b_1*age + b_2*het + id1 + id2 + e}$$

where Y is the ACC, **sex** is the fixed effect of sex, $\mathbf{b_1}$ is the fixed regression of age of the FID when the offspring is born (from ages 1 to 13), $\mathbf{b_2}$ is the regression of ACC on **het** of the FID,

and **e** is the residual effect, **id1** is the random additive genetic effect of the FID with a covariance matrix proportional to the numerator relationship matrix, **id2** is the random effect of the FIDs permanent environment (i.e. individual identity, capturing environmental effects that are constant across repeated measures on an FID), **het** is the method-of-moments F coefficient estimates (i.e. observed homozygosity count – expected homozygosity count) / total observations – expected homozygosity count) calculated with the *--het* function in PLINK1.9 (Chang et al., 2015), The narrow-sense heritability (h^2) was defined as the proportion of phenotypic variance explained by the additive genetic effect.

Genome-wide associations with individual recombination rates

The fastGWA module implemented in GCTA (J. Yang et al., 2011) was used to look for potential associations between any of the SNP markers from the 600K datasets and mean individual ACC. This is a mixed models-based tool that uses a sparse genomic relatedness matrix (GRM) to correct for relatedness and principal components to control for population stratification. A sparse genomic relatedness matrix was created with the *--make-bK-sparse* option using a cut-off value of 0.05. i.e., off-diagonal elements below 0.05 were set to 0. The genome-wide significance threshold = 0.05/Number of markers per analysis.

Results

Broad and fine scale recombination rates

The total genetic length of the Norwegian Red autosomes is 2492.9 cM in males and 2308.9 cM in females, equating to an overall rate of autosomal recombination of 1 cM/Mb in males and 0.93 cM/Mb in females (Table 2). The total recombination rate per chromosome in cM per Mb varies from 0.81 to 1.38 in males and from 0.78 to 1.34 in females in the 29 autosomes (Table 2). Sex differences in recombination rate are mainly driven by elevated male recombination in sub-telomeric regions of some chromosomes (i.e. the last 10-30 Mb), with the largest effects seen in chromosomes 13, 17 and 19, which are 14, 14 and 20 % longer in males than in females, respectively (Figures 2 & 3). However, there are generally no sex-differences variation in recombination rate along the remainder of the autosomal chromosomes (Figure 3). Chromosome 15 is the only chromosome where the female rate is elevated in the sub-telomeric region, although the male map length is still longer overall (Figures 2 & 3). The relationship between the physical length (Mb) and genetic length (cM) of the autosomes was close to linear in both males and females (adjusted $R^2 = 0.92$ and 0.97, respectively) (Figure 4).

Table 2. Male and female linkage map lengths for all 29 autosomes. Male and female cM is the estimated total genetic length of the autosomes in centiMorgans. cM/Mb is the recombination rate in centiMorgan per Megabase. The physical map length is relative to the ARS-UCD1.2 genome.

Chr	physical length (Mb)	Male cM	Male cM/Mb	female cM	female cM/Mb	%
1	158.30	128.8	0.81	123.4	0.78	4
2	136.23	114.6	0.84	111.3	0.82	3
3	121.01	110.7	0.91	103.6	0.86	6
4	120.00	107.9	0.90	104.7	0.87	3
5	120.09	110.2	0.92	101.1	0.84	8
6	117.81	104.4	0.89	99.3	0.84	5
7	110.68	106.9	0.97	102.8	0.93	4
8	113.32	101.4	0.89	96.9	0.86	4
9	105.45	95.7	0.91	86.9	0.82	9
10	103.31	102.0	0.99	94.1	0.91	8
11	106.98	99.2	0.93	92.6	0.87	7
12	87.22	87.1	1.00	78.2	0.90	10
13	83.47	93.7	1.12	80.7	0.97	14
14	82.40	84.5	1.03	76.1	0.92	10
15	85.01	86.3	1.02	78.3	0.92	9
16	81.01	86.1	1.06	75.9	0.94	12
17	73.17	78.3	1.07	67.5	0.92	14
18	65.82	79.8	1.21	69.8	1.06	13
19	63.45	89.8	1.42	71.2	1.12	21
20	71.97	67.7	0.94	64.7	0.90	4
21	69.86	74.1	1.06	68.3	0.98	8
22	60.77	73.6	1.21	65.5	1.08	11
23	52.50	63.1	1.20	57.9	1.10	8
24	62.32	62.3	1.00	60.6	0.97	3
25	42.35	58.3	1.38	56.7	1.34	3
26	51.99	57.5	1.11	56.3	1.08	2
27	45.61	55.8	1.22	54.7	1.20	2
28	45.94	55.5	1.21	54.3	1.18	2
29	51.10	57.6	1.13	55.5	1.09	4
TOTAL	2 489.14	2492.9	1.00	2308.9	0.93	7

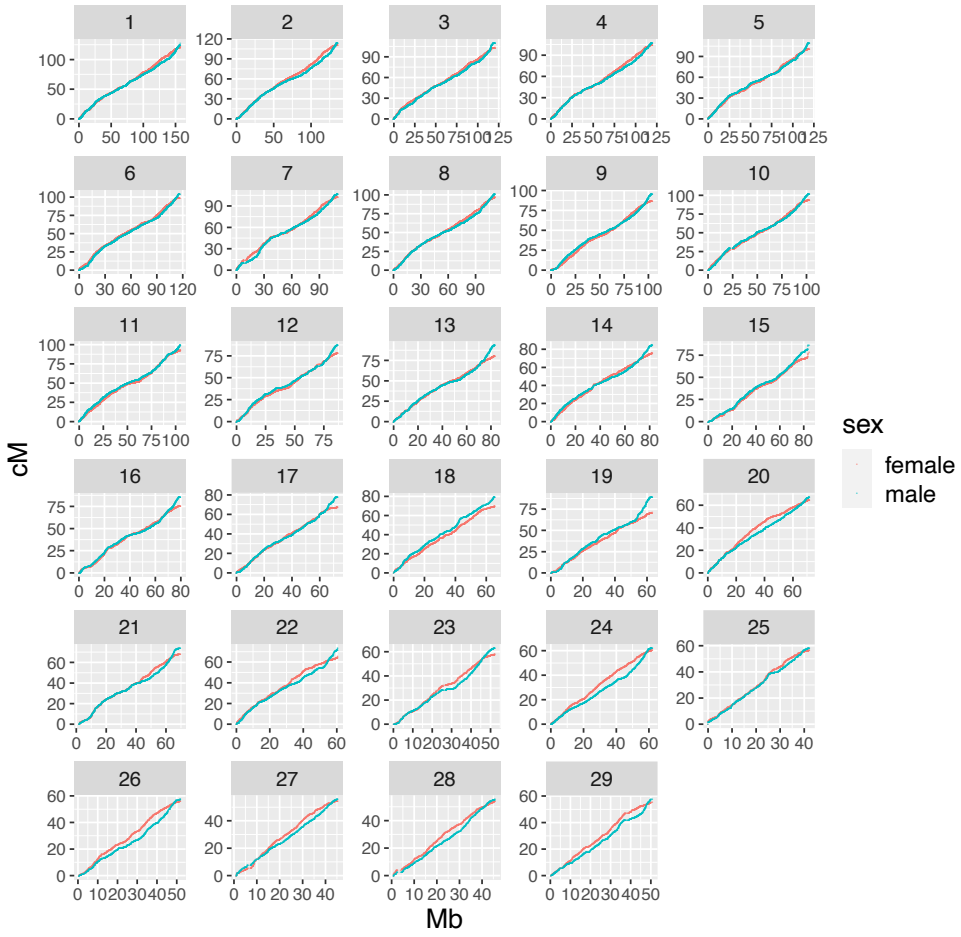


Figure 2. Sex-specific autosomal linkage maps for cattle autosomes. The genetic position of each SNP marker in cM is plotted against its physical position on the 29 autosomes. Male positions in blue and female positions in red. Cattle autosomes are acrocentric, with the left and right hand ends of each map corresponding to the centromere and sub-telomeric regions, respectively.

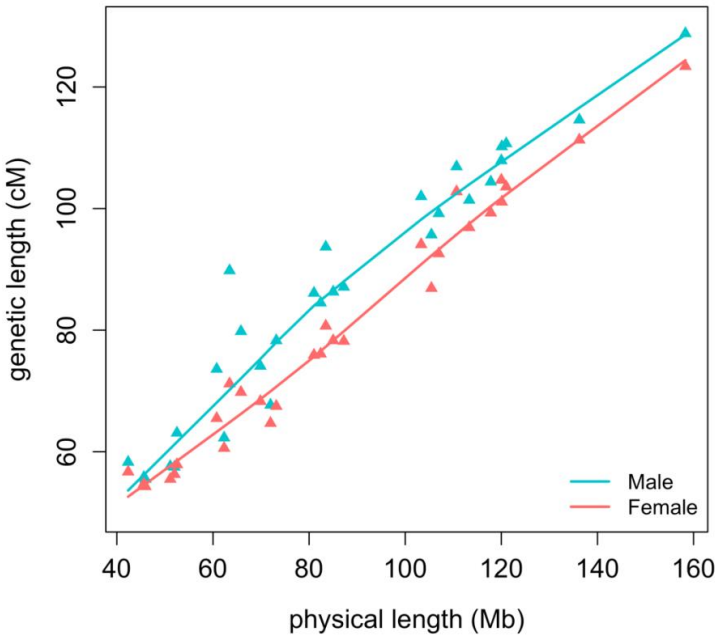


Figure 3. The relationship between the linkage map length (cM) and physical length (Mb) of the 29 autosomes. The relationship is plotted with robust locally weighted regression using the Lowess smoother in R.

Individual recombination rates

The ACC per gamete is close to normally distributed in both sexes, with mean of 24.3 (± 4.3) for males and 22.2 (± 4.9) for females (Figure 5). Gametes with <6 or >50 crossovers were excluded (plot with distribution of ACC before filtering can be found in S1), resulting in a total of 19,861 ACC measures for males and 19,824 for females. The distribution of ACC per gamete before filtering is plotted in Figure S1. The heritability (h^2) for ACC was 9% in males and 4% in females (Table 3). Most of the phenotypic variance is explained by the error term in both sexes (Table 3). Inbreeding significantly decreased ACC in both sexes (-9.8 and -13.7 ACC per unit $F_{hom,}$ in males and females, respectively), equivalent to a difference of 2.35

ACC for males and 6.35 ACC for females between individuals with the minimum and maximum levels of inbreeding in our data. There was no effect of age on ACC in either sex in our analysis. Results from the analysis of genetic variation in individual crossover count can be found in Table 3.

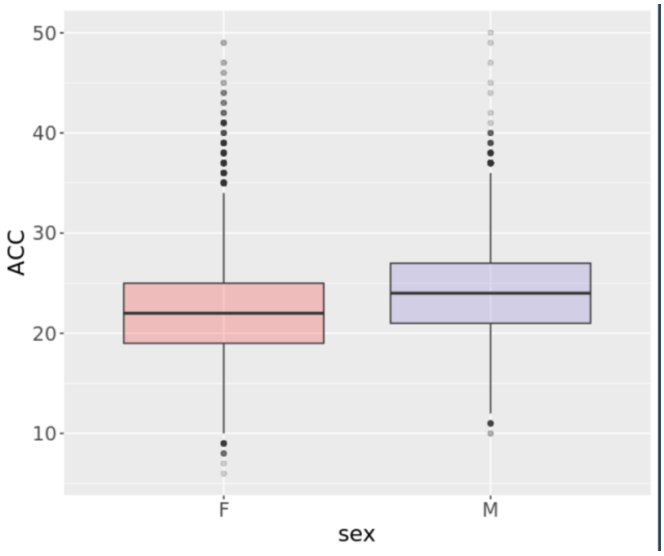


Figure 4. Median and Distribution of autosomal crossover count in females (red) and males (blue). The midline is the median and the box is from the 25th percentile to the 75th percentile.

Table 3. Results from the analysis of genetic variance of individual ACC. N_{FIDs} is the number of focal individuals in the respective model, N_{obs} is the total number of ACC measures, Mean is the mean ACC with standard deviations in parentheses, h^2 is the heritability estimate with standard errors in parentheses. V_P and V_e is the phenotypic and error variances, respectively. F_{hom} is the slope of the inbreeding coefficient when fit as a fixed effect.

Sex	**N_{FIDs}**	**N_{obs}**	**Mean (sd)**	**h^2 (SE)**	**V_P**	**V_e**	**$f_{hom\ effect}$**
Female	14 815	19824	22.2 (4.9)	0.04 (0.01)	24.0	21.2	-13.7 (1.25)
Male	603	19861	24.3 (4.3)	0.09 (0.02)	18.2	16.01	-9.8 (2.8)

Genome wide associations with ACC

We found 3 genomic regions significantly associated with ACC in females, with two on chromosome 10 and one on chromosome 26 (Figure 6, Table 4). The top SNP in the first peak on chromosome 10 is 85.0 kb away from *RNF212B*, a gene that has been previously associated with individual recombination rates in domestic sheep, Soay sheep, cattle, and red deer (S. Johnston et al., 2020; S. E. Johnston et al., 2018; Kadri et al., 2016; Petit et al., 2017). Under the second peak on chromosome 10, there are two candidate genes; *MHL3* and *NEK9*, which are 20.7 kb and 45.0 kb away from the top SNP in that region (Figure 6). *MLH3* and *NEK9* have been reported as being associated with cattle recombination rates in Kadri et al. (2016) and Ma et al.(2015), respectively. The most highly associated SNP on chromosome 26 is 45.6 kb away from the gene *CEP55*, which is associated with individual recombination rates in Holstein (Ma et al. 2015). No markers showed significant associations with ACC in males; however, the most highly associated SNP corresponds to a region on chromosome 6 containing *RNF212* previously associated with cattle recombination rate (Kadri et al., 2016).

Table 4. Top SNPs in genome-wide association with ACC.
SNP positions in base pairs, MAF is the minor allele frequency. Beta is the additive effect size (i.e. the slope) of allele, SE is the standard error of Beta, P is the P value.

Sex	Chr	SNP	bp	MAF	Beta	SE	P	Candidate Gene(s)
Female	10	BovineHD1000006959	21574661	0.12	0.48	0.09	5.91E-08	*RNF212B*
Female	10	BTA-78285-no-rs	86322591	0.27	0.45	0.07	7.93E-12	*NEK9* *MLH3*
Female	26	BovineHD2600003818	14891061	0.18	0.56	0.08	1.34E-13	*CEP55*

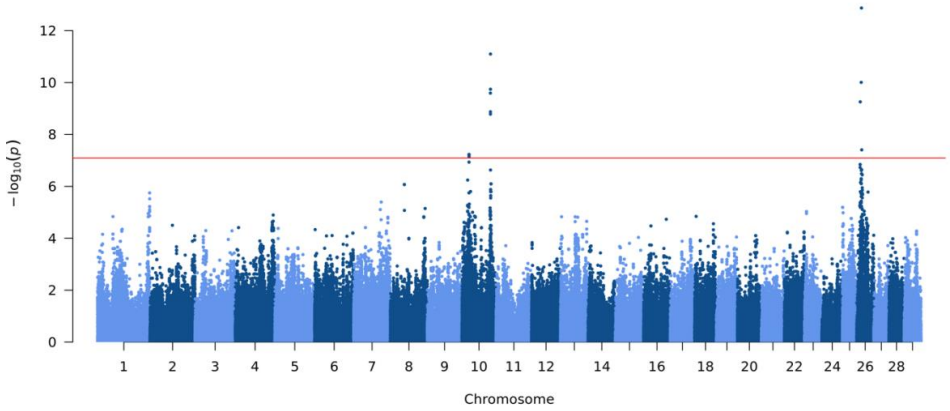


Figure 5. Genome-wide associations between individual ACC in females and the SNP markers from the 600K dataset. The red line is the statistical significance threshold equivalent to $P < 0.05$ after multiple testing.

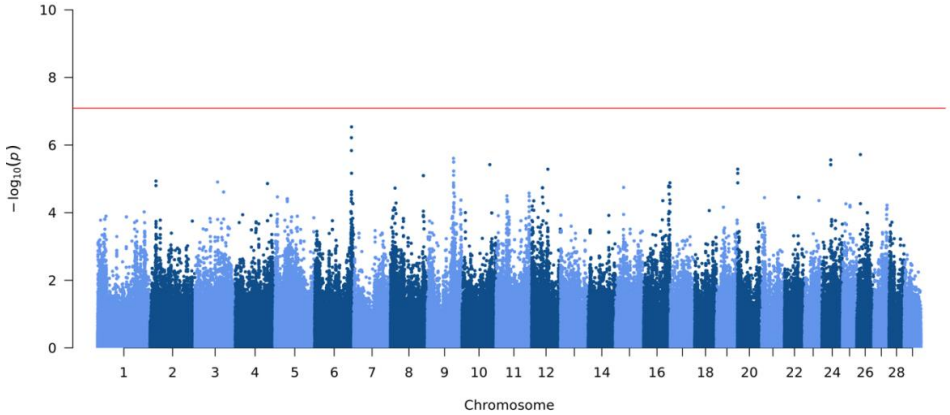


Figure 6. Genome-wide associations between individual ACC in males and the SNP markers on from the 600K dataset. The red line is the statistical significance threshold equivalent to $P < 0.05$ after multiple testing.

Discussion

In this study we took advantage of the extensive genotyping in the breeding programme of Norwegian Red cattle to study individual recombination rates. We found that recombination rates in this breed vary between the sexes and within and between chromosomes. Individual recombination rates are heritable in both sexes, and we found three loci significantly associated with the trait in females, all close to genes that have previously been associated with recombination rate variation in vertebrates. In the following section, we discuss the results in more detail and consider the possible causes and implications of variation in recombination rates for a breeding population.

Cattle have relatively low sexual dimorphism in recombination landscapes

In total, the genetic length of the male autosomes is 8% longer than the female autosomes, which is similar to other cattle breeds studied, such as Holstein, Holstein Frisian and Jersey where male maps were 8-9% longer (Kadri et al., 2016; Ma, O'Connell, et al., 2015). The genetic length of individual autosomes was consistently higher in males than in females and is driven by the sub-telomeric regions. Along the rest of the autosomes, the sex specific patterns of recombination rate are almost overlapping and sometimes slightly higher in females (Figure 3). Overall, these patterns were similar to those found in previous cattle studies, indicating conservation of recombination patterns and their sex differences across breeds and autosomes. The modest sex difference observed in cattle is in contrast to other mammal species, such as in sheep where the male maps are 24% longer (S. E. Johnston et al., 2016), or in pigs (Johnsson et al., 2021), red deer (S. E. Johnston et al., 2017) and humans (Broman et al., 1998; Halldorsson et al., 2019b), where the female maps are 20-27%, 18%, and 38-39% longer, respectively.

Autosomal crossover count is associated with RNF212B, CEP55 and NEK9 or MLH3

We find three loci that exceeds the significance threshold in females, but no significant peaks in males. The sex difference may be due to the difference in the number of FIDs, but comparisons with other studies indicate that there are likely to be actual sex difference in genetic architecture of the trait. The first peak on chromosome 10 only barely exceeds the significance threshold, but this peak is near the gene *RNF212B* that has showed association with individual crossover rates in both male and female cattle in Kadri et al (2016), as well as other ruminants like red deer females (S. Johnston et al., 2020) and both male and female in sheep (S. E. Johnston et al., 2018; Petit et al., 2017). *RNF212B* is a close paralog to *RNF212*, which is a gene that is well known to influence recombination rates in mammals (Reynolds et al., 2013). In the next region on chromosome 10 the top SNP is at position 86.32 Mb, relatively close to both *NEK9* at 86.33 – 86.37 Mb and *MLH3* at 86.28 – 86.30 Mb. *NEK9* and *MLH3* were associated with individual recombination rates in both sexes in the study by Ma et al. (2015) and Kadri et al. (2016). *NEK9* is involved in spindle organization and alignment and segregation of the chromosomes during oocyte meiosis (S. W. Yang et al., 2012). *MLH3* is a MutL homolog involved in post replicative mismatch repair which has been shown to interact with the meiosis specific protein MSH4, that has a well-documented role in recombination (Santucci-Darmanin et al., 2002). The third peak, on chromosome 26, is near the gene *CEP55*. This gene is involved in spindle organization (Xu et al., 2015). This association was also female specific in Ma et al(2015).

Sex differences in recombination in a broader context.

Most species studied to date show a difference in recombination rate between males and females, in the overall rate and/or the pattern of recombination along the genome, but it is more common among mammals to have a female biased heterochiasmy (Halldorsson et al.,

2019a; Johnsson et al., 2021; S. E. Johnston et al., 2017; Wang et al., 2017). A long-standing theory has been that selection and domestication increases recombination rates (Burt & Bell, 1987), and Ma et al (2015) suggests that a higher selection pressure in bulls may explain the higher male recombination rates. However, in domestic pigs, where there is also a higher selection pressure in males, heterochiasmy is female biased (Brekke et al 2022, Johnsson et al., 2021). Furthermore, the theory of higher recombination rates in domesticated species versus their wild counterparts have been challenged in at least three pairs of species (dog vs. wolf, sheep vs. mouflon, and goat vs. ibex) where no difference in recombination rate was observed (Munoz-Fuentes et al., 2015). There may also be sexual dimorphism in the selection at the gamete level that differs between species, in a study on Soay sheep, they hypothesised that the higher male recombination rates may be due to high levels of sperm competition in Soay sheep males as they have a highly promiscuous mating system, although this could not be formally tested (S. E. Johnston et al., 2016).

In Ma et al. (2015), they also looked at development of recombination rates over time and found a steady decrease in recombination rates in males the last 40 years, which coincided with a decrease in fertility. A recent study of fertility in Norwegian Red bulls also show a slight, but significant unfavourable genetic trend from 1994 – 2016 (Olsen et al., 2020). One hypothesis could be that historically, selection has led to higher recombination rates because individuals with unique combinations of alleles are selected (Charlesworth & Barton, 1996), which may indirectly select for alleles associated with higher recombination. However, once bulls have favourable haplotypes, the gametes from low recombination individuals are the ones leading to favourable offspring – leading to selection for lower recombination rates, at least on some chromosomes. A study that compared recent and historical recombination in Norwegian Red bulls found lower recent recombination rates compared to historical rates on

chromosomes with important quantitative trait loci for milk production (Sodeland et al., 2011). This could also mean that the sex difference in recombination rate may have been larger. Overall, biological explanations for difference in recombination rates between sexes remains to be understood, as most hypotheses have species that contradict the theory, and there is a lack of empirical studies on heterochiasmy.

There is genetic variation in individual rates in both sexes.

The heritability is higher in males than in females, which is consistent with what has been found in other studies in cattle, but the heritability estimates in our study is lower than what other studies have found (Kadri et al., 2016; Sandor et al., 2012). Sandor et al. (2012) only have bulls in their study and estimated $h^2 = 0.22$, whereas Kadri et al. (2016) finds heritability 0.13 in males and 0.08 in females which is closer to what we find, but still substantially higher. This could be due to breed differences in allele frequencies for loci affecting recombination rates. The standard error is slightly higher in males, probably due to the number of FIDs being much lower (603) than in females (14,815), but the standard errors are not very high in either sex (0.01 and 0.02 in females and males respectively).

The observed effect of inbreeding on recombination rates may be due to long runs of homozygosity affecting the ability to detect crossovers, rather than a true effect of inbreeding itself on reducing recombination rates. Most of the phenotypic variance is explained by the error terms (Table 3). An explanation to this may be that because we are studying recombination in gametes in live offspring, i.e. successful gametes, it may not be a random sample from the meiosis, but there may be selection at the haploid level in both males and females. Studies in human show signs of selection against non-recombinant chromatids in meiosis II (Ottolini et al., 2015).

Reflections / implications for selective breeding.

Recombination is one of the main contributors to within-family genetic variation and therefore important for breeding. With a greater use of genomic information in the breeding work and selection based on genomic evaluations, insights into the breed specific patterns of variation in recombination rates may be of great importance. Recombination rate is relatively evenly distributed along the autosome compared to many other species and more closely resembles how recombination rates are typically modelled in phasing and imputation software like for example Beagle (Browning & Browning, 2007) and SHAPEIT (O'Connell et al., 2014), that defaults to a constant recombination rate of 1 cM per Mb. This might suggest that there is not much to gain by using breed and sex specific linkage maps in phasing and imputation in cattle. However, there are clear sex differences in the subtelomeric regions, but to our knowledge there are no phasing or imputation software available to date that can take sex-specific linkage maps. Indeed, the sex averaged map would not show this pattern at the ends as most of the autosomes have opposite pattern in males and females.

Individual recombination rate is a heritable trait in cattle and could therefore be selected on to increase genetic variance. However, studies show that to really have an effect on the selection response, the emphasis on this trait in the breeding work would have to be unrealistically high (Battagin et al., 2016). Also, there may be disadvantages to having high recombination rates like increased mutation rates (Arbeithuber et al., 2015) and break up of favourable linkage (Charlesworth & Barton, 1996). Knowing that at least one recombination is needed for proper segregation of chromosomes in meiosis and that aneuploidy is one of the main reasons for pregnancy loss in human (Fledel-Alon et al., 2009; Hassold et al., 1995; Koehler et al., 1996; Sherman et al., 1991) it may be important for fertility to make sure that the rates are not too low. It would be of great interest to specifically look at effects of low recombination rates on

fertility as well as potential selection between gametes maybe by studying recombination in sperm and egg cells instead of in live offspring.

Conclusion

In this study, we find that both patterns and rates of recombination differ between the sexes in the subtelomeric regions but are relatively evenly distributed and overlapping between the sex in the rest of the autosomes. The genome wide recombination rates in Norwegian Red cattle are comparable to other cattle breeds studied. In agreement with other studies, we find that there is a low but significant heritable component to the genetic variation, but that the trait has a large error variance. It would be of interest to study sperm and egg cells to possibly explain more of the phenotypic variance in recombination rates in cattle. We find genes associated with individual recombination rates also found in other studies suggesting that some of the genetic architecture of the trait is well conserved across cattle breeding populations and other ruminants. Further studies should also look at the mechanisms leading to the difference in recombination rates between sex as well as how individual rates may affect fertility. We here provide an example of the genetic basis of recombination rates in a domesticated breed under strong selection, contributing to the understanding of the underlying mechanisms of individual recombination in cattle as well as in mammals in general.

References

Arbeithuber, B., Betancourt, A. J., Ebner, T., & Tiemann-Boege, I. (2015). Crossovers are associated with mutation and biased gene conversion at recombination hotspots. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(7), 2109–2114. https://doi.org/10.1073/pnas.1416622112

Barton, N. H., & Charlesworth, B. (1998). *Why Sex and Recombination?* https://doi.org/10.1126/science.281.5385.1986

Battagin, M., Gorjanc, G., Faux, A. M., Johnston, S. E., & Hickey, J. M. (2016). Effect of manipulating recombination rates on response to selection in livestock breeding programs. *Genetics Selection Evolution*, *48*(1). https://doi.org/10.1186/s12711-016-0221-1

Booker, T. R., Ness, R. W., & Keightley, P. D. (2017). The recombination landscape in wild house mice inferred using population genomic data. *Genetics*, *207*(1), 297–309. https://doi.org/10.1534/genetics.117.300063

Broman, K. W., Murray, J. C., Sheffield, V. C., White, R. L., & Weber, J. L. (1998). Comprehensive Human Genetic Maps: Individual and Sex-Specific Variation in Recombination. In *Am. J. Hum. Genet* (Vol. 63).

Browning, S. R., & Browning, B. L. (2007). Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *American Journal of Human Genetics*, *81*(5), 1084–1097. https://doi.org/10.1086/521987

Burt, A., & Bell, G. (1987). Mammalian chiasma frequencies as a test of two theories of recombination. *Nature*, *326*, 803–805.

Chang, C. C., Chow, C. C., Tellier, L. C. A. M., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience*, *4*(1). https://doi.org/10.1186/s13742-015-0047-8

Charlesworth, B., & Barton, N. H. (1996). Recombination load associated with selection for increased recombination. *Genetical Research*, *67*(1), 27–41. https://doi.org/10.1017/s0016672300033450

Dumont, B. L., Broman, K. W., & Payseur, B. A. (2009). Variation in genomic recombination rates among heterogeneous stock mice. *Genetics*, *182*(4), 1345–1349. https://doi.org/10.1534/genetics.109.105114

Fledel-Alon, A., Leffler, E. M., Guan, Y., Stephens, M., Coop, G., & Przeworski, M. (2011). Variation in human recombination rates and its genetic determinants. *PLoS ONE*, *6*(6). https://doi.org/10.1371/journal.pone.0020321

Fledel-Alon, A., Wilson, D. J., Broman, K., Wen, X., Ober, C., Coop, G., & Przeworski, M. (2009). Broad-scale recombination patterns underlying proper disjunction in humans. *PLoS Genetics*, *5*(9). https://doi.org/10.1371/journal.pgen.1000658

Gao, N., Teng, J., Ye, S., Yuan, X., Huang, S., Zhang, H., Zhang, X., Li, J., & Zhang, Z. (2018). Genomic prediction of complex phenotypes using genic similarity based relatedness matrix. *Frontiers in Genetics*, *9*(AUG). https://doi.org/10.3389/fgene.2018.00364

Götherström, A., Anderung, C., Hellborg, L., Elburg, R., Smith, C., Bradley, D. G., & Ellegren, H. (2005). Cattle domestication in the Near East was followed by hybridization with aurochs bulls in Europe. *Proceedings of the Royal Society B: Biological Sciences*, *272*(1579), 2345–2350. https://doi.org/10.1098/rspb.2005.3243

Halldorsson, B. v., Hardarson, M. T., Kehr, B., Styrkarsdottir, U., Gylfason, A., Thorleifsson, G., Zink, F., Jonasdottir, A., Jonasdottir, A., Sulem, P., Masson, G., Thorsteinsdottir, U.,

Helgason, A., Kong, A., Gudbjartsson, D. F., & Stefansson, K. (2016). The rate of meiotic gene conversion varies by sex and age. *Nature Genetics*, *48*(11), 1377–1384. https://doi.org/10.1038/ng.3669

Halldorsson, B. v., Palsson, G., Stefansson, O. A., Jonsson, H., Hardarson, M. T., Eggertsson, H. P., Gunnarsson, B., Oddsson, A., Halldorsson, G. H., Zink, F., Gudjonsson, S. A., Frigge, M. L., Thorleifsson, G., Sigurdsson, A., Stacey, S. N., Sulem, P., Masson, G., Helgason, A., Gudbjartsson, D. F., … Stefansson, K. (2019a). Human genetics: Characterizing mutagenic effects of recombination through a sequence-level genetic map. *Science*, *363*(6425). https://doi.org/10.1126/science.aau1043

Halldorsson, B. v., Palsson, G., Stefansson, O. A., Jonsson, H., Hardarson, M. T., Eggertsson, H. P., Gunnarsson, B., Oddsson, A., Halldorsson, G. H., Zink, F., Gudjonsson, S. A., Frigge, M. L., Thorleifsson, G., Sigurdsson, A., Stacey, S. N., Sulem, P., Masson, G., Helgason, A., Gudbjartsson, D. F., … Stefansson, K. (2019b). Human genetics: Characterizing mutagenic effects of recombination through a sequence-level genetic map. *Science*, *363*(6425). https://doi.org/10.1126/science.aau1043

Hassold, T., Merrill, M., Adkins, K., Freeman, S., & Sherman2, S. (1995). Recombination and Maternal Age-Dependent Nondisjunction: Molecular Studies of Trisomy 16. In *Am. J. Hum. Genet* (Vol. 57).

Hunter, C. M., Huang, W., Mackay, T. F. C., & Singh, N. D. (2016). The Genetic Architecture of Natural Variation in Recombination Rate in Drosophila melanogaster. *PLoS Genetics*, *12*(4). https://doi.org/10.1371/journal.pgen.1005951

Johnsson, M., Whalen, A., Ros-Freixedes, R., Gorjanc, G., Chen, C. Y., Herring, W. O., de Koning, D. J., & Hickey, J. M. (2021). Genetic variation in recombination rate in the pig. *Genetics Selection Evolution*, *53*(1). https://doi.org/10.1186/s12711-021-00643-0

Johnston, S. E., Bérénos, C., Slate, J., & Pemberton, J. M. (2016). Conserved genetic architecture underlying individual recombination rate variation in a wild population of soay sheep (Ovis aries). *Genetics*, *203*(1), 583–598. https://doi.org/10.1534/genetics.115.185553

Johnston, S. E., Huisman, J., Ellis, P. A., & Pemberton, J. M. (2017). A high-density linkage map reveals sexual dimorphism in recombination landscapes in red deer (Cervus elaphus). *G3: Genes, Genomes, Genetics*, *7*(8), 2859–2870. https://doi.org/10.1534/g3.117.044198

Johnston, S. E., Huisman, J., & Pemberton, J. M. (2018). A genomic region containing REC8 and RNF212B is associated with individual recombination rate variation in a wild population of red deer (Cervus elaphus). *G3: Genes, Genomes, Genetics*, *8*(7), 2265–2276. https://doi.org/10.1534/g3.118.200063

Johnston, S., Stoffel, M., & Pemberton, J. (2020). *Variants at RNF212 and RNF212B are associated with recombination rate variation in Soay sheep (Ovis aries)*. https://doi.org/10.1101/2020.07.26.217802

Kadri, N. K., Harland, C., Faux, P., Cambisano, N., Karim, L., Coppieters, W., Fritz, S., Mullaart, E., Baurain, D., Boichard, D., Spelman, R., Charlier, C., Georges, M., & Druet, T. (2016). Coding and noncoding variants in HFM1, MLH3, MSH4, MSH5, RNF212, and RNF212B affect recombination rate in cattle. *Genome Research*, *26*(10), 1323–1332. https://doi.org/10.1101/gr.204214.116

Koehler, K. E., Hawley, R. S., Sherman, S., & Hassold, T. (1996). Recombination and nondisjunction in humans and flies. In *Human Molecular Genetics* (Vol. 5). Oxford University Press. https://academic.oup.com/hmg/article/5/Supplement_1/1495/663563

Kong, A., Thorleifsson, G., Stefansson, H., Masson, G., Helgason, A., Gudbjartsson, D. F., Jonsdottir, G. M., Gudjonsson, S. A., Sverrisson, S., Thorlacius, T., Jonasdottir, A., Hardarson, G. A., Palsson, S. T., Frigge, M. L., Gulcher, J. R., Thorsteinsdottir, U., &

Stefansson, K. (2008). Sequence Variants in the RNF212 Gene Associate with Genome-Wide Recombination Rate. *Science*, *319*(5868), 1398–1401. https://doi.org/10.1126/science.1152422

Lenormand, T., & Dutheil, J. (2005). Recombination difference between sexes: A role for haploid selection. *PLoS Biology*, *3*(3), 0396–0403. https://doi.org/10.1371/journal.pbio.0030063

Lenormand, T., Engelstädter, J., Johnston, S. E., Wijnker, E., & Haag, C. R. (2016). Evolutionary mysteries in meiosis. In *Philosophical Transactions of the Royal Society B: Biological Sciences* (Vol. 371, Issue 1706). Royal Society of London. https://doi.org/10.1098/rstb.2016.0001

Lotterhos, K. E. (2019). The effect of neutral recombination variation on genome scans for selection. *G3: Genes, Genomes, Genetics*, *9*(6), 1851–1867. https://doi.org/10.1534/g3.119.400088

Ma, L., O'connell, J. R., Vanraden, P. M., Shen, B., Padhi, A., Sun, C., Bickhart, D. M., Cole, J. B., Null, D. J., Liu, G. E., Da, Y., & Wiggans, G. R. (2015). Recombination and Genetic Control from a Large Pedigree Analysis. *PLoS Genet*, *11*(11), 1005387. https://doi.org/10.1371/journal.pgen.1005387

Ma, L., O'Connell, J. R., VanRaden, P. M., Shen, B., Padhi, A., Sun, C., Bickhart, D. M., Cole, J. B., Null, D. J., Liu, G. E., Da, Y., & Wiggans, G. R. (2015). Cattle Sex-Specific Recombination and Genetic Control from a Large Pedigree Analysis. *PLoS Genetics*, *11*(11). https://doi.org/10.1371/journal.pgen.1005387

Madsen, P., Jensen, J., Labouriau, R., Christensen, O. F., & Sahana, G. (2014). DMU - A Package for Analyzing Multivariate Mixed Models in quantitative Genetics and Genomics. *Journal of Applied Statistics*, *41*(6), 1286–1306.

Mank, J. E. (2009). The evolution of heterochiasmy: The role of sexual selection and sperm competition in determining sex-specific recombination rates in eutherian mammals. *Genetics Research*, *91*(5), 355–363. https://doi.org/10.1017/S0016672309990255

Munoz-Fuentes, V., Marcet-Ortega, M., Alkorta-Aranburu, G., Forsberg, C. L., Morrell, J. M., Manzano-Piedras, E., Soderberg, A., Daniel, K., Villalba, A., Toth, A., di Rienzo, A., Roig, I., & Vila, C. (2015). Strong artificial selection in domestic mammals did not result in an increased recombination rate. *Molecular Biology and Evolution*, *32*(2), 510–523. https://doi.org/10.1093/molbev/msu322

O'Connell, J., Gurdasani, D., Delaneau, O., Pirastu, N., Ulivi, S., Cocca, M., Traglia, M., Huang, J., Huffman, J. E., Rudan, I., McQuillan, R., Fraser, R. M., Campbell, H., Polasek, O., Asiki, G., Ekoru, K., Hayward, C., Wright, A. F., Vitart, V., … Marchini, J. (2014). A General Approach for Haplotype Phasing across the Full Spectrum of Relatedness. *PLoS Genetics*, *10*(4). https://doi.org/10.1371/journal.pgen.1004234

Olsen, H. B., Heringstad, B., & Klemetsdal, G. (2020). Genetic analysis of semen characteristic traits in young Norwegian Red bulls. *Journal of Dairy Science*, *103*(1), 545–555. https://doi.org/10.3168/jds.2019-17291

Ottolini, C. S., Newnham, L. J., Capalbo, A., Natesan, S. A., Joshi, H. A., Cimadomo, D., Griffin, D. K., Sage, K., Summers, M. C., Thornhill, A. R., Housworth, E., Herbert, A. D., Rienzi, L., Ubaldi, F. M., Handyside, A. H., & Hoffmann, E. R. (2015). Genome-wide maps of recombination and chromosome segregation in human oocytes and embryos show selection for maternal recombination rates. *Nature Genetics*, *47*(7), 727–735. https://doi.org/10.1038/ng.3306

Paigen, K., & Petkov, P. M. (2018). PRDM9 and Its Role in Genetic Recombination. In *Trends in Genetics* (Vol. 34, Issue 4, pp. 291–300). Elsevier Ltd. https://doi.org/10.1016/j.tig.2017.12.017

Petit, M., Astruc, J. M., Sarry, J., Drouilhet, L., Fabre, S., Moreno, C. R., & Servin, B. (2017). Variation in recombination rate and its genetic determinism in sheep populations. *Genetics*, *207*(2), 767–784. https://doi.org/10.1534/genetics.117.300123

Rastas, P. (2017). Lep-MAP3: Robust linkage mapping even for low-coverage whole genome sequencing data. *Bioinformatics*, *33*(23), 3726–3732. https://doi.org/10.1093/bioinformatics/btx494

Reynolds, A., Qiao, H., Yang, Y., Chen, J. K., Jackson, N., Biswas, K., Holloway, J. K., Baudat, F., de Massy, B., Wang, J., Höög, C., Cohen, P. E., & Hunter, N. (2013). RNF212 is a dosage-sensitive regulator of crossing-over during mammalian meiosis. *Nature Genetics*, *45*(3), 269–278. https://doi.org/10.1038/ng.2541

Ritz, K. R., Noor, M. A. F., & Singh, N. D. (2017). Variation in Recombination Rate: Adaptive or Not? In *Trends in Genetics* (Vol. 33, Issue 5, pp. 364–374). Elsevier Ltd. https://doi.org/10.1016/j.tig.2017.03.003

Samuk, K., Manzano-Winkler, B., Ritz, K. R., & Noor, M. A. F. (2020). Natural Selection Shapes Variation in Genome-wide Recombination Rate in Drosophila pseudoobscura. *Current Biology*, *30*(8), 1517-1528.e6. https://doi.org/10.1016/j.cub.2020.03.053

Sandor, C., Li, W., Coppieters, W., Druet, T., Charlier, C., & Georges, M. (2012). Genetic variants in REC8, RNF212, and PRDM9 influence male recombination in cattle. *PLoS Genetics*, *8*(7). https://doi.org/10.1371/journal.pgen.1002854

Santucci-Darmanin, S., Neyton, S., Oise Lespinasse, F., Sauniè, A., Gaudray, P., & Ronique Paquis-Flucklinger, V. (2002). *The DNA mismatch-repair MLH3 protein interacts with MSH4 in meiotic cells, supporting a role for this MutL homolog in mammalian meiotic recombination*. https://academic.oup.com/hmg/article/11/15/1697/636769

Shen, B., Jiang, J., Seroussi, E., Liu, G. E., & Ma, L. (2018). Characterization of recombination features and the genetic basis in multiple cattle breeds. *BMC Genomics*, *19*(1). https://doi.org/10.1186/s12864-018-4705-y

Sherman, S. L., Takaesu, N., Freeman, S. B., Grantham, M., Phillips, C., Blackston, R. D., Jacobs, P. A., Cockwell, A. E., Freeman, V., Uchida4, $ 1, Mikkelsen, M., Kurnit, D. M., Buraczynska, I. M., Keats, I. B. J. B., & Hassold, T. J. (1991). Trisomy 21: Association between Reduced Recombination and Nondisjunction. In *Am. J. Hum. Genet* (Vol. 49).

Sodeland, M., Kent, M., Hayes, B. J., Grove, H., & Lien, S. (2011). Recent and historical recombination in the admixed Norwegian Red cattle breed. *BMC Genomics*, *12*. https://doi.org/10.1186/1471-2164-12-33

Sved, J. A., & Hill, W. G. (2018). One hundred years of linkage disequilibrium. *Genetics*, *209*(3), 629–636. https://doi.org/10.1534/genetics.118.300642

Wang, R. J., Gray, M. M., Parmenter, M. D., Broman, K. W., & Payseur, B. A. (2017). Recombination rate variation in mice from an isolated island. *Molecular Ecology*, *26*(2), 457–470. https://doi.org/10.1111/mec.13932

Weng, Z. Q., Saatchi, M., Schnabel, R. D., Taylor, J. F., & Garrick, D. J. (2014). Recombination locations and rates in beef cattle assessed from parent-offspring pairs. *Genetics Selection Evolution*, *46*(1). https://doi.org/10.1186/1297-9686-46-34

Weng, Z., Wolc, A., Su, H., Fernando, R. L., Dekkers, J. C. M., Arango, J., Settar, P., Fulton, J. E., O'Sullivan, N. P., & Garrick, D. J. (2019). Identification of recombination hotspots and quantitative trait loci for recombination rate in layer chickens. *Journal of Animal Science and Biotechnology*, *10*(1). https://doi.org/10.1186/s40104-019-0332-y

Winbush, A., & Singh, N. D. (2021). Genomics of Recombination Rate Variation in Temperature-Evolved Drosophila melanogaster Populations. *Genome Biology and Evolution*, *13*(1). https://doi.org/10.1093/gbe/evaa252

Xu, Z. Y., Ma, X. S., Qi, S. T., Wang, Z. B., Guo, L., Schatten, H., Sun, Q. Y., & Sun, Y. P. (2015). Cep55 regulates spindle organization and cell cycle progression in meiotic oocyte. *Scientific Reports*, *5*. https://doi.org/10.1038/srep16978

Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2011). GCTA: A tool for genome-wide complex trait analysis. *American Journal of Human Genetics*, *88*(1), 76–82. https://doi.org/10.1016/j.ajhg.2010.11.011

Yang, S. W., Gao, C., Chen, L., Song, Y. L., Zhu, J. L., Qi, S. T., Jiang, Z. Z., Wang, Z. W., Lin, F., Huang, H., Xing, F. Q., & Sun, Q. Y. (2012). Nek9 regulates spindle organization and cell cycle progression during mouse oocyte meiosis and its location in early embryo mitosis. *Cell Cycle*, *11*(23), 4366–4377. https://doi.org/10.4161/cc.22690

Supplementary

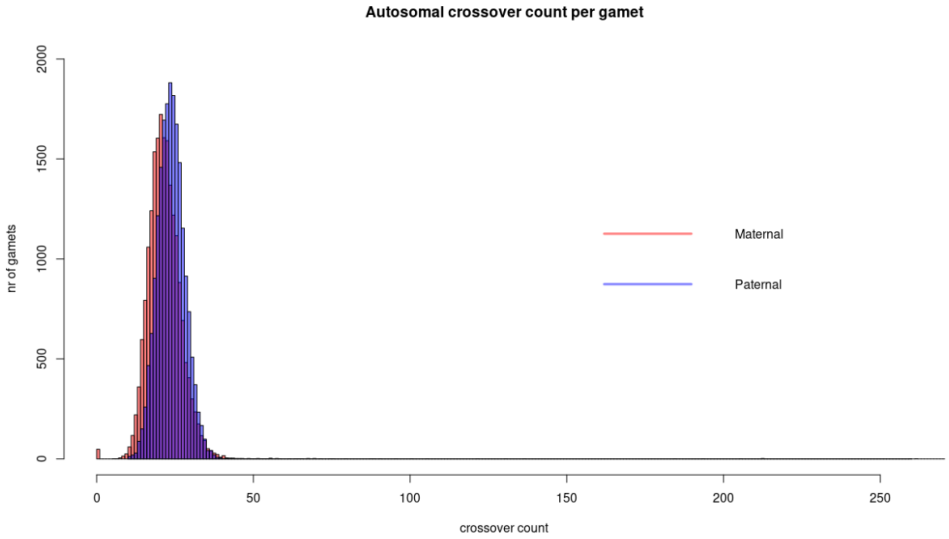


Figure S1. Distribution of crossover count in maternal (red) and paternal (blue) gamets before filtering.

Paper III.

Genetic variation in genome wide recombination rates and genetic shuffling in an Atlantic salmon breeding population.

Brekke, C., Johnston, S.E., Knutsen, T., Berg, P. (2022).
Manuscript.

Genetic variation in genome wide recombination rates and genetic shuffling in an Atlantic salmon breeding population

Cathrine Brekke[1*], Susan E. Johnston[2], Tim Martin Knutsen[3], Peer Berg[1]

Abstract

Meiotic recombination ensures proper segregation of homologous chromosomes in meiosis while also breaking down linkage disequilibrium and shuffling alleles of genes located on the same chromosome. The Atlantic salmon genome shows patterns of delayed diploidization following a whole genome duplication event ~80 mya in the common ancestor of the salmonid lineage, which is likely the cause of the marked differences in distribution of recombination rate between males and females. In this study we use more than 1900 full sib families with large offspring groups to investigate the genetic variation in individual measures of recombination in an Atlantic salmon breeding population with high-density SNP genotypes. The heritability of genome-wide rates of recombination was low but significant in both sexes (h^2 = 0.11 and 0.12 in females and males, respectively) and is similar to estimates in mammalian species. The extreme patterns in males where recombination is restricted to telomeric regions leads to substantially lower levels of genetic shuffling in paternal gametes compared to maternal gametes. In females, the variation in shuffling between individuals is larger than in males and 16 % of the phenotypic variation is explained by genetic variation, suggesting a potential for genetic change in overall rates and distribution of recombination in the Atlantic salmon.

Background

Meiotic recombination is the event where crossing over occurs between the maternal and paternal chromosomes during synapsis in the early prophase I of meiosis, leading to novel combinations of alleles in the gametes transmitted to the next generation. Recombination is of large interest in studies of both wild and domestic species as it breaks up linkage between loci located on the same chromosome and is therefore an important force in creating novel genetic variation for selection to act upon. Several studies show that recombination also has a mechanistic role in the proper alignment of chromosomes during meiosis and that a lack of recombination usually leads to aneuploidy (Sherman *et al.*, 1991; Hassold *et al.*, 1995; Koehler *et al.*, 1996; Fledel-Alon *et al.*, 2009). However, formation of double strand breaks to form crossovers is mutagenic (Halldorsson *et al.*, 2019) and can also break apart beneficial combinations of alleles (Charlesworth and Barton, 1996). Yet, the number and position of crossovers can vary within and between species, sexes, individuals and chromosomes (Ritz, Noor and Singh, 2017; Stapley *et al.*, 2017).

Various studies in eukaryotes find a genetic component to individual variation in recombination rate. Several genes have been associated with crossover count in a number of species, suggesting that it might be under oligogenic control and have a conserved genetic architecture (Dumont, Broman and Payseur, 2009; Ma *et al.*, 2015; Hunter *et al.*, 2016; Johnston *et al.*, 2016; Petit *et al.*, 2017; Wang *et al.*, 2017; Halldorsson *et al.*, 2019; Johnsson *et al.*, 2021). Genes like *RNF212*, *REC8* and *RNF212B* are often associated with recombination rate in mammals (Kong *et al.*, 2008; Sandor *et al.*, 2012; Kadri *et al.*, 2016; Johnston, Huisman and Pemberton, 2018; Johnston, Stoffel and Pemberton, 2020) and *PRDM9* has been identified as a gene that determines recombination hotspot location in mammals (Baudat *et al.*, 2010). A common feature of recombination is that most species

have differences in recombination rates between sexes (known as heterochiasmy) which can be male or female biased (Lenormand and Dutheil, 2005; Brandvain and Coop, 2012; Stapley *et al.*, 2017; Sardell *et al.*, 2018). An although the causes and consequences of this sexual dimorphism has been of interest for decades (Dunn and Bennett, 1967; Nei, 1969; Burt and Bell, 1987; Burt, Bell and Harvey, 1991; Barton and Charlesworth, 1998; Mank, 2009), the mechanistic and evolutionary drivers that lead to heterochiasmy are not yet well understood.

Because of their importance in aquaculture and evolutionary biology, there are many linkage maps in teleost fish, but little is known about how recombination differs at the individual level, and if it is heritable. One striking feature is the marked sex differences and that the direction of heterochiasmy can vary even between closely related species, but the evolutionary mechanisms remain unclear (Cooney, Mank and Wright, 2021) A limitation of previous studies using population-level map lengths alone is that whilst they may show differences in overall rate, they do not quantify the landscape of heterochiasmy, which can vary to a large degree in teleosts (e.g. salmonids where males recombine almost exclusively in telomeric regions, versus sticklebacks where recombination is more uniform (Rastas *et al.*, 2015). These sex differences in landscape may have adaptive importance in terms of differences in genetic shuffling between the sexes i.e. how often pairs of loci are broken apart due to the positioning of crossovers on the chromosomes (Veller, Kleckner and Nowak, 2019). Overall, extending the exploration to look at recombination within gametes can allow us to investigate the extent to which crossover count and genetic shuffling varies at the individual level and what the genetic architecture is, but also the extent to which the genetic architecture of recombination differs between the sexes and its potential role in driving heterochiasmy.

Atlantic salmon (*Salmo salar*) are known to have distinct differences in recombination landscape between males and females (Lien *et al.*, 2011; Gonen *et al.*, 2014). Early linkage maps in Atlantic salmon showed little to no recombination in males because the analyses were based on few markers and information in the telomeric regions were lacking (Moen *et al.*, 2008). As marker density increased, it was shown that males recombine with an overall rate relatively close to that of females in most chromosomes, but that male recombination occurs almost exclusively in the telomeric regions (Lien *et al.*, 2011). Similar sex-specific recombination patterns are also found in rainbow trout *Oncorhynchus mykiss* (Sakamoto *et al.*, 2000). The Salmonidae family (salmon, char, trout, whitefish and graylings) share an ancestor that underwent a whole genome duplication (WGD) event some 50-100 million years ago (Ohno, Wolf and Atkin, 1968). Studies in salmonids find that homeologous chromosome arms (i.e. those that were originally derived from the same chromosome during WGD events) that still show high sequence similarity can exchange genetic material during meiosis in a quadrivalent formation (Allendorf *et al.*, 2015), which appears to be almost exclusive to male meiosis (Timusk *et al.*, 2011; Waples, Seeb and Seeb, 2016) and is therefore likely to result in different recombination patterns between the sex.

In this study, we investigate individual recombination rate variation and genetic shuffling in a large pedigree of Atlantic salmon. We construct sex specific linkage maps from more than 1,900 full sib families with genotypes on ~36,000 SNP markers to quantify individual rates of recombination and investigate the genetic architecture of the trait. We then investigate how differences in male and female crossover positioning affect the frequency of genetic shuffling in maternal and paternal gametes.

Methods

Genotype data

A total of 128 363 individuals from the Norwegian AquaGen Atlantic Salmon (*Salmo Salar*) breeding line were studied. This breeding program started in 1970 and the founders of the population stem from 41 Norwegian rivers(Gjedrem, Gjøen and Gjerde, 1991). Fish from the 2018- and 2020-year classes were included. Genotypes for the 2018 individuals in the study was obtained from a database of genotypes from routine fish genotyping performed by AquaGen as part of their breeding program. The individuals were genotyped on a mix of three custom Thermo Fisher genotyping arrays developed for salmon, with marker densities of approx. 50,000 for one array and two arrays with 70,000 markers. All genotype calls were made using a pipeline developed according the Thermo Fisher Best practices Genotyping Analysis workflow. Only markers common to all three arrays were included, and markers in the good categories PolyHighRes and NoMinorHom were kept for further analysis. Filtering based on these quality parameters, gave a total genotyping call rate for the data set of 0.997. The genotypes for the 2020 year class, was obtained in the same manner as described above, but in this case using only the two genotyping arrays with 70,000 markers density, and therefore ~57,000 markers instead of ~36,000. The datasets from the two year classes were merged, and only markers common to both datasets were used in further analysis, resulting a datset with 36 394 markers, referred to as the 36K dataset (Table 1).

Table 1. Genotype summary for each data-set.

Data-set	Number of individuals	Number of markers
AS18	83 132	36 394
AS20	45 231	57 648
36K (AS18 and AS20)	128 363	35 715

Three generation full-sib families

In order to carry out linkage mapping and crossover estimation, the pedigree was ordered into three generation full-sib families in the following manner: for each unique sire – dam mating pair (hereafter referred to as focal individuals or FIDs), a family was constructed including their offspring and potentially genotyped parents (see Figure 1 for illustration of the family structure). This enables phasing of the gametes transmitted from the FIDs to their offspring, and in turn allows the counting of crossovers inherited within that gamete from meiosis occurring within the FIDs. An FID can be in several families if the individual is mated with several individuals in the pedigree, or as an offspring or grandparent, but each gamete transmitted to an offspring is only counted once using this approach. The crossover count is assigned as a phenotype to the FID. The 36K set had 1,952 full-sib families with number of offspring in a family ranging from 1 to 543. Because the number of full-sibs were high in most of the families, genotyped grandparents were not vital for proper phasing of the offspring gametes, and therefore having grandparents within the family was not set as a fixed criteria for this analysis.

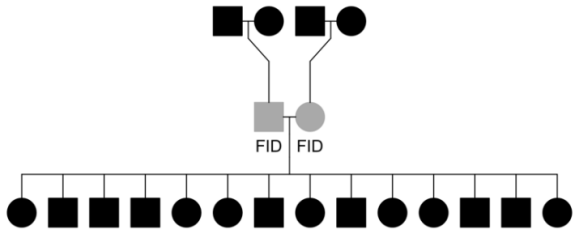


Figure 1 Illustration of the full sib family structures. The focal individuals (FIDs) in the study are the parents, in black and the recombination events studied are the ones transmitted in gametes from the FIDs to the offspring.

Linkage mapping

The SNP markers on each genotyping array have their physic positions mapped to the Atlantic salmon reference genome (assembly ICSASG_v2) (Lien *et al.*, 2016). Linkage mapping was done with the LepMap3 software (Rastas, 2017). The *filtering2* module was run as suggested for multi-family datasets with a *datatolerance* = 0.01 to filter markers based on segregation distortion. The *seperatechromosomes2* module was run within each linkage group and markers that were not assigned to the main group (i.e. LOD score < 5) were excluded, as suggested for analysis of species where chromosome assignments are well established. The number of markers in the final linkage map after these quality control steps can be found in Table 2. The *Ordermarkers2* module was run with the option to evaluate the given marker order, i.e. to calculate the centimorgan (cM) positions for the markers based on the pre-ordering of the markers using the Haldane mapping function option.

Fine-scale recombination rates along the genome

Fine scale recombination rates were measured as the cM distance per megabase (cM/Mb) within each 1Mb bin, where the cM distance was measured between the first and last SNP marker within each bin.

Individual recombination rates

Individual recombination rates were measured as autosomal crossover counts (ACC). Crossovers were counted from the gamete phase from the output of the *orderMarkers2* module, and assigned to each parent, i.e., FIDs in which the meiosis took place. There were a total of 145,590 observations from 621 unique females and 148,060 from 416 unique males.

Genetic shuffling within individuals

Genetic shuffling was calculated as the probability that a randomly chosen pair of loci was shuffled during gamete production (in meiosis) following the method suggested by Veller et al. (2019). We defined the parameter $\bar{r}$ as the genetic shuffling within a gamete generated by recombination alone and independent of shuffling between loci generated by independent assortment of chromosomes. The $\bar{r}$ for one gamete is calculated using the following equation (Veller et al. 2019):

$$\mathbb{E}[\bar{r}] = \sum_{k=1}^{n} 2p_k(1 - p_k)\, L_k^2$$

where k is the chromosomes 1-29, p_k is the proportion of grand-paternal alleles on chromosome k, $1\text{-}p_k$ is the proportion of grand-maternal alleles on chromosome k, and L_k is the length of the chromosome k as a fraction of the total length of the genome. For each phased offspring gamete in the output from the *Ordermarkers2* module in Lepmap3, $\bar{r}$ assigned as an observation to the paternal or maternal FID, resulting in multiple observations of $\bar{r}$ for each FID.

Genetic variance in measures of recombination

Variance components for individual ACC and $\bar{r}$ were estimated in DMU v6 (Madsen *et al.*, 2014) with a repeatability model using the Restricted Maximum Likelihood (REML) and average information (AI) algorithm. The model was:

$$\mathbf{Y = sex + b_1{*}age + id1 + id2 + b_2{*}het + e}$$

where Y is the ACC or $\bar{r}$, **sex** is the fixed effect of sex, **id1** is the random additive genetic effect of the FID, **id2** is the random effect of the FID permanent environment (i.e. environmental effects that are constant across repeated measures on an FID), **het** is the method-of-moments F coefficient estimates calculated with the *--het* function in PLINK1.9 (Chang *et al.*, 2015), $\mathbf{b_2}$ is the regression of ACC or $\bar{r}$ on **het** of the FID, and **e** is the residual effect. The narrow-sense heritability (h^2) was defined as the proportion of phenotypic variance explained by the additive genetic effect and was estimated separately for each sex. Repeatability was measured as the sum of the genetic variance and permanent environment variance divided by the total phenotypic variance.

Genome-Wide Associations between markers and measures of recombination

The MLMA-LOCO (Leave One Chromosome Out) module implemented in in GCTA (Yang *et al.*, 2011) was used to look for loci associated with the mean ACC and $\bar{r}$ per individual. The module runs a mixed linear model-based association where the chromosome that the candidate marker is located on is left out of the calculation of the genomic relationship matrix, to correct for relatedness. The significance threshold equivalent to P = 0.05 after correcting for multiple testing was set as P = 0.05/Number of markers.

Results

Linkage mapping

The sex-specific autosomal linkage maps spanned a total of 2294.3 cM in females and 1593.5 cM in males, with the female map 1.44 times longer than that of males (Table 2, Figure 2). The biggest sex differences are on chromosomes 2, 8 and 17, where the female maps are 10.08, 12.27 and 6.86 times longer than the male maps, respectively. These large differences agree with a previous linkage mapping study in Atlantic salmon (Lien et al. 2011). It is likely that these large differences in length are because we are unable to detect the majority of male recombination events on these chromosomes, as the map lengths are only 4-10 cM long and therefore much shorter than the minimum bound of 50cM for linkage map lengths as a result of obligate crossing-over during meiosis. In the rest of the chromosomes, the sex difference is between 1 and 2.2 fold longer in females, and with slightly longer male maps for chromosomes 21, 23 and 24 (Table 2, Figure 2).

Table 2. Summary of linkage map results by sex and chromosome.

Chr	Physical length (Mb)	Male linkage map length (cM)	Female linkage map length (cM)	Female/ Male Ratio	Male recombination rate (cM/Mb)	Female recombination rate (cM/Mb)
1	159.04	109.94	155.49	1.41	0.69	0.98
2	72.94	10.04	101.17	10.08	0.14	1.39
3	92.50	59.95	114.38	1.91	0.65	1.24
4	82.40	64.43	105.69	1.64	0.78	1.28
5	80.50	65.80	110.47	1.68	0.82	1.37
6	87.04	56.83	113.38	1.99	0.65	1.30
7	58.79	63.99	114.36	1.79	1.09	1.95
8	26.43	4.26	52.29	12.27	0.16	1.98
9	141.71	49.82	110.48	2.22	0.35	0.78
10	116.14	72.43	85.56	1.18	0.62	0.74
11	93.89	58.19	81.51	1.40	0.62	0.87
12	91.88	62.24	83.73	1.35	0.68	0.91
13	107.76	73.51	82.42	1.12	0.68	0.76
14	93.90	75.88	80.35	1.06	0.81	0.86
15	103.96	56.09	80.30	1.43	0.54	0.77
16	87.80	60.04	71.23	1.19	0.68	0.81
17	57.68	9.08	62.25	6.86	0.16	1.08
18	70.70	60.19	64.44	1.07	0.85	0.91
19	82.98	52.68	59.84	1.14	0.63	0.72
20	86.80	58.04	59.78	1.03	0.67	0.69
21	58.02	53.77	53.67	1.00	0.93	0.93
22	63.42	53.56	61.69	1.15	0.84	0.97
23	49.85	64.04	54.76	0.86	1.28	1.10
24	48.65	55.14	55.09	1.00	1.13	1.13
25	51.48	47.81	58.74	1.23	0.93	1.14
26	47.90	41.06	52.60	1.28	0.86	1.10
27	43.94	48.97	51.82	1.06	1.11	1.18
28	39.60	51.43	58.64	1.14	1.30	1.48
29	42.49	54.30	58.14	1.07	1.28	1.37
Total	2,240.20	1,593.50	2,294.27	1.44	0.71	1.02

Total is the total lengths and rates of the 29 chromosomes. The physical lengths (Mb) of chromosomes are taken from the Atlantic salmon reference genome (assembly ICSASG_v2) (Lien *et al.*, 2016).

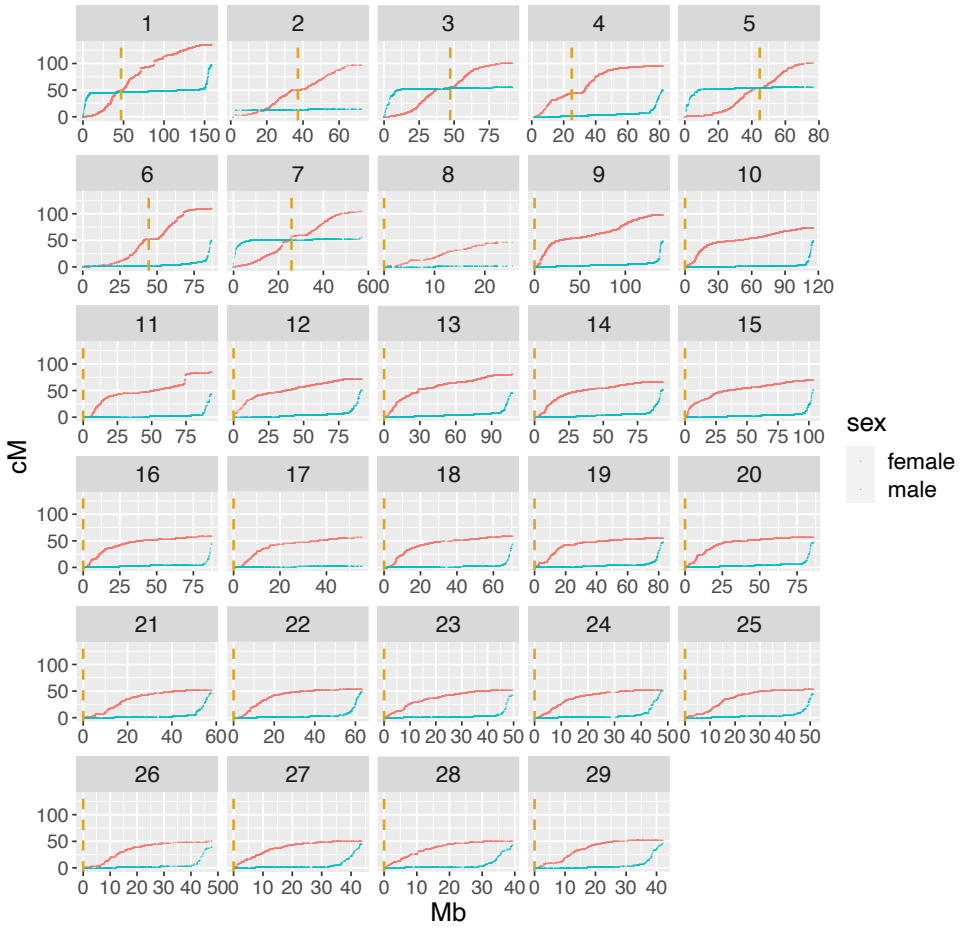


Figure 2. Male and female linkage maps for the 29 Atlantic salmon chromosomes. The physical position in Mb is plotted against the genetic position in cM for the SNP markers within each linkage group. Female positions are in red and male positions are in blue. The vertical dashed line in yellow is the centromere position from Lien et al (2016)

Fine scale recombination rates along the genome

The pattern of recombination rate across the genome is strikingly different between males and females in agreement with previous studies (Lien *et al.*, 2011; Gonen *et al.*, 2014). Male recombination rates are highly elevated in the sub-telomeric regions with an almost complete absence of recombination across the rest of the chromosomes (Figure 3). Conversely, female recombination is more evenly distributed along the genome, is greatly reduced in sub-

telomeric regions, and is slightly elevated in some regions closer to the centromeres (Figure 3). Recombination is suppressed in immediate proximity to the centromeres in both sexes on almost all chromosomes.

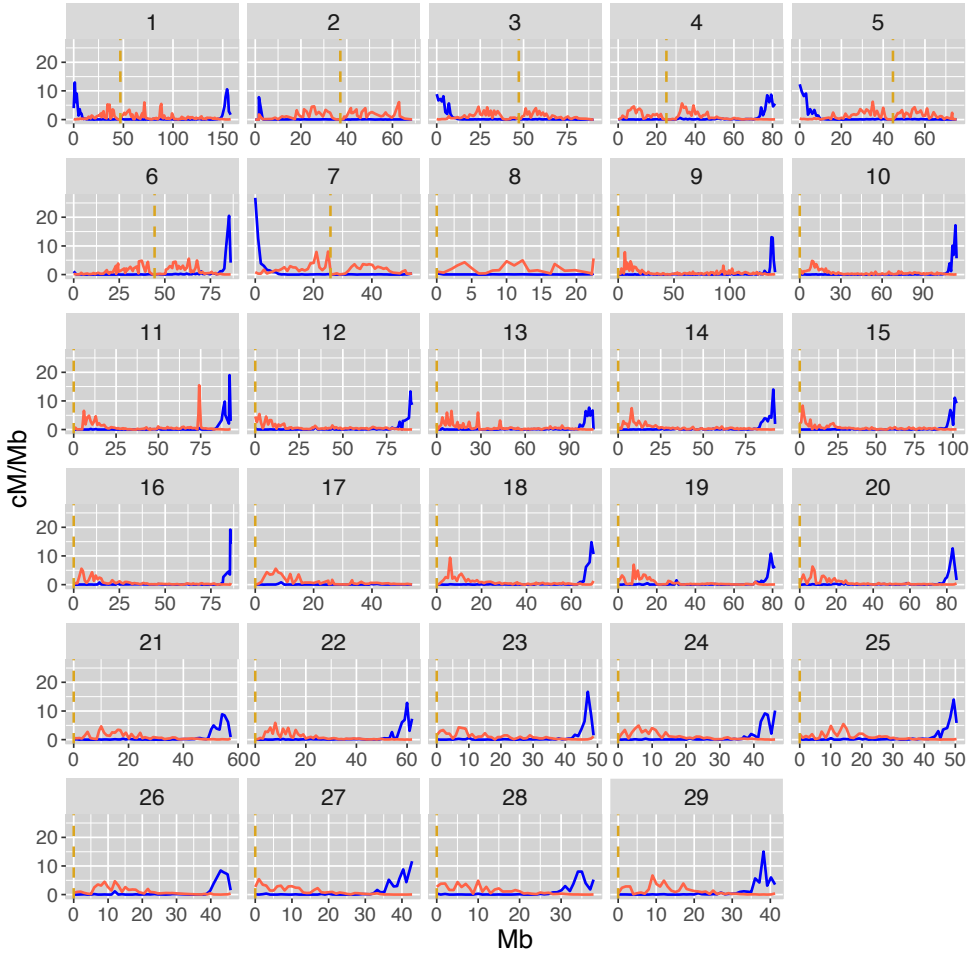


Figure 3. Fine-scale sex-specific recombination rate along the 29 Atlantic salmon chromosomes. The recombination rate in cM/Mb within each 1Mb bin, with males in blue and females in red. The dashed vertical lines in yellow are the centromere positions from Lien et al. (2016).

Individual crossover rates and genetic shuffling.

Both the autosomal crossover counts (ACC) and intra-chromosomal genetic shuffling values ($\bar{r}$) were approximately normally distributed in both males and females. The mean ACC, with standard deviation in parentheses, was 19.7 (4.7) in females and 11.9 (3.8) in males (Figure 4). Means and standard deviation for $\bar{r}$ was 0.00759 (0.002.06) for females and 0.00113 (0.00068) for males. (Figure 5, Table 3)

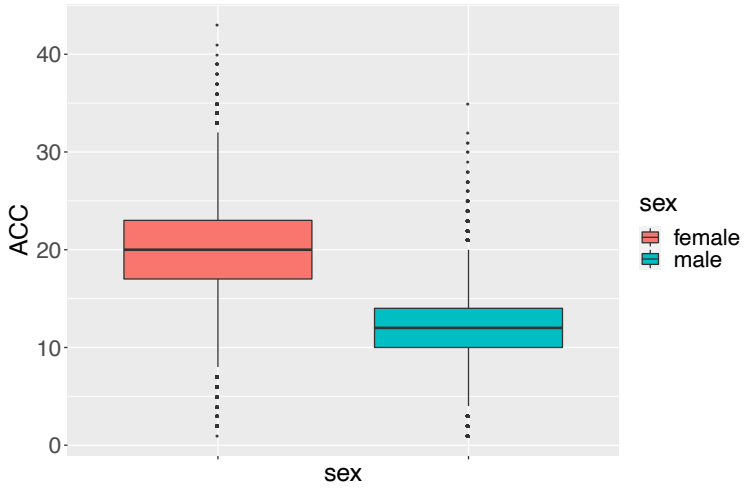


Figure 4. Distribution of autosomal crossover count (ACC) for each sex.

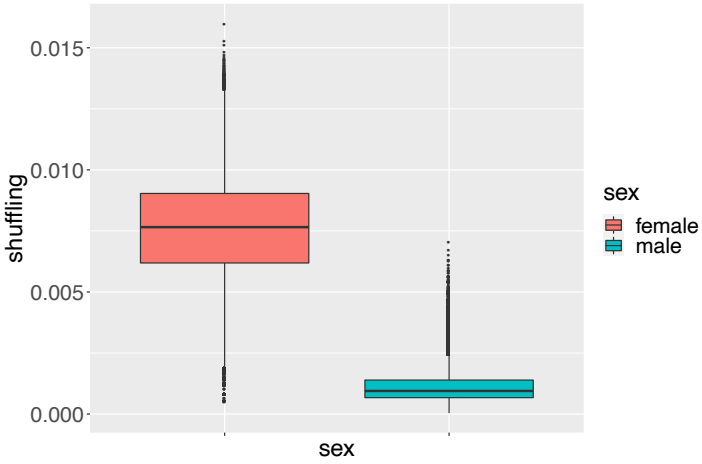


Figure 5. Distribution of intrachromosomal genetic shuffling $\bar{r}$ in each sex.

Heritability and repeatability of ACC and $\bar{r}$

The heritability (h^2) for ACC was 0.12 (SE = 0.03) in males and 0.11 in females (SE = 0.02; Table 3). In both sexes, most of the phenotypic variance is explained by the error term. There was an effect of inbreeding on ACC, where higher inbreeding coefficients were associated with reduced crossover counts. Genetic shuffling $\bar{r}$ was significantly heritable in females (h^2 =0.16, SE = 0.03), but not in males (h^2 = 0.05, SE = 0.06). The effect of inbreeding on $\bar{r}$ is lower than on ACC in females, whereas in males higher inbreeding has the opposite effect, leading to slightly lower $\bar{r}$. The genetic correlations between ACC and $\bar{r}$ was 0.86 (0.01) in females, but only 0.42 (0.05) in males. The genetic correlations between male and female ACC and $\bar{r}$ was 0.17 (0.17) and 0.11 (0.15), respectively.

Table 3. Results from variance component estimation of ACC and $\bar{r}$ (x1000).

Trait	sex	N_{FIDs}	N_{obs}	mean (sd)	h^2(SE)	effect F
ACC	female	621	145590	19.71 (4.65)	0.11 (0.02)	-12.94
ACC	male	416	148060	11.92 (3.81)	0.12 (0.03)	-5.85
$\bar{r}$	female	621	145590	7.59 (2.06)	0.16 (0.03)	-5.14
$\bar{r}$	male	416	148060	1.13 (0.68)	0.05 (0.06)	0.34

N_{FIDs} are the total number of FIDs (with repeated observations) and N_{obs} is the total number of observations (i.e. gametes) for each sex. Mean is the mean ACC or $\bar{r}$ (x1000) with standard deviations in parentheses. h^2 is the heritability estimate with standard errors in parentheses. Effect F is the estimated effect of inbreeding for each trait and sex.

Table 4. Genetic and phenotypic correlations between male and female ACC and $\bar{r}$.

	Male $\bar{r}$	Female $\bar{r}$	Male ACC	Female ACC	
Male $\bar{r}$		-	0.13	-	*Phenotypic correlations*
Female $\bar{r}$	0.11 (0.15)		-	0.74	
Male ACC	0.42 (0.05)	-		-	
Female ACC	-	0.86 (0.01)	0.17 (0.17)		
Genetic correlations					

Genetic correlations with standard error in parenthesis between the two traits for each sex and between males and females for both traits in the lower triangle. Phenotypic correlations between the two traits for each sex in the upper triangle.

Genome wide association analyses of ACC and $\bar{r}$

Two SNP loci were significantly associated with male ACC. The top SNP occurred at ~132 Mb on chromosome 1 (slope = 1.886, SE = 0.359, P = 1.47×10^{-7}; Table 5, Figure 6). There were no clear candidate genes with functions related to meiosis. The locus *SWI5*, which is a homologous recombination repair gene (Gene ID: 100195612, Benson et al., 2017) is 7.7 Mb downstream from this marker. The other significant SNP occurred at ~33Mb on chromosome 11 (slope = 2.109, SE = 0.416, P = 4.05×10^{-7}; Table 5, Figure 6). However, again, there were no clear candidate genes with functions related to meiosis. There were no further significant SNPs associated with ACC or $\bar{r}$ in males or females in this study (see Table 5 for the top 5 SNPs for each trait; full results are plotted in Figures 6 & 7). However, it is worth noting that for male $\bar{r}$, the two top SNPs (located at ~17 and 20Mb on chromosome 11) have a meiotic recombination protein coding gene, *REC114*, between them (Gene ID: 1065621289, Benson et al., 2017).

Table 5. Top five SNPs from genome-wide association analysis with Male and female autosomal crossover count (ACC) and intrachromosomal shuffling (r).

Sex	Trait	Chr	SNP	bp	MAF	Beta (SE)	P
Male	ACC	1	ctg7180001873044_6275_SAC*	132926135	0.05	1.886 (0.359)	1.47E-07
		11	ctg7180001855396_4974_SAG*	32967616	0.03	2.109 (0.416)	4.05E-07
		6	ctg7180001878488_1814_SAG	24177632	0.07	1.305 (0.290)	6.89E-06
		1	ctg7180001868490_5665_SCT	30757854	0.10	1.017 (0.234)	1.37E-05
		3	ctg7180001881801_5357_SAG	68273136	0.03	1.894 (0.444)	1.96E-05
	$\bar{r}$	11	ctg7180001809099_11901_SCT	20172905	0.20	0.268 (0.061)	1.27E-05
		11	ctg7180001928069_1860_SAG	17951854	0.23	0.255 (0.059)	1.42E-05
		13	ctg7180001890089_1886_SCT	84699720	0.05	0.451 (0.104)	1.44E-05
		21	ctg7180001808275_4599_SAC	16403453	0.20	0.248 (0.058)	1.76E-05
		20	ctg7180001910409_1746_SGT	78888079	0.09	0.334 (0.082)	4.85E-05
Female	ACC	15	ctg7180001922064_3259_SAG	100964747	0.38	-0.800 (0.170)	2.53E-06
		18	ctg7180001924742_3375_SGT	29080197	0.26	0.823 (0.177)	3.49E-06
		19	ctg7180001836738_1599_SAG	26049954	0.06	2.381 (0.513)	3.50E-06
		18	ctg7180001566763_174_SAG	15856593	0.23	0.815 (0.192)	2.25E-05
		12	ctg7180001302330_2749_SAC	85852951	0.28	-0.776 (0.184)	2.53E-05
	$\bar{r}$	9	ctg7180001896984_3593_SGT	121046130	0.26	-0.002 (0.000)	2.31E-06
		9	ctg7180001896984_3747_SAC	121046284	0.26	-0.002 (0.000)	2.41E-06
		9	ctg7180001833464_333_SAG	129063540	0.22	-0.002 (0.000)	6.50E-06
		9	ctg7180001850551_2639_SCT	129630313	0.34	-0.002 (0.000)	6.95E-06
		5	ctg7180001819026_9900_SAC	49941475	0.08	-0.003 (0.001)	2.26E-05

SNP is the SNP marker names; bp is the base pair position of the marker; MAF is the minor allele frequency of this marker; beta is the additive effect size (i.e. the slope) of the effect allele with standard error in parentheses; P is the P value for the marker. *Marker was significantly associated with the phenotype (equivalent to $P < 0.05$ after correcting for multiple testing).

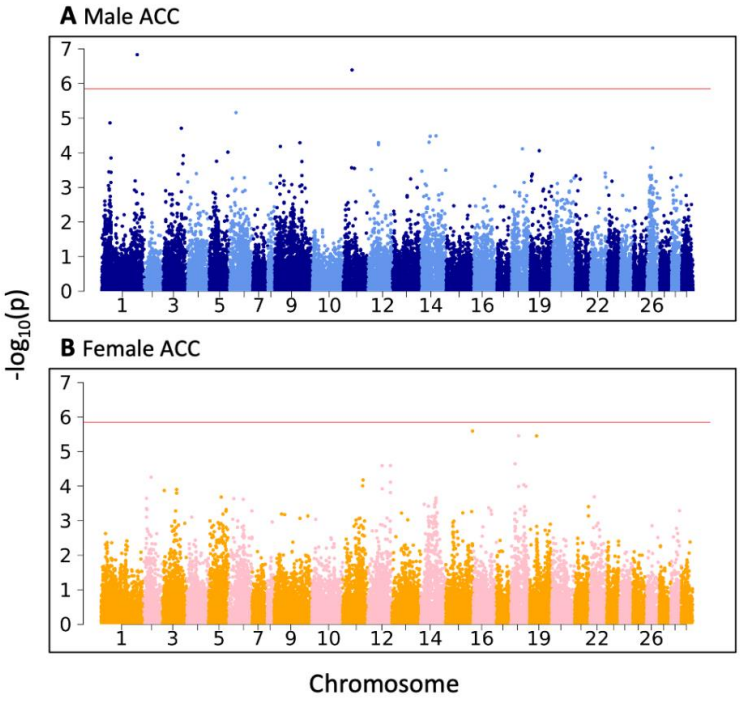


Figure 6. Manhattan plots of genome-wide association between markers and autosomal crossover count (ACC) for males (A) and females (B). The red line is the genome-wide significance threshold equivalent to P = 0.05 after accounting for multiple testing. The y-axis is the negative $\log_{10}$ of the p-values, and the x-axis is the physical positions of the markers with alternating colors for chromosomes 1-29.

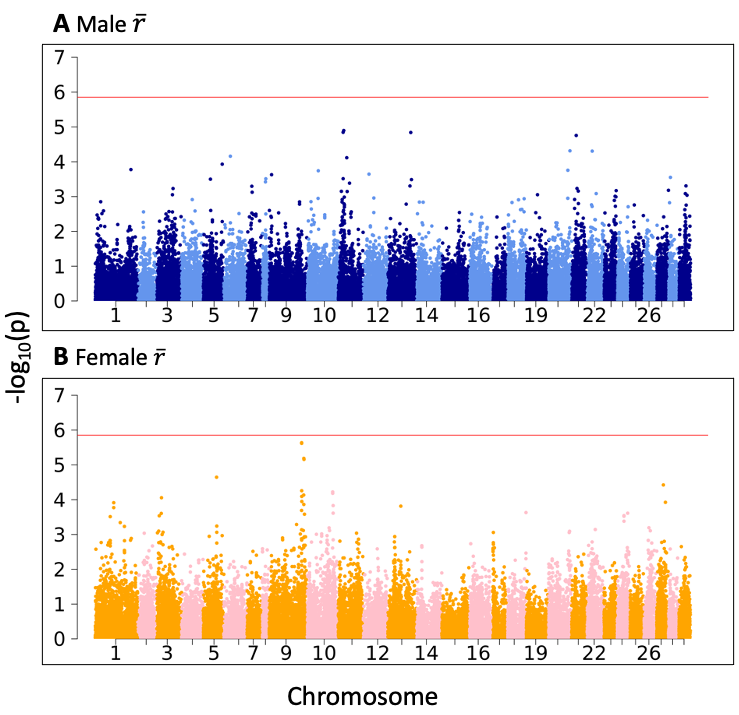


Figure 7. Manhattan plots of genome-wide association between markers and intrachromosomal shuffling ($\bar{r}$) for males (A) and females (B). The red line is the genome-wide significance threshold equivalent to P = 0.05 after accounting for multiple testing. The y-axis is the negative $\log_{10}$ of the p-values, and the x-axis is the physical positions of the markers with alternating colors for chromosomes 1-29.

Discussion

In this study, we confirm sex differences in genome wide recombination rate and substantial differences in recombination landscapes between males and females in Atlantic salmon. We show that there is variation in individual recombination rates and that they are heritable in both males and females, but with a low genetic correlation between the sexes, suggesting that parts of the genetic architecture of the trait is sex-specific. We also measure individual intra-chromosomal shuffling of loci and find that this is also a significantly heritable trait in females, but not in males. Females contributed substantially more genetic shuffling of alleles from one generation the next due to higher rates of recombination and more even distribution of crossover locations. Two regions of the genome were associated with male autosomal crossover count but did not correspond to clear candidate genes; there were no significant associations with female crossover rate or genetic shuffling in both sexes. Here, we explore the results in more detail and discuss how the findings may be relevant in the breeding work on Atlantic salmon as well as contributing to the general understanding of variation in recombination rates and patterns.

Sexual dimorphism in recombination rates and patterns.

The sex-specific linkage maps show a female-biased heterochiasmy in salmon and a strikingly large differences in the distribution of crossovers along the chromosome. This agrees with previous studies (Lien *et al.*, 2011; Gonen *et al.*, 2014). The linkage maps in this study do not have a higher number of markers than already published linkage maps of Atlantic salmon, but the number of individuals used is much higher, capturing a higher number of meioses and allowing more fine-scale resolution of distances between markers. Figure 2 shows that the markers are well-distributed without large jumps of cM distance between adjacent markers and/or inflation in the genetic length, which can be indicative of

mapping errors. The total length in cM of the male and female linkage maps is shorter than previous studies have found (Lien *et al.*, 2011; Gonen *et al.*, 2014) which may be due to more accurate physical positions of the markers. However, the sex difference (1:1.38) is the same as determined by Lien et al (2011). On the chromosomes with the largest sex differences in total genetic length, the male maps show little to no recombination in one or two sub-telomeric regions (i.e. chromosomes 2, 8 and 17; Figure 2). These regions coincide with regions reported to have > 90% sequence similarity with regions on other chromosomes (Lien *et al.*, 2016). The combination of both high similarity and some reduction of recombination between homologues suggests that these chromosomes may be experiencing delayed rediploidization and are forming quadrivalents during meiosis (Lien *et al.*, 2016). Cytological studies in different salmonid species find that multivalent pairing happens primarily in males (Timusk *et al.*, 2011; Allendorf *et al.*, 2015). The high sequence similarity and tetrasomic inheritance also make these areas difficult to genotype and to map, and these regions also coincide with regions characterised by low marker density in our dataset (Figure S1). Therefore, we cannot exclude that there might be recombination events occurring in these regions that we cannot detect, suggesting that genome wide male recombination rates may be underestimated in Atlantic salmon.

Genetic variation in individual measures of recombination

To the best of our knowledge, our study provides the first estimates of additive genetic effects for individual measures of recombination in Atlantic salmon. The heritability of ACC is moderate, with measures of 0.12 in males and 0.11 in females. Estimates in other species range from around 0.05 – 0.18 in pigs (Johnsson *et al.*, 2021), sheep (Johnston *et al.*, 2016; Petit *et al.*, 2017), cattle (Kadri *et al.*, 2016) and red deer (Johnston, Huisman and Pemberton, 2018) to as high as 0.41 and 0.46 in some *Drosophila* strains (Hunter *et al.*, 2016) and mouse

lines (Dumont, Broman and Payseur, 2009) respectively. The genetic correlation between male and female ACC was positive but low ($r_A = 0.17$), indicating that there are different genetic architectures affecting crossover rates in males and females. This is perhaps not surprising, considering the large differences in recombination patterns between the sexes. Using a genome-wide association analysis, we identified two SNP markers that exceeded the significance threshold in for male ACC. However, some caution must be exerted in interpreting this finding. First, these loci had relatively low minor allele frequencies (MAF ≤ 0.05); as using mean values led to a relatively low sample size for the GWA analysis, this can increase the probability of obtaining false positive results at rare alleles due to sampling effects. Secondly there are no candidate genes associated with meiosis in close proximity to these two SNPs.

Studies in other systems imply that there are a number of moderate to large effect loci controlling recombination, so perhaps it is surprising not to detect loci for this trait. However, those studies were carried out in mammals, where conservation of genetic architecture may be stronger than between mammals and fish (Sandor *et al.*, 2012; Johnston *et al.*, 2016; Halldorsson *et al.*, 2019; Johnsson *et al.*, 2021). We may have reduced power in our data to identify SNP associations: first, there is relatively low marker density in some areas of the genome, particularly in regions with high sequence similarity between chromosomes (Lien et al 2016); and second, the individual sample sizes are comparatively low to other GWAS studies, with 416 and 621 FIDs with phenotype in males and females respectively. A next step would be to measure this trait in more individuals genotyped on a higher density chip, and to expand the GWAS to incorporate multiple measures of recombination per individual, rather than using the individual means as in this study; this was not carried out here due to time constraints and the analysis being highly computationally

intensive. Furthermore, a newly released reference genome has enabled development of SNP arrays with better physical mapping of a high number of markers, which may mitigate the issues of marker density.

In addition to the genes affecting recombination rate often found in mammals (see previously cited examples), future work could also focus on candidate genes for variation in males associated with homologous vs. homeologous chromosome pairing. For example, in both hexaploid bread wheat and tetraploid pasta wheat, a major gene *Ph1* has been found to control correct pairing of homologous chromosomes in meiosis, resulting in a stable diploid meiosis despite genome duplication (Griffiths *et al.*, 2006). This could be done by mapping homeologous recombination as carried out in previous studies of salmon (Waples, Seeb and Seeb, 2016) and investigating loci associated with potential variation in the number of homologous vs. homeologous recombination events.

Sexual dimorphism in the shuffling of alleles

Intra-chromosomal shuffling ($\bar{r}$) is significantly heritable in females but not in males. This trait reflects genetic variation in crossover locations, but high phenotypic and genetic correlations between $\bar{r}$ and ACC in females (0.74 & 0.86, respectively; Table 4) indicates that these traits are not independent, where the distribution of crossovers along the chromosome is more evenly distributed. However, in males there is relatively little intra-chromosomal shuffling (in both the mean and variance) due to recombination occurring almost exclusively at the sub-telomeric regions. When more than one crossover occurs between a pair of chromatids with an even distribution as in females, the additional crossover will result in more shuffling (Veller, Kleckner and Nowak, 2019), but if two crossovers happen either in very close proximity or at the chromosome ends, most shuffling is essentially eliminated,

which may explain the differences in correlations between the two traits between males and females. In females, it may be of interest to better define a trait that aims to pick up purely the genetic variation in crossover position, as this seems to be the main driver of the observed difference in shuffling and could potentially show more accurate association with loci affecting distribution of crossovers along the genome.

Most of the variation of ACC and $\bar{r}$ remains unexplained

We find that most of the phenotypic variance is explained by the error term. When studying recombination events in liveborn offspring, we cannot obtain information from of all the products of meiosis. In females, only ¼ of chromatids end up in an oocyte and studies in human suggest that there is variation in number of crossovers between oocytes and polar bodies, suggesting selection against non-recombinants in the final product of meiosis (Ottolini *et al.*, 2015) Or there may be a potential role of meiotic drive(Brandvain and Coop, 2012) Similarly in males, it is unlikely that more than one of the four products of a meiosis will be sampled in offspring. In both sexes, the release of eggs and sperm in external fertilization also introduces the potential for selection to act on recombination during the haploid and/or zygote phase, such as against aneuploid gametes (as a result of lower recombination) or those that have accumulated mutations or unfavourable combinations of alleles (as a result of higher recombination). Therefore, even in the absence of selection, gametes only represent a sample of recombinant and non-recombinant chromatids assigned to the gamete by Mendelian segregation, meaning that ACC and $\bar{r}$ are only a proxy for the true rate of recombination and intra-chromosomal shuffling, which may be incorporated into the error variance in the animal models.

What are the implications for animal breeding?

Investigating recombination and the rate of genetic shuffling is of interest for breeding as it can lead to the creation of novel haplotypes to exploit in selection, which is the essence of breeding in animals and plants (Battagin *et al.*, 2016). Here, the marked differences in recombination patterns between the sexes leads to considerable differences in the probability of shuffling between linked loci in male and female meioses. However, each offspring born in a population will always inherit exactly one paternal and one maternal gamete, and in the subsequent generation alleles should segregate independently of which sex they were transmitted from in previous generations. Therefore, the genetic shuffling on population level may be more affected by the variation within males and females, when a limited number of individuals is contributing to the next generation.

Our finding that ACC and $\bar{r}$ are heritable suggest a potential to increase the amount of recombination and genetic shuffling via selection on standing genetic variation within the population (Battagin *et al.*, 2016). However, little is known about the potential consequences of selecting for more shuffling. Despite some evidence for a potential to increase genetic gain with higher recombination rates (Battagin *et al.*, 2016), selecting for higher recombination rates may not be beneficial. The negative consequences of extensive recombination rates remain to be understood, but increased mutation rates (Halldorsson *et al.*, 2019) combined with the fact that there seems to be an upper limit for number of crossovers per chromosome shared among species across a broad range of taxa (Ritz, Noor and Singh, 2017), suggests that the negative effects of recombination. Similarly, "recombination load" may occur through breaking up beneficial allele combinations, potentially outweighing the positive effects of increased rate (Charlesworth and Barton , 1996). The mechanisms leading to the observed variation in shuffling are probably mechanisms controlling crossover locations (as

discussed above) and particularly in males, the factors controlling the amount of shuffling may be genes associated with homologue vs homeologue pairing. It would be of interest to further explore potential variation in the amount of multivalent pairing in males, and to explore the potential to select for preferred homologue pairing and induce stable diploidy in meiosis, as in wheat. More generally, our findings of a clear difference between male and female genetic shuffling demonstrates how important the location of the crossover is for the probability of generating novel combinations of linked loci, a factor that may be frequently overlooked in breeding.

Conclusions

In conclusion, this study shows that there is genetic variation in genome-wide rates and distribution of crossover events in the Atlantic salmon. Consistent with other studies, most of the phenotypic variance is explained by the error variance, suggesting that studying crossover events in offspring is not adequate to fully understand the genetic mechanisms controlling the number and location of crossovers during meiosis. Potentially more important in breeding, there are marked differences in the genetic shuffling produced by males and females, but more studies should aim to include the inter-chromosomal shuffling in males as a consequence of the multivalent pairing in male meiosis. Overall, these findings provide a basis to understanding the evolution of recombination rates and distribution in Atlantic salmon specifically, and more generally in species going through rediploidization.

References

Allendorf, F.W. *et al.* (2015) "Effects of crossovers between homeologs on inheritance and population genomics in polyploid-derived salmonid fishes," *Journal of Heredity*. Oxford University Press, pp. 217–227. doi:10.1093/jhered/esv015.

Barton, N.H. and Charlesworth, B. (1998) *Why Sex and Recombination?* doi:10.1126/science.281.5385.1986.

Battagin, M. *et al.* (2016) "Effect of manipulating recombination rates on response to selection in livestock breeding programs," *Genetics Selection Evolution*, 48(1). doi:10.1186/s12711-016-0221-1.

Baudat, F. *et al.* (2010) "PRDM9 Is a Major Determinant of Meiotic Recombination Hotspots in Humans and Mice," *Science*, 327(5967), pp. 836–840. doi:10.1126/science.1183439.

Benson, D.A. *et al.* (2017) "GenBank," *Nucleic Acids Research*, 45(D1), pp. D37–D42. doi:10.1093/nar/gkw1070.

Brandvain, Y. and Coop, G. (2012) "Scrambling eggs: Meiotic drive and the evolution of female recombination rates," *Genetics*, 190(2), pp. 709–723. doi:10.1534/genetics.111.136721.

Burt, A. and Bell, G. (1987) "Mammalian chiasma frequencies as a test of two theories of recombination," *Nature*, 326, pp. 803–805.

Burt, A., Bell, G. and Harvey, P.H. (1991) *Sex differences in recombination*, *J. evol. Biol*. doi:https://doi.org/10.1046/j.1420-9101.1991.4020259.x.

Chang, C.C. *et al.* (2015) "Second-generation PLINK: Rising to the challenge of larger and richer datasets," *GigaScience*, 4(1). doi:10.1186/s13742-015-0047-8.

Charlesworth, B. and Barton, N.H. (1996) "Recombination load associated with selection for increased recombination," *Genetical Research*, 67(1), pp. 27–41. doi:10.1017/s0016672300033450.

Cooney, C.R., Mank, J.E. and Wright, A.E. (2021) "Constraint and divergence in the evolution of male and female recombination rates in fishes," *Evolution*, 75(11), pp. 2857–2866. doi:10.1111/evo.14357.

Dumont, B.L., Broman, K.W. and Payseur, B.A. (2009) "Variation in genomic recombination rates among heterogeneous stock mice," *Genetics*, 182(4), pp. 1345–1349. doi:10.1534/genetics.109.105114.

Dunn, L.C. and Bennett, D. (1967) "Sex differences in recombination of linked genes in animals," *Genetical Research*, 9(2), pp. 211–220. doi:10.1017/S0016672300010491.

Fledel-Alon, A. *et al.* (2009) "Broad-scale recombination patterns underlying proper disjunction in humans," *PLoS Genetics*, 5(9). doi:10.1371/journal.pgen.1000658.

Gonen, S. *et al.* (2014) "Linkage maps of the Atlantic salmon (Salmo salar) genome derived from RAD sequencing," *BMC Genomics*, 15(1). doi:10.1186/1471-2164-15-166.

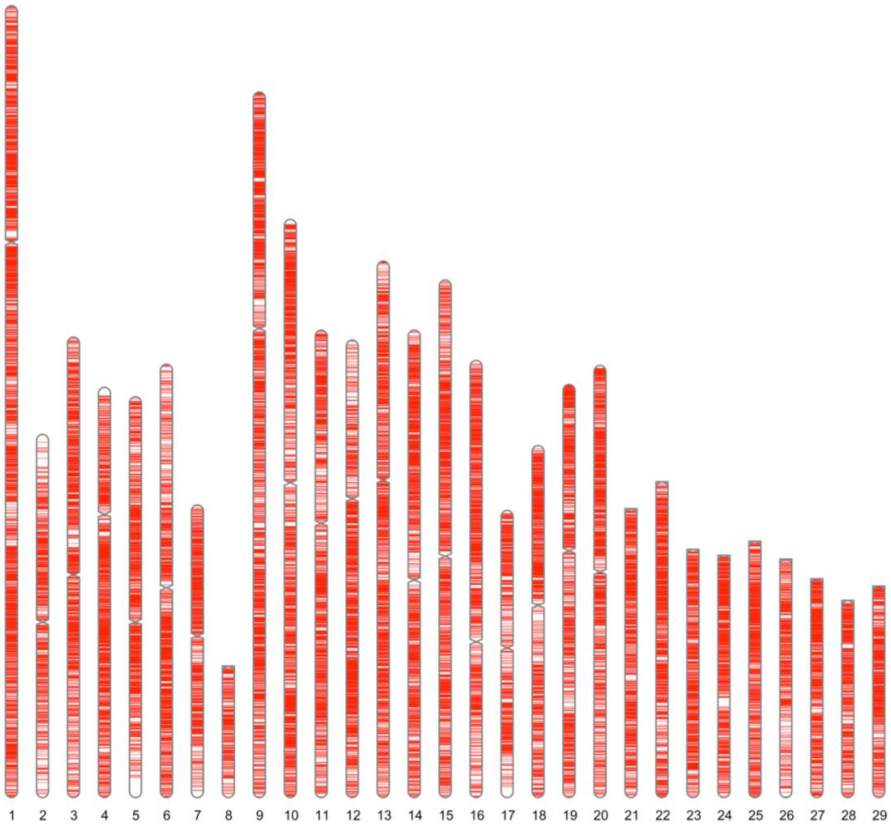
Griffiths, S. *et al.* (2006) "Molecular characterization of Ph1 as a major chromosome pairing locus in polyploid wheat," *Nature*, 439(7077), pp. 749–752. doi:10.1038/nature04434.

Halldorsson, B. v. *et al.* (2019) "Human genetics: Characterizing mutagenic effects of recombination through a sequence-level genetic map," *Science*, 363(6425). doi:10.1126/science.aau1043.

Hassold, T. *et al.* (1995) *Recombination and Maternal Age-Dependent Nondisjunction: Molecular Studies of Trisomy 16*, *Am. J. Hum. Genet*.
Hunter, C.M. *et al.* (2016) "The Genetic Architecture of Natural Variation in Recombination Rate in Drosophila melanogaster," *PLoS Genetics*, 12(4). doi:10.1371/journal.pgen.1005951.
Johnsson, M. *et al.* (2021) "Genetic variation in recombination rate in the pig," *Genetics Selection Evolution*, 53(1). doi:10.1186/s12711-021-00643-0.
Johnston, S., Stoffel, M. and Pemberton, J. (2020) "Variants at RNF212 and RNF212B are associated with recombination rate variation in Soay sheep (Ovis aries)." doi:10.1101/2020.07.26.217802.
Johnston, S.E. *et al.* (2016) "Conserved genetic architecture underlying individual recombination rate variation in a wild population of soay sheep (Ovis aries)," *Genetics*, 203(1), pp. 583–598. doi:10.1534/genetics.115.185553.
Johnston, S.E., Huisman, J. and Pemberton, J.M. (2018) "A genomic region containing REC8 and RNF212B is associated with individual recombination rate variation in a wild population of red deer (Cervus elaphus)," *G3: Genes, Genomes, Genetics*, 8(7), pp. 2265–2276. doi:10.1534/g3.118.200063.
Kadri, N.K. *et al.* (2016) "Coding and noncoding variants in HFM1, MLH3, MSH4, MSH5, RNF212, and RNF212B affect recombination rate in cattle," *Genome Research*, 26(10), pp. 1323–1332. doi:10.1101/gr.204214.116.
Koehler, K.E. *et al.* (1996) *Recombination and nondisjunction in humans and flies*, *Human Molecular Genetics*. Oxford University Press. Available at: https://academic.oup.com/hmg/article/5/Supplement_1/1495/663563.
Kong, A. *et al.* (2008) "Sequence Variants in the RNF212 Gene Associate with Genome-Wide Recombination Rate," *Science*, 319(5868), pp. 1398–1401. doi:10.1126/science.1152422.
Lenormand, T. and Dutheil, J. (2005) "Recombination difference between sexes: A role for haploid selection," in *PLoS Biology*, pp. 0396–0403. doi:10.1371/journal.pbio.0030063.
Lien, S. *et al.* (2011) "A dense SNP-based linkage map for Atlantic salmon (Salmo salar) reveals extended chromosome homeologies and striking differences in sex-specific recombination patterns," *BMC Genomics*, 12. doi:10.1186/1471-2164-12-615.
Lien, S. *et al.* (2016) "The Atlantic salmon genome provides insights into rediploidization," *Nature*, 533, pp. 200–205. doi:10.1038/nature17164.
Ma, L. *et al.* (2015) "Cattle Sex-Specific Recombination and Genetic Control from a Large Pedigree Analysis," *PLoS Genetics*, 11(11). doi:10.1371/journal.pgen.1005387.
Madsen, P. *et al.* (2014) "DMU - A Package for Analyzing Multivariate Mixed Models in quantitative Genetics and Genomics," in *Journal of Applied Statistics*. Routledge, pp. 1286–1306.
Mank, J.E. (2009) "The evolution of heterochiasmy: The role of sexual selection and sperm competition in determining sex-specific recombination rates in eutherian mammals," *Genetics Research*, 91(5), pp. 355–363. doi:10.1017/S0016672309990255.
Moen, T. *et al.* (2008) "A linkage map of the Atlantic salmon (Salmo salar) based on EST-derived SNP markers," *BMC Genomics*, 9. doi:10.1186/1471-2164-9-223.
Nei, M. (1969) "Linkage modification and sex difference in recombination," pp. 681–689.

Ohno, S., Wolf, U. and Atkin, N.B. (1968) “Evolution from fish to mammals by gene D,” *Hereditas*, 59(1), pp. 169–187. Available at: https://doi.org/10.1111/j.1601-5223.1968.tb02169.x (Accessed: January 2, 2022).
Ottolini, C.S. *et al.* (2015) “Genome-wide maps of recombination and chromosome segregation in human oocytes and embryos show selection for maternal recombination rates,” *Nature Genetics*, 47(7), pp. 727–735. doi:10.1038/ng.3306.
Petit, M. *et al.* (2017) “Variation in recombination rate and its genetic determinism in sheep populations,” *Genetics*, 207(2), pp. 767–784. doi:10.1534/genetics.117.300123.
Rastas, P. *et al.* (2015) “Construction of Ultradense Linkage Maps with Lep-MAP2: Stickleback F 2 Recombinant Crosses as an Example,” *Genome Biology Evolution* [Preprint]. doi:10.5061/dryad.f8j71.
Rastas, P. (2017) “Lep-MAP3: Robust linkage mapping even for low-coverage whole genome sequencing data,” *Bioinformatics*, 33(23), pp. 3726–3732. doi:10.1093/bioinformatics/btx494.
Ritz, K.R., Noor, M.A.F. and Singh, N.D. (2017) “Variation in Recombination Rate: Adaptive or Not?,” *Trends in Genetics*. Elsevier Ltd, pp. 364–374. doi:10.1016/j.tig.2017.03.003.
Sakamoto, T. *et al.* (2000) *A Microsatellite Linkage Map of Rainbow Trout (Oncorhynchus mykiss) Characterized by Large Sex-Specific Differences in Recombination Rates*. doi:10.1093/genetics/155.3.1331.
Sandor, C. *et al.* (2012) “Genetic variants in REC8, RNF212, and PRDM9 influence male recombination in cattle,” *PLoS Genetics*, 8(7). doi:10.1371/journal.pgen.1002854.
Sardell, J.M. *et al.* (2018) “Sex differences in recombination in sticklebacks,” *G3: Genes, Genomes, Genetics*, 8(6), pp. 1971–1983. doi:10.1534/g3.118.200166.
Sherman, S.L. *et al.* (1991) *Trisomy 21: Association between Reduced Recombination and Nondisjunction*, *Am. J. Hum. Genet*.
Stapley, J. *et al.* (2017) “Variation in recombination frequency and distribution across eukaryotes: Patterns and processes,” *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1736). doi:10.1098/rstb.2016.0455.
Timusk, E.R. *et al.* (2011) “Genome evolution in the fish family salmonidae: Generation of a brook charr genetic map and comparisons among charrs (Arctic charr and brook charr) with rainbow trout,” *BMC Genetics*, 12. doi:10.1186/1471-2156-12-68.
Veller, C., Kleckner, N. and Nowak, M.A. (2019) “A rigorous measure of genome-wide genetic shuffling that takes into account crossover positions and Mendel’s second law,” *Proceedings of the National Academy of Sciences*, 116(5), pp. 1659–1668. doi:10.1073/PNAS.1817482116.
Wang, R.J. *et al.* (2017) “Recombination rate variation in mice from an isolated island,” *Molecular Ecology*, 26(2), pp. 457–470. doi:10.1111/mec.13932.
Waples, R.K., Seeb, L.W. and Seeb, J.E. (2016) “Linkage mapping with paralogs exposes regions of residual tetrasomic inheritance in chum salmon (Oncorhynchus keta),” *Molecular Ecology Resources*, 16(1), pp. 17–28. doi:10.1111/1755-0998.12394.
Yang, J. *et al.* (2011) “GCTA: A tool for genome-wide complex trait analysis,” *American Journal of Human Genetics*, 88(1), pp. 76–82. doi:10.1016/j.ajhg.2010.11.011.

S1.



SNP marker density along the 29 linkage groups. The position of each SNP marker is plotted in red.

Paper IV.

Variation in patterns of recombination result in genetic variation in intrachromosomal shuffling in the domestic pig.

Brekke, C., Johnston, S.E., Gjuvsland, A.B., Berg, P. (2022).

Submitted manuscript to WCGALP

Variation in patterns of recombination result in genetic variation in intrachromosomal shuffling in the domestic pig

C.Brekke[1*], S.E. Johnston[2], A.B. Gjuvsland[3] and P. Berg[1]

[1] Norwegian University of Life Sciences, Oluf Thesens vei 6, 1433 Ås, Norway; [2] University of Edinburgh, EH9 3FL, Edinburgh, United Kingdom; [3] Norsvin SA, Storhamargata 44, 2317 Hamar, Norway; [*]cathrine.brekke@nmbu.no

Abstract
Meiotic recombination leads to shuffling of loci located on the same chromosome. The amount of intrachromosomal shuffling from one generation to the next is affected by the number of crossovers, location of crossovers and crossover interference. In the domestic pig, *Sus Scrofa*, genome wide recombination rates are higher in females than in males. However, in this study we find that the genome wide intrachromosomal shuffling between pairs of loci is higher in males than in females due to difference in distribution of crossovers along the chromosome in the sexes. We show that this pattern is consistent in four of five different pig breeds and that there is a genetic component to the variation in genetic shuffling.

Introduction
Meiotic recombination is when crossing over occurs between the homologues maternal and paternal chromosomes during synapsis resulting in gametes with novel haplotypes. This event breaks down linkage disequilibrium and lead to haplotypic diversity that can be exploited in selection. Recombination also has a vital role in the proper alignment and segregation of homologues chromosomes. (Sherman et al., 1991; Koehler et al., 1996; Hassold et al., 1995; Fledel-Alon et al., 2009). Recombination can however also break up beneficial linkage previously built up by selection (Charlesworth and Barton, 1996), and there is evidence for increased mutation rates in recombination hotspots (Halldorsson et al., 2019; Arbeithuber et al., 2015). Recombination rates vary between taxa and species, and even within and between closely related populations (reviewed in Ritz et al., 2017; and Stapley et al., 2017). Most species show some level of heterochiasmy (Burt et al., 1991) and in some species there is also a substantial sexual dimorphism in the distribution of crossovers along the genome (Sakamoto et al., 2000; Lien et al., 2011; Tortereau et al., 2012; Johnston et al., 2016). In breeding this variation is of interest because it affects the production of novel allelic combinations. Veller et al (2019) suggests shuffling of maternal and paternal alleles from one generation to the next as an alternative measure that pics up both the number and distribution of crossovers. The aim of this study was to compare our previous results of variation in recombination rates within and between five domestic pig breeds with measures of intrachromosomal genetic shuffling and investigate whether there is variation between sex and breeds.

Materials & Methods
Data. This study focused on five purebred commercial breeding populations with pedigree and genotype data: two sow breeds, Landrace (LR) and Large White (LW); and three boar breeds, Duroc (DU), Synthetic (SY) and Pietrain (PI). The genotype data and filtering are described in detail in Brekke et al. (2022).
Linkage mapping and crossover detection. Detailed description of linkage mapping, gamete phasing, crossover detection and fine scale recombination mapping can be found in Brekke et al. (2022).

Genetic shuffling within individuals. Genetic shuffling was calculated as the probability that a randomly chosen pair of loci was shuffled during gamete production following the method suggested by Veller et al. (2019). We defined the parameter $\bar{r}$ as the probability of two alleles at a chromosome being shuffled due to recombination, i.e excluding the part due to independent assortment in equation (4) in Veller et al. (2019):

$$\mathbb{E}[\bar{r}] = \sum_{k=1}^{n} 2p_k(1-p_k)\,L_k^2$$

where k is the chromosome number (1-18), p is the proportion of grandpaternal alleles, $1\text{-}p$ is the proportion of grandmaternal alleles and L is the length of the chromosome as a fraction of the total length of the genome. For each phased gamete transmitted from a mother or father to an offspring, shuffling was calculated following equation 1 and assigned as an observation to the parent, resulting in multiple observations for the phenotype "genetic shuffling", hereafter referred to as $\bar{r}$, for each parent (hereafter focal individual, or FID).

Genetic variation. We estimated variance components for individual $\bar{r}$ with a repeatability model using the restricted maximum likelihood (REML) method and average information (AI) algorithm in DMU v 6 (Madsen et al., 2014). The following model was used:

$\bar{r}$=*sex*+*id1*+*id2*+**e**

where **sex** is the fixed effect of sex, **id1** is the random additive genetic effect of the FID, **id2** is the random effect of the FID permanent environment (i.e. individual and/or environmental effects affecting all gametes from an FID) and **e** is the residual effect. The narrow-sense heritability (h^2) was defined as the proportion of phenotypic variance explained by the additive genetic effect and was estimated separately for each breed and sex.

Results

Mean $\bar{r}$ is significantly higher in males than in females in the LR ($p=3.12e^{-105}$), DU ($p=2.51e^{-85}$), PI ($p=3.70e^{-72}$) and SY ($p=2.38e^{-53}$) breed. In the LW line $\bar{r}$ is only slightly, but significantly ($p = 1.40e^{-21}$) higher in females. Means and distribution are plotted in Figure 1. Intrachromosomal shuffling is a heritable trait in females in all breeds, but only in the LR breed in males. These results are presented in Table 1.

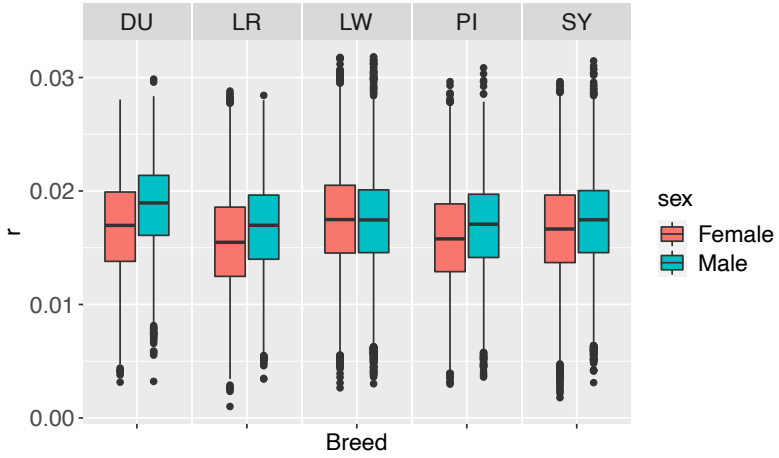


Figure 1. Sex difference in mean and distribution of shuffling.

Table 1. Results from variance component estimation of $\bar{r}$.

Breed	Sex	Mean(SD)	h^2(SE)	No obs[1]	No ind[2]
LR	M	0.0167(0.0039)	0.03(0.01)	11805	155
LR	F	0.0155(0.0043)	0.13(0.02)	11805	1960
DU	M	0.0186(0.0038)	0.03(0.02)	4090	89
DU	F	0.0168(0.0043)	0.09(0.02)	4090	661
LW	M	0.0172(0.0039)	0.04(0.01)	41237	273
LW	F	0.0175(0.0042)	0.15(0.01)	41237	4704
PI	M	0.0168(0.0040)	0.03(0.01)	12159	196
PI	F	0.0158(0.0042)	0.09(0.01)	12159	1355
SY	M	0.0172(0.0039)	0.03(0.01)	25705	224
SY	F	0.0166(0.0043)	0.05(0.01)	25705	2635

[1]Total number of observations (gametes).
[2]Total number of unique males or females with repeated observations.

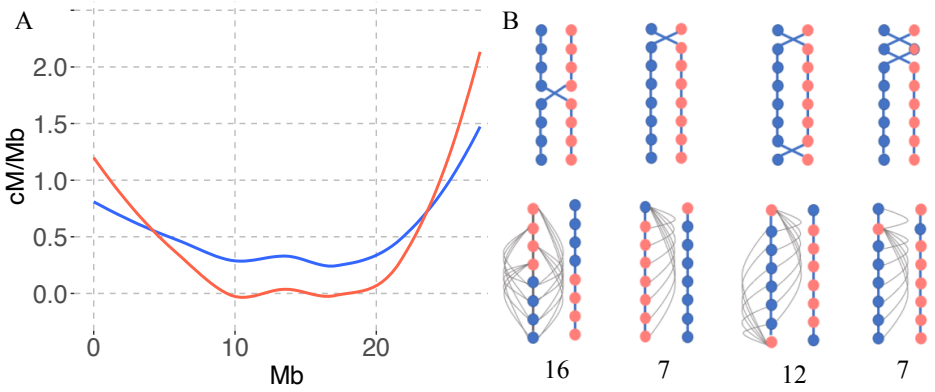


Figure 2. A) Variation in recombination rates along chromosome 1 in the PI breed for males in blue and females in red plotted with the loess method in the geom_smooth function in ggplot2 (Wickham, 2007). B) Example of crossover positions and the resulting number of pairs of loci shuffled (modified from Veller et al. (2019)).

Discussion

Our results show that the probability of shuffling between two loci on the same chromosomes is low relative to the shuffling of alleles at different chromosomes (average of 0.5 due to Mendelian segregation), but also that the intra-chromosomal shuffling differs between breeds and sexes and is a heritable trait in the domestic pig. Genome wide recombination rates in these pig breeds are higher in females than in males, and rates tend to be elevated in the telomeric regions in both sexes (Brekke et al. 2022). In some chromosomes this pattern is more extreme in females, i.e. recombination rates are higher than males in the telomeric regions, but lower than in males closer to the centromere (e.g. as in Figure 2a). This could explain why $\bar{r}$ is lower in females despite higher genome wide recombination rates. Figure 2b illustrates why a central crossover leads to more shuffling and why the position of the crossover may have a higher impact on the probability of shuffling between two loci than the number of crossovers. It is however puzzling that one of the breeds show the opposite sex difference in $\bar{r}$ (Figure 1). This breed (LW) is the breed with the highest overall rate (Brekke et al. 2022). More evenly spread crossovers lead to more shuffling (Figure 2b). Differences in

genetic shuffling can thus also be caused by differences in crossover interference between the breeds, explaining part of the difference in $\bar{r}$. Our results show that even if overall levels of sex-differences in recombination is the same in closely related populations, the shuffling might be different, potentially because of rapidly evolving hotspot usage (Paigen and Petkov, 2010; Weng et al., 2014, 2019). It is not clear, however, how the difference between the sex is maintained in a population from generation to generation as each offspring receives a paternal and maternal gamete. Even if a different number of unique sires and dams mated in each generation, the number of maternal and paternal gametes in each generation is always exactly the same. A next step could be to look at differences in recombination and genetic reshuffling between X and Y paternal gametes. The population level shuffling might be more influenced by variation in $\bar{r}$ within the sex, and in males in particular in pigs as the selection pressure is higher. In conclusion this study shows that variation in crossover distribution affects the production of novel haplotypes from one generation to the next and that there is variation in the shuffling caused by recombination between breeds, sex and individuals in the domestic pig.

References

Arbeithuber, B. *et al.* (2015) PNAS, 112(7), pp. 2109–2114. doi:10.1073/pnas.1416622112.

Burt, A., Bell, G. and Harvey, P.H. (1991), *J. evol. Biol*., pp. 259-277, doi: 10.1046/j.1420-9101.1991.4020259.x

Brekke et al. 2022. Recombination rates vary between breeds, sex and individuals in the domestic pig and is associated with *RNF212*, *SYCP2*, *PRDM7, MEI1* and *MSH4*. [Manuscript submitted for publication]. Norwegian University of Life Sciences.

Charlesworth, B. and Barton, N.H. (1996), Gen. Res., 67(1), pp. 27–41. doi:10.1017/s0016672300033450.

Fledel-Alon, A. *et al.* (2009) *PLoS Gen.*, 5(9). doi:10.1371/journal.pgen.1000658.

Halldorsson, B. v. *et al.* (2019), *Science*, 363(6425). doi:10.1126/science.aau1043.

Hassold, T. *et al.* (1995), *Am. J. Hum. Genet*. 57(4): 867–874. PMCID: PMC1801507

Johnston, S.E. *et al.* (2016), *Gen.*, 203(1), pp. 583–598. doi:10.1534/genetics.115.185553.

Koehler, K.E. *et al.* (1996) Hum. Mol. Gen. 5.1495-1504. doi:10.1093/hmg/5.Supplement_1.1495

Lien, S. *et al.* (2011) *BMC Genomics*, 12. doi:10.1186/1471-2164-12-615.

Madsen, P., Jensen, J. et al. (2014). *Proceedings of the 10th world congress of genetics applied to livestock production* (pp. 18-22).

Paigen, K. and Petkov, P. (2010) *Nat. Rev. Gen.*, pp. 221–233. doi:10.1038/nrg2712.

Team, R. Core.(2013) "A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing." *URL http://www.R-project.org*

Ritz, K.R., Noor, M.A.F. and Singh, N.D. (2017) *Trends in Genetics*. Elsevier Ltd, pp. 364–374. doi:10.1016/j.tig.2017.03.003.

Sakamoto, T. *et al.* (2000) Genetics155:1331–1345, doi: 10.1093/genetics/155.3.1331.

Sherman, S.L. *et al.* (1991) *Am. J. Hum. Genet*. 49(3):608-620, PMCID: PMC1683138.

Stapley, J. *et al.* (2017) *Phil. Trans. Roy. Soc. B.*, 372(1736). doi:10.1098/rstb.2016.0455.

Tortereau, F. *et al.* (2012) *BMC Genomics*, 13(1). doi:10.1186/1471-2164-13-586.

Veller, C., Kleckner, N. and Nowak, M.A. (2019) *Proceedings of the National Academy of Sciences*, 116(5), pp. 1659–1668. doi:10.1073/PNAS.1817482116.

Weng, Z.Q. *et al.* (2014) *Gen. Sel. Evo.*, 46(1). doi:10.1186/1297-9686-46-34.

Weng, Z. *et al.* (2019) *J. An. Sci. Bio.*, 10(1). doi:10.1186/s40104-019-0332-y.

Wickham, H., & Wickham, M. H. (2007). The ggplot package.

4. Results

In this section results that are comparable across species and papers are presented. Detailed information on the results for each species is in the respective papers.

4.1 Difference between species

4.1.1 Genome-wide recombination rate

Total sex averaged recombination rate in cM per Mb is highest in Cattle and lowest in Atlantic salmon. When comparing sex specific recombination rate, female pigs and female Atlantic salmon have the highest rate of 1.02 cM/Mb and Atlantic salmon males have the lowest rate of 0.71 cM/Mb. There is a bigger difference between male rates in the three species than female rates. Sex averaged and sex specific total rates for all three species are presented in table 2.

	cM/Mb		
Species	**Sex average rate**	**Female rate**	**Male rate**
Pig	0.91	1.02	0.80
Cattle	0.96	0.93	1.00
Atlantic salmon	0.87	1.02	0.71

Table 2. cM/Mb is the total genetic length in centiMorgan of the autosome in pigs and cattle and genome in Atlantic salmon divided by the total physical length in Megabases.

4.1.2. Genetic map lengths and number of chromosome arms

Number of chromosome arms has been suggested as a better indicator of genetic length of a genome (Coop and Przeworski, 2007). Presented here is the relationship between genetic length and number of chromosome arms in the three species. Short arms of acrocentric chromosomes are excluded. The cattle autosome consists of 29 acrocentric chromosomes, i.e. 29 chromosome arms (Liu *et al.*, 2009). The pig autosome consists of 18 chromosomes where 12 are metacentric or submetacentric and the rest are acrocentric resulting in 30 chromosome arms (Hansen-Melander *et al.*, 1974). The Atlantic salmon genome does not have sex chromosomes and consist of 29 chromosomes where seven are metacentric or submetacentric and 22 are acrocentric resulting in 36 chromosome arms (Lien *et al.*, 2016). Centimorgan per

chromosome arm is highest in cattle for both sex averaged and sex specific estimates, in both males and female the estimates suggest on average more than 1,5 crossovers per chromosome arm per meiosis. The estimates are lowest for Atlantic salmon, in males an estimate of 44.26 cM signify on average less than one crossover per chromosome arm per meiosis. All estimates are presented in table 3 and plotted in figure 2. The difference in cM per chromosome arm in males is even bigger than the difference in cM/Mb.

		cM/chromosome arm		
Species	**Chromosome arms**	**Sex averaged**	**Male**	**Female**
Cattle	29	82.79	85.96	79.62
Pigs	30	68.82	60.25	77.39
Atlantic salmon	36	54.00	44.26	63.73

Table 3. average genetic length in cM per chromosome arm. Number of chromosome arms is excluding short arms of acrocentric chromosomes.

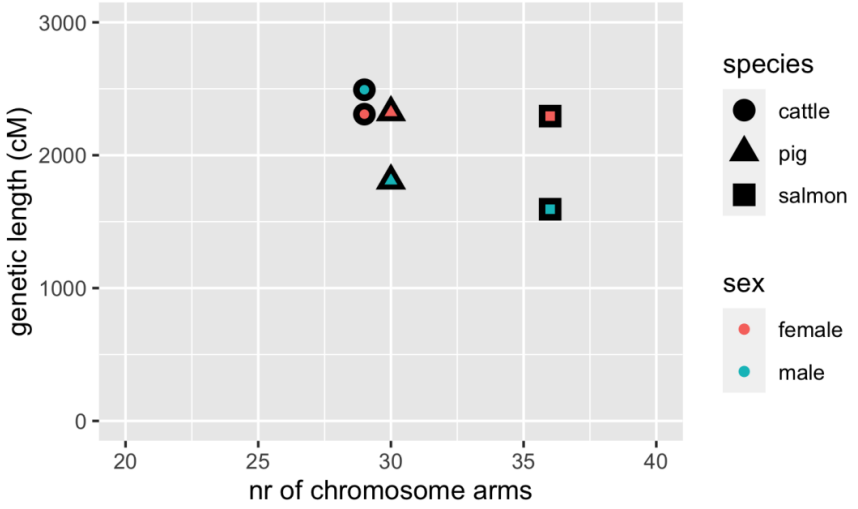


Figure 2. Estimate of sex specific total genetic length plotted against number of chromosome arms for males in blue and females in red. Circles for cattle estimates, triangle for pig estimates (averaged across five breeds) and squares are atlantic salmon estimates.

4.1.3 Relationship between physical length and genetic length of the chromosomes

The relationship between the physical length and genetic length of the autosomes were tested with linear regression. The relationship was close to linear in both male and female cattle (adjusted $R^2 = 0.92$ and 0.97, respectively) (Table 3). In salmon however, the physical length of the chromosomes is not a very good indicator of the genetic length with R2 = 0.5 for

females and 0.33 for males. In the pig, the relationship between physical and genetic length of the autosomes is very different in males and females. In males, the relationship is not far from linear ($R^2 = 0.84$), but in females R^2 for linear fit is close to 0. Adjusted R^2 and p values for linear fit between genetic and physical length of the autosomes is presented in table 3. The relationship is plotted in Figure 3.

Species	Sex	Adjusted R^2	P-value
Pig	Female	0.11	0.1022
	Male	0.84	6.2e-08
Cattle	Female	0.97	2.2e-16
	Male	0.92	2.2e-16
Atlantic salmon	Female	0.50	9.6e-06
	Male	0.33	0.6e-03

Table 3. Adjusted R squared and respective p values for linear fit between genetic and physical length of the autosomes.

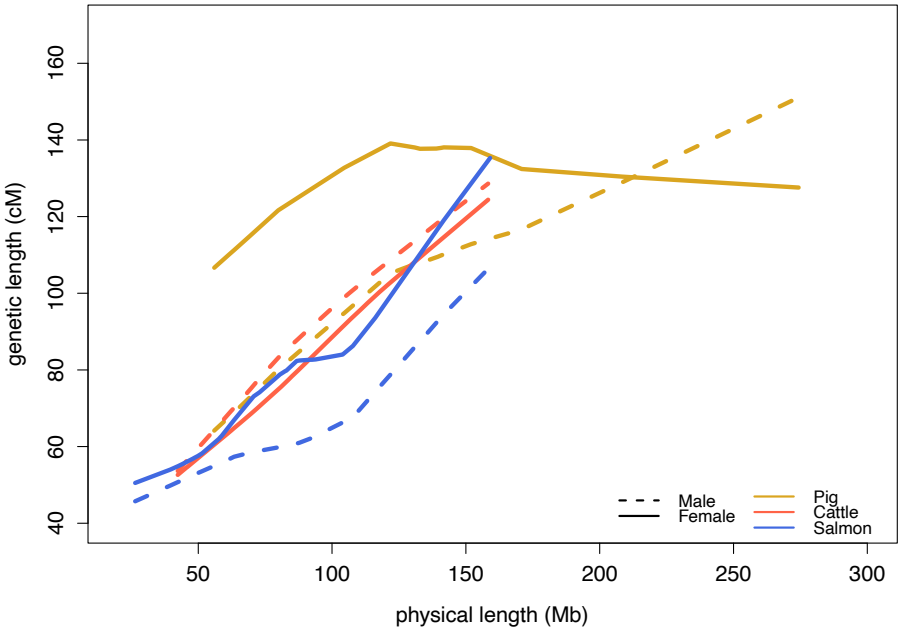


Figure 3. The physical length in Megabases is plotted against genetic length in centiMorgan for each autosome for pigs in yellow (averaged across five breeds), cattle in red and Atlantic salmon in blue. A loess smoother is fitted for solid line for female estimated genetic length and stapled line for male estimates.

4.2 Sex differences

All three species show difference between male and female recombination rates, however to varying degree and direction. In cattle, the genome wide rate is 8% higher in males and although the sex difference is vague it is consistent across all 29 autosomes. In cattle, crossovers are relatively evenly distributed along the genome and almost overlapping in males and females (Figure 4A). In pigs, the genome wide recombination rate is 20-27 % higher in females than in males. However, two of the chromosomes, 1 and 13, show higher rates in males than in females for all five breeds, also in agreement with previous studies in other pig breeds (Tortereau *et al.*, 2012; Johnsson *et al.*, 2021). Both male and female pigs have slightly elevated recombination rates towards the telomeric regions, but this pattern is stronger in females and there is a clear difference between the sex in all five breeds (Figure

4B). In Atlantic salmon total rate is also higher in females, but the striking sex difference in salmon however is the difference in distribution across the genome; in males crossovers are almost exclusively occurring in the sub-telomeric regions, in females the pattern is a bit less extreme, but opposite from males, i.e. rates are elevated closer to the centromere (Figure 4C).

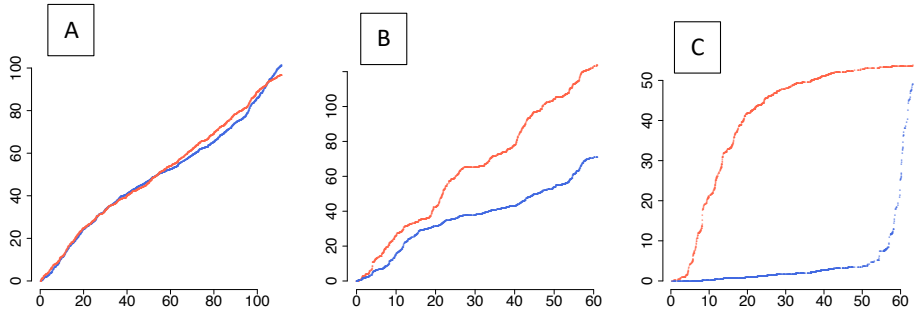


Figure 4. Physical positions in Mb on the x axis and genetic positions in cM on the y axis for all SNP markers on an acrocentric chromosome from A) Cattle (chromosome 8), B) Pig (chromosome 12) and C) Atlantic salmon (chromosome 22). The red line is the female linkage map, and the blue line is the male linkage map. The pig male and female maps is the average maps of the five breeds in paper I.

4.3 Genetic variance and architecture of individual recombination rates

Heritability estimates of autosomal recombination rate is low to moderate for both sex and all three species, ranging from 4% to 12 %. In pigs, the trait is more heritable in females than in males, but in cattle and Atlantic salmon, heritability is higher in males. All estimates are presented in table 5.

	Cattle	Pig	Atlantic Salmon
♀ h^2	4%	8 - 11 % (5 breeds)	11%
♂ h^2	9%	4 – 7 % (5 breeds)	12%

Table 5. estimates of percentage of phenotypic variation explained by genetic variation for autosomal crossover count for males and females in Cattle, pigs and Atlantic salmon. (Detailed information on results are in papers I, II and III.)

We do not pick up signals of association for any of the same genes with individual recombination rates in the three species. In pigs, 11 loci show significant association with the genetic variation, where six loci have clear candidate genes, namely MEI1, PRDM7, MSH4, RNF212, REC8 and SYCP2. In cattle, significant association is found in three loci, with candidate genes also reported in previous studies in cattle, namely NEK9, RNF212b, MLH3 and CEP55. In Atlantic salmon, two loci are significant, but only supported by one marker and without clear candidate genes close by.

5. General discussion

5.1 Data quality

In this study datasets with 17 000- 130 000 individuals in each breed/population with 35K – 50K SNP genotypes was used to study recombination rate variation. All three breeding companies routinely genotype a handful of individuals each year on high density SNP arrays (700-900K), however, to maximize number of FID with genotyped parents, mates and offspring, it was prioritized to have datasets with a high number of individuals that were genotyped rather than a high number of markers. Considering that crossovers are usually spaced apart due to crossover interference(Otto and Payseur, 2019), the marker density is expected to be sufficient. However, it is possible that recombination events occurring very close to one end of a chromosome are missed, particularly in Atlantic salmon, where males recombine almost exclusively in the sub-telomeric region and because of high similarity between homeologues, some telomeric regions are hard to map and consequently marker density is low (Lien *et al.*, 2016). (See marker density plot in supplementary S1 in paper III.).

Due to different breeding practices as well as biological differences in litter size, the family structures in the three different species are different. Most of the cattle full-sib families had only one offspring, which may make phasing of the FIDs less accurate. However, in the cattle dataset, only families with genotyped grandparents (i.e. parents of the FIDs) were included. Male cattle FIDs still had a high number of repeated measures from different families, but most females only had one or two observations. In pigs, the number of offspring was between 1 and 27 in a full-sib family, suggesting that there is sufficient information in most families for correct phasing of the FIDs. For analysis of genetic variation, the number of unique FID males was substantially lower than for females in both cattle and pigs, because selection is stronger in males. The power for estimation of genetic variation and detection of regions associated with the trait is therefore better in females in the pig and cattle datasets. Which may be why we almost exclusively pick up significant loci in the association analysis in females. For salmon the number of unique males and females was similar, but lower than number of cattle and pig females (Table 1).

Pedigree and genotype data have been used in many similar studies to this one (Ma *et al.*, 2015; Johnston *et al.*, 2016, 2017; Petit *et al.*, 2017; Johnsson *et al.*, 2021) and the results

were therefore comparable with a lot of other species. However, a problem with this method is that observations are biased towards successful gametes, i.e. gametes that turn into viable offspring. A great addition would be to study oocytes and spermatocytes from the same populations (this topic is further discussed in section 5.4).

5.2 Software used

The software Lepmap3 (Rastas, 2017) was used to construct the linkage maps and infer individual recombination events. This software allowed for inclusion of pedigree and genotype information and estimate sex specific linkage maps in addition to individual recombination events. One potential downside with Lepmap3 is that it does not take half sib information into account. Meaning that if an FID only has one offspring with each mate, genotypes from all halfsibs is not utilized in the phasing of the FID. Other linkage map softwares like for example CRIMAP (Green, Crooks and Falls K., 1990) have that option. However, in the salmon families, number of offspring were most likely more than high enough for proper phasing of the FIDs and in pigs and cattle, the FID always had genotyped parents. In the different analyses, recombination rates were estimated based on a given order of the markers. Lepmap3 also has the option to estimate the order of the markers within a linkage group, this was tested but gave a poorer result (inflations in the map), suggesting that the physical mapping of the markers in all three datasets was already good.

GCTA (Yang *et al.*, 2011) was used for genome wide associations between marker alleles and trait variation. A restriction with GCTA is that it is not able to take repeated observations, for each individual an average across repeated observation was used in the analysis. This will remove information about variation within individual, and it would be of interest to test the same analysis again with a repeatability model.

5.3 Recombination and genetic shuffling

Availability of data from several breeds, three different kinds of species, and two different taxa, allowed for direct comparison of the distribution and magnitude of variation in recombination rates both between closely related populations and very distantly related populations in an evolutionary perspective. In this section, the discussion focuses on variation

across the three species. Discussion of all results for each species and analysis is in the respective papers. Paper I. also discuss the variation between the different pig breeds in more detail.

5.3.1 Variation between species

There is a difference between the three species in sex averaged genome wide recombination rate (total genetic length in cM/total physical length in Mb). The difference is 0.09 cM/Mb between the lowest (Atlantic salmon) and the highest (Cattle) estimate. When comparing sex spesific estimates there is a bigger difference between male rates than female rates. The difference between the lowest and highest male rate is 3 times higher than the difference between the lowest and highest female rate (Table 2). When comparing recombination rates across species, it has been suggested that number of chromosome arms is a better indicator than physical length in Mb (Coop and Przeworski, 2007). When looking at cM per arm, the difference between these three species is higher than in cM/Mb. Cattle have the highest rate per arm for both sex-averaged and sex specific estimates. Also with this measure the difference between males is higher than the difference between females. Cattle male cM per chromosome arm is almost double than male Atlantic salmon (Table 4). It should be noted that the male salmon rate might be slightly underestimated (see discussion paper III.) When comparing these three species, the driver of variation in direction and magnitude of heterochiasmy between the species is the male rates, weather this is just by chance or a potential general pattern of heterochiasmy would have to validated by looking at more species. However, variation in ACC within females is higher than within males in all three species (See paper I., II. And III.).

The relationship between physical length and genetic length of the chromosomes is linear in Cattle, but in Atlantic salmon and pig, physical length of the chromosomes is a poorer indicator of genetic length. In female pigs R^2 for linear fit is close to 0 (Table 3 and Figure 2). This is mainly driven by chromosome 1 and 13, that are a notably larger chromosome than the other chromosomes and much larger than any chromosome in cattle and salmon. Other studies also find that recombination rates are often lower in larger chromosomes ((Johnston *et al.*, 2017). Also, chromosomes 1 and 13 are the only two chromosomes where male recombination rate is higher than female recombination rate in pigs (See paper I.). A possible explanation why physical length better explain genetic length in cattle may be that all

autosomes are acrocentric, whereas pigs and Atlantic salmon both have several metacentric autosomes, resulting in a higher number of chromosome arms.

5.3.2 Difference in magnitude and direction of heterochiasmy

A general observation from previous studies in recombination rate variation is that biological sex is the main factor of variation in recombination rates within species (Dunn and Bennett, 1967; Burt, Bell and Harvey, 1991; Barton and Charlesworth, 1998; Lenormand and Dutheil, 2005a; Mank, 2009; Halldorsson *et al.*, 2016; Johnston *et al.*, 2017). This seems to hold in these studies as well, but the level and direction of heterochiasmy is different between the three species. Male maps are longer in cattle and female maps are longer in pigs and Atlantic salmon. The male to female ratio of total genetic length is 1:1.3, 1:0.9 and 1:1.4 in pigs, cattle and Atlantic salmon respectively. In mammals, female biased heterochiasmy is most common (Halldorsson *et al.*, 2016; Johnston *et al.*, 2017; Johnsson *et al.*, 2021). Why recombination rates vary between sex has been a topic of high interest, and one of the suggestions that have received more attention the last decades is the difference in opportunities for selection at the haploid stage (Lenormand and Dutheil, 2005). In eukaryotic systems female meiosis experience meiotic arrest and meiosis is not completed until fertilization (Lenormand *et al.*, 2016), hence the combinations of alleles are not expressed in the haploid stage as in spermatocytes. However, in human there is some evidence of selection on recombination rate during female meiosis where elevated rates are observed in the oocytes compared to the polar bodies (Ottolini *et al.*, 2015). If this occurs in more eukaryotic systems, selection on rates of recombination in female meiosis might be for mechanistic purposes to ensure proper segregation and production of a gamete that can result in viable offspring, whereas in males, selection affecting recombination in the haploid stage could be more affected by the actual allele combinations in the gamete (Parker, 1990). Following on this speculation, the direction and magnitude of heterochiasmy in a population would be driven by male rates, and females would have a shared optimum that is relatively stable across species. This would fit well with the observation in this thesis that there is a bigger difference between male rates, and female rates are more stable across species (Table 1 and 2).

More striking than the difference in Genome-wide rate is the clear difference between species in the distribution of crossovers. The difference in recombination patterns is most extreme in Atlantic salmon, where male rates are highly elevated in telomeric regions and female rates are elevated in centromeric regions (Figure 2 and 3 paper III.) Atlantic salmon is also the

species in this study with the biggest difference between male and female meiosis with the partial tetrasomic inheritance in male salmonids (Allendorf *et al.*, 2015). In pigs the variation in overall rates and distribution of crossovers is less consistent, two autosomes exhibit higher crossover rates in male and the difference in distribution along the chromosomes between males and females is fluctuating (Figure 4, paper I.). Interestingely though, the patterns are very well conserved across breeds (Figure 2 and 3, paper I.). The cattle sex specific distributions of crossovers are overlapping and comparably more evenly distributed along the genome than in the other species. Both domestic cattle and pigs are under strong selection and both species were domesticated around 10 000 years ago (Götherström et al., 2005). Hence, attempts to connect the observed patterns to patterns of selection sweeps becomes hard to justify when comparing to the other species. In fact, some recent studies find little correlation between patterns of recombination and signatures of selection during speciation (Turbek *et al.*, 2021). The extreme pattern in Atlantic salmon suggests that patterns of recombination rates may be more connected to genome evolution as salmonids have undergone a relatively recent WGD event and studies suggests that stable diploidy is not yet full retained (Timusk *et al.*, 2011; Allendorf *et al.*, 2015; May and Delany, 2015; Lien *et al.*, 2016).

5.3.3 Genetic basis of variation in recombination

The phenotypic variation in recombination rate explained by genetic variation is low to moderate in all three species (4-12%) and most of the variation remains unexplained. This is in agreement with many other studies that have used the same approach to study heritability for recombination rates, i.e. genotype data on individuals in a population with known pedigree) (Ma *et al.*, 2015; Kadri *et al.*, 2016; Petit *et al.*, 2017; Johnston, Huisman and Pemberton, 2018; Johnston, Stoffel and Pemberton, 2020; Johnsson *et al.*, 2021). A shared bias in these studies is that we are only sampling gametes from livebirths. There is great evidence in human that a large percentage of miscarriages are due to aneuploidy, which is usually caused by a lack of recombination between one of the homologues chromosome pairs (Sherman *et al.*, 1991; Hassold *et al.*, 1995). Studies in more products of the meiosis in more species is needed to determine if the genetics of an individual has a greater influence on recombination rate or not. Sandor et al (2012), studied genome wide recombination rate in

sperm in Holstein Frisian bulls and estimated heritability (h^2) at 0.22, which is substantially higher than 0.09 in bulls in this study. There is no effect of age in any of our studies (could only be tested in pigs and cattle). The age affects observed in human (Halldorsson *et al.*, 2016) may not be relevant in either of our species, as they do not reproduce at such high age in the breeding programs.

We do not pick up any common genes affecting recombination in cattle, pigs and Atlantic salmon. In pigs however, we pick up association at two genes that are also associated with recombination rate variation cattle in a different study (Sandor *et al.*, 2012), namely REC8 and RNF212. These two genes have also been associated with the trait in several other species, like human (Fledel-Alon *et al.*, 2011; Halldorsson *et al.*, 2019), red deer (Johnston, Huisman and Pemberton, 2018) and sheep (Petit *et al.*, 2017; Johnston, Stoffel and Pemberton, 2020). Alleles at these loci may not be segregating in our population, or we don't have enough power to pick up the association in our dataset. In pigs we pick up an association in a region close to a gene annotated as PRDM7, which could potentially be PRDM9, (see discussion in paper I.) PRDM9 is usually a gene associated with hotspot usage (Paigen and Petkov, 2018), but in cattle, PRDM9 has also been associated with variation in genome-wide rates of recombination (Ma *et al.*, 2015; Shen *et al.*, 2018). We did not investigate hotspot usage in these studies, which would be an interesting next step that would require higher density genotypes. Also, more investigation into the genetic control of recombination rates in salmon with imputed or higher density data and/or with repeated observations is required to further scrutinize common genetic architecture between the species. (See discussion in paper III).

5.4. perspectives

It is established that recombination rates vary between and within species and sex and along the genome and that it is a heritable trait, but how is this relevant for breeding?

5.4.1. Potential to increase genetic variation

The genetic variation in these breeding populations as well as in most other natural and domestic populations studied, suggests that there is a potential to increase recombination rates to increase genetic variation. However, studies have shown that to achieve significant increase in response to selection, recombination rates would have to be increased 10-20-fold (Battagin et al., 2016). The consequences of increasing recombination rates to that level is not known, but there is compelling evidence that substantial increase in recombination rates is not beneficial. In humans, high levels of recombination have been associated with cancer (Mao *et al.*, 2009) and there is evidence of increased mutation rates in recombination hotspots (Arbeithuber *et al.*, 2015; Halldorsson *et al.*, 2019). Because recombination can also break up favorable linkage previously built up by selection, the benefits are likely to be outweighed by the negative consequences.

5.4.2 Fertility

The genetic variation in recombination rates also suggest that there is a potential to decrease population level rates of recombination. One important aspect in breeding is avoiding indirect selection on a trait with a negative correlation with a trait of interest. A lack of recombination between two homologue chromosomes during meiosis is often detrimental as it leads to non-disjunction and aneuploidy in the resulting gamete (Sherman *et al.*, 1991; Hassold *et al.*, 1995). This suggests that there might be an association between low recombination rates and reduced fertility. The ability to produce offspring is fundamental to the success of both natural and domestic populations, and fertility is thus an important trait in all breeding programs. There has been some evidence of a correlation between reduced fertility and reduced recombination rate (Ma *et al.*, 2015) which would be of interest to investigate in these species. For one of the pig breeds there were breeding values available for fertility traits. A simple correlation test was performed between individual ACC and breeding values for these traits, failing to find a relationship (S1). However, there may be a potential to compare results from this study with other data, for example from a recent study in sperm quality in Norwegian Red cattle (Olsen, Heringstad and Klemetsdal, 2020). However, to

avoid the potential bias in studying recombination in successful gametes, it would be of interest to study crossovers in sperm and egg cells to further investigate the association with fertility. If there is an association, it would be important to avoid indirectly selecting for reduced recombination in a breeding population.

5.4.3 Genetic shuffling

Most of the shuffling of alleles in a genome from one generation to the next is caused by independent assortment of chromosomes. Indeed, when calculating probability of shuffling between any pair of loci in the data in this study, estimates are close to 0.5 (see figure S2), in agreement with Veller et al (2019). The estimates of genetic shuffling caused by recombination in paper III. And IV. are very low in comparison, however, it can be argued that it is the intrachromosomal shuffling that should be of interest in a breeding population, because alleles on different chromosomes have a 50% chance of being reshuffled in every generation, but if a reshuffling of linked alleles leads to a beneficial novel haplotype it may percist and be passed on in future generations.

Individual intrachromosomal shuffling was only analyzed in Atlantic salmon (paper III.) and pigs (paper IV.). The intrachromosomal shuffling is much higher in males than in females in Atlantic salmon, which is caused by the marked difference in distribution of crossovers. In pigs, it is less evident from looking at the crossover distribution that intrachromosomal shuffling is higher in males. The sex difference is much higher in Atlantic salmon than in pigs, but the results in pig demonstrate that genetic shuffling is not necessarily higher in the sex with the highest genome wide recombination rate. An interesting observation is that having terminal crossovers is a more effective way of reducing shuffling than reducing the overall rate.

It is unclear what the consequences are of the sex difference in genetic shuffling in a breeding population. Every generation, each offspring inherits one maternal gamete and one paternal gamete. New gametes are produced in this individual where rates and distribution of crossovers is affected by the sex of the focal individual and not which sex the chromatid was inherited from in the previous generation. This may suggest that the population level shuffling of alleles should be equally affected by male and female recombination patterns.

However, variation in genetic shuffling within sex may have a higher impact on the population level shuffling if one of the sexes is under strong selection.

5.5 Suggestions for future studies

In this study, most of the phenotypic variation in recombination rates is explained by the error variance. Future studies should analyze a broader sample of gametes, for example from sperm and egg cells, to try to further understand the genetic control of recombination rates.

Further studies in the datasets already at hand should aim to define a crossover position trait to study the genetic basis of the variation in recombination landscapes. In this study we find that the variation in individual measures of genetic shuffling has a heritable component which may reflect genetic variation in crossover positioning.

Another point of interest would be to look at how directional selection may affect recombination rates and distribution as well as direction and magnitude of heterochiasmy in real or simulated data.

6. Conclusions

In this study, we provide first estimates of heritability of individual crossover rates and distribution in Atlantic salmon. In agreement with previous studies in pigs and cattle, recombination rates are also heritable in the Norwegian red cattle breed and five pig breeds: Norwegian landrace, Norwegian Duroc, Dutch Large White, Piétrain and a synthetic mixed breed. In pigs and cattle, we detect association with genes that have been identified to affect recombination rates in other breeds as well as in other species, suggesting some conserved genetic control of this trait across species. However, we also detect two novel associations in pig and fail to detect associations with some loci detected in other cattle breeds. Only two out of five genes affect the trait in all pig breeds, and there is a difference in overall rate between the breeds consistent across chromosomes, suggesting that there is some variation in genetic architecture of recombination rates between closely related breeds and that overall rates can also evolve relatively rapidly. Most of the phenotypic variation in rates of recombination observed remains to be explained. In Atlantic salmon, regions without recombination in males coincide with regions of high sequence similarity between homeologs known to act differently in male and female meiosis in salmonids, suggesting that differences in the meiotic process between sex can play an important role in the marked sexual dimorphism in genome-wide rates and distribution of recombination observed in many species. The sex specific patterns of recombination lead to marked differences in the amount of intrachromosomal shuffling of alleles in maternal and paternal gametes, however it is not clear how this affects the overall genetic shuffling in a breeding population across generations. Overall, the findings in this study contribute to the understanding of the genetic mechanisms underlying recombination rate variation and how this affects genetic shuffling in a breeding population as well as in natural populations.

7. References

Allendorf, F.W. *et al.* (2015) “Effects of crossovers between homeologs on inheritance and population genomics in polyploid-derived salmonid fishes,” *Journal of Heredity*. Oxford University Press, pp. 217–227. doi:10.1093/jhered/esv015.

Arbeithuber, B. *et al.* (2015) “Crossovers are associated with mutation and biased gene conversion at recombination hotspots,” *Proceedings of the National Academy of Sciences of the United States of America*, 112(7), pp. 2109–2114. doi:10.1073/pnas.1416622112.

Barton, N.H. and Charlesworth, B. (1998) *Why Sex and Recombination?* doi:10.1126/science.281.5385.1986.

Battagin, M. *et al.* (2016) “Effect of manipulating recombination rates on response to selection in livestock breeding programs,” *Genetics Selection Evolution*, 48(1). doi:10.1186/s12711-016-0221-1.

Baudat, F. *et al.* (2010) “PRDM9 Is a Major Determinant of Meiotic Recombination Hotspots in Humans and Mice,” *Science*, 327(5967), pp. 836–840. doi:10.1126/science.1183439.

Booker, T.R., Ness, R.W. and Keightley, P.D. (2017) “The recombination landscape in wild house mice inferred using population genomic data,” *Genetics*, 207(1), pp. 297–309. doi:10.1534/genetics.117.300063.

Brooks’ And, L.D. and Marks, R.W. (1986) *THE ORGANIZATION OF GENETIC VARIATION FOR RECOMBINATION IN DROSOPHILA MELANOGASTER*.

Burt, A., Bell, G. and Harvey, P.H. (1991) *Sex differences in recombination*, *J. evol. Biol*. doi:https://doi.org/10.1046/j.1420-9101.1991.4020259.x.

Coop, G. and Przeworski, M. (2007) “An evolutionary view of human recombination,” *Nature Reviews Genetics*, pp. 23–34. doi:10.1038/nrg1947.

Dunn, L.C. and Bennett, D. (1967) “Sex differences in recombination of linked genes in animals,” *Genetical Research*, 9(2), pp. 211–220. doi:10.1017/S0016672300010491.

Fledel-Alon, A. *et al.* (2009) “Broad-scale recombination patterns underlying proper disjunction in humans,” *PLoS Genetics*, 5(9). doi:10.1371/journal.pgen.1000658.

Fledel-Alon, A. *et al.* (2011) “Variation in human recombination rates and its genetic determinants,” *PLoS ONE*, 6(6). doi:10.1371/journal.pone.0020321.

Götherström, A. *et al.* (2005) “Cattle domestication in the Near East was followed by hybridization with aurochs bulls in Europe,” *Proceedings of the Royal Society B: Biological Sciences*, 272(1579), pp. 2345–2350. doi:10.1098/rspb.2005.3243.

Green, P., Crooks, S. and Falls K. (1990) “Documentation for CRIMAP, Version 2.4.” Washington University School of Medicine,St. Louis, MO.

Halldorsson, B. v. *et al.* (2016) “The rate of meiotic gene conversion varies by sex and age,” *Nature Genetics*, 48(11), pp. 1377–1384. doi:10.1038/ng.3669.

Halldorsson, B. v. *et al.* (2019) “Human genetics: Characterizing mutagenic effects of recombination through a sequence-level genetic map,” *Science*, 363(6425). doi:10.1126/science.aau1043.

Hansen-Melander, Eva *et al.* (1974) “The kar yo type of the pig,” *Hereditas*, 77, pp. 149–158.

Hassold, T. *et al.* (1995) *Recombination and Maternal Age-Dependent Nondisjunction: Molecular Studies of Trisomy 16*, *Am. J. Hum. Genet*.

Hunter, C.M. *et al.* (2016) “The Genetic Architecture of Natural Variation in Recombination Rate in Drosophila melanogaster,” *PLoS Genetics*, 12(4). doi:10.1371/journal.pgen.1005951.

Isberg, S.R. *et al.* (2006) “First evidence of higher female recombination in a species with temperature-dependent sex determination: The saltwater crocodile,” *Journal of Heredity*, 97(6), pp. 599–602. doi:10.1093/jhered/esl035.

Johnsson, M. *et al.* (2021) “Genetic variation in recombination rate in the pig,” *Genetics Selection Evolution*, 53(1). doi:10.1186/s12711-021-00643-0.

Johnston, S., Stoffel, M. and Pemberton, J. (2020) "Variants at RNF212 and RNF212B are associated with recombination rate variation in Soay sheep (Ovis aries)." doi:10.1101/2020.07.26.217802.
Johnston, S.E. *et al.* (2016) "Conserved genetic architecture underlying individual recombination rate variation in a wild population of soay sheep (Ovis aries)," *Genetics*, 203(1), pp. 583–598. doi:10.1534/genetics.115.185553.
Johnston, S.E. *et al.* (2017) "A high-density linkage map reveals sexual dimorphism in recombination landscapes in red deer (Cervus elaphus)," *G3: Genes, Genomes, Genetics*, 7(8), pp. 2859–2870. doi:10.1534/g3.117.044198.
Johnston, S.E., Huisman, J. and Pemberton, J.M. (2018) "A genomic region containing REC8 and RNF212B is associated with individual recombination rate variation in a wild population of red deer (Cervus elaphus)," *G3: Genes, Genomes, Genetics*, 8(7), pp. 2265–2276. doi:10.1534/g3.118.200063.
Kadri, N.K. *et al.* (2016) "Coding and noncoding variants in HFM1, MLH3, MSH4, MSH5, RNF212, and RNF212B affect recombination rate in cattle," *Genome Research*, 26(10), pp. 1323–1332. doi:10.1101/gr.204214.116.
Kawakami, T. *et al.* (2019) "Substantial heritable variation in recombination rate on multiple scales in honeybees and bumblebees," *Genetics*, 212(4), pp. 1101–1119. doi:10.1534/genetics.119.302008.
Koehler, K.E. *et al.* (1996) *Recombination and nondisjunction in humans and flies*, *Human Molecular Genetics*. Oxford University Press. Available at: https://academic.oup.com/hmg/article/5/Supplement_1/1495/663563.
Kong, A. *et al.* (2014) "Common and low-frequency variants associated with genome-wide recombination rate," *Nature Genetics*, 46(1), pp. 11–16. doi:10.1038/ng.2833.
Lenormand, T. *et al.* (2016) "Evolutionary mysteries in meiosis," *Philosophical Transactions of the Royal Society B: Biological Sciences*. Royal Society of London. doi:10.1098/rstb.2016.0001.
Lenormand, T. and Dutheil, J. (2005a) "Recombination difference between sexes: A role for haploid selection," in *PLoS Biology*, pp. 0396–0403. doi:10.1371/journal.pbio.0030063.
Lenormand, T. and Dutheil, J. (2005b) "Recombination difference between sexes: A role for haploid selection," in *PLoS Biology*, pp. 0396–0403. doi:10.1371/journal.pbio.0030063.
Lien, S. *et al.* (2011) "A dense SNP-based linkage map for Atlantic salmon (Salmo salar) reveals extended chromosome homeologies and striking differences in sex-specific recombination patterns," *BMC Genomics*, 12. doi:10.1186/1471-2164-12-615.
Lien, S. *et al.* (2016) "The Atlantic salmon genome provides insights into rediploidization," *Nature*, 533, pp. 200–205. doi:10.1038/nature17164.
Liu, Y. *et al.* (2009) "Bos taurus genome assembly," *BMC Genomics*, 10. doi:10.1186/1471-2164-10-180.
Ma, L. *et al.* (2015) "Cattle Sex-Specific Recombination and Genetic Control from a Large Pedigree Analysis," *PLoS Genetics*, 11(11). doi:10.1371/journal.pgen.1005387.
Madsen, P. *et al.* (2014) "DMU - A Package for Analyzing Multivariate Mixed Models in quantitative Genetics and Genomics," in *Journal of Applied Statistics*. Routledge, pp. 1286–1306.
Mank, J.E. (2009) "The evolution of heterochiasmy: The role of sexual selection and sperm competition in determining sex-specific recombination rates in eutherian mammals," *Genetics Research*, 91(5), pp. 355–363. doi:10.1017/S0016672309990255.
Mao, Z. *et al.* (2009) "DNA repair by homologous recombination, but not by nonhomologous end joining, is elevated in breast cancer cells," *Neoplasia*, 11(7), pp. 683–691. doi:10.1593/neo.09312.

Martinez-Perez, E. and Colaiácovo, M.P. (2009) "Distribution of meiotic recombination events: talking to your neighbors," *Current Opinion in Genetics and Development*, pp. 105–112. doi:10.1016/j.gde.2009.02.005.
May, B. and Delany, M.E. (2015) "Meiotic models to explain classical linkage, pseudolinkage, and chromosomal pairing in tetraploid derivative salmonid genomes: II. Wright is still right," *Journal of Heredity*. Oxford University Press, pp. 762–766. doi:10.1093/jhered/esv056.
Muñoz-Fuentes, V., Rienzo, A. and Vilà, C. (2011) "Prdm9, a major determinant of meiotic recombination hotspots, is not functional in dogs and their wild relatives, wolves and coyotes," *PLoS ONE*, 6(11). doi:10.1371/journal.pone.0025498.
Olsen, H.B., Heringstad, B. and Klemetsdal, G. (2020) "Genetic analysis of semen characteristic traits in young Norwegian Red bulls," *Journal of Dairy Science*, 103(1), pp. 545–555. doi:10.3168/jds.2019-17291.
Otto, S.P. and Payseur, B.A. (2019) "Annual Review of Genetics Crossover Interference: Shedding Light on the Evolution of Recombination." doi:10.1146/annurev-genet-040119.
Ottolini, C.S. *et al.* (2015) "Genome-wide maps of recombination and chromosome segregation in human oocytes and embryos show selection for maternal recombination rates," *Nature Genetics*, 47(7), pp. 727–735. doi:10.1038/ng.3306.
Paigen, K. and Petkov, P.M. (2018) "PRDM9 and Its Role in Genetic Recombination," *Trends in Genetics*. Elsevier Ltd, pp. 291–300. doi:10.1016/j.tig.2017.12.017.
Parker, G.A. (1990) "Sperm competition games: raffles and roles," *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 242(1304). doi:10.1098/rspb.1990.0114.
Petit, M. *et al.* (2017) "Variation in recombination rate and its genetic determinism in sheep populations," *Genetics*, 207(2), pp. 767–784. doi:10.1534/genetics.117.300123.
R Core Team (2020) "R: A Language and Environment for Statistical Computing." Vienna.
Ramey, C. (2011) *The Bourne-Again SHell*.
Rastas, P. (2017) "Lep-MAP3: Robust linkage mapping even for low-coverage whole genome sequencing data," *Bioinformatics*, 33(23), pp. 3726–3732. doi:10.1093/bioinformatics/btx494.
Ritz, K.R., Noor, M.A.F. and Singh, N.D. (2017) "Variation in Recombination Rate: Adaptive or Not?," *Trends in Genetics*. Elsevier Ltd, pp. 364–374. doi:10.1016/j.tig.2017.03.003.
Sakamoto, T. *et al.* (2000) *A Microsatellite Linkage Map of Rainbow Trout (Oncorhynchus mykiss) Characterized by Large Sex-Specific Differences in Recombination Rates*. doi:10.1093/genetics/155.3.1331.
Sandor, C. *et al.* (2012) "Genetic variants in REC8, RNF212, and PRDM9 influence male recombination in cattle," *PLoS Genetics*, 8(7). doi:10.1371/journal.pgen.1002854.
Shen, B. *et al.* (2018) "Characterization of recombination features and the genetic basis in multiple cattle breeds," *BMC Genomics*, 19(1). doi:10.1186/s12864-018-4705-y.
Sherman, S.L. *et al.* (1991) *Trisomy 21: Association between Reduced Recombination and Nondisjunction*, *Am. J. Hum. Genet*.
Smukowski, C.S. and Noor, M.A.F. (2011) "Recombination rate variation in closely related species," *Heredity*, pp. 496–508. doi:10.1038/hdy.2011.44.
Smukowski Heil, C.S. *et al.* (2015) "Recombining without Hotspots: A Comprehensive Evolutionary Portrait of Recombination in Two Closely Related Species of Drosophila," *Genome Biol. Evol.*, 7(10), pp. 2829–2842. doi:10.5061/dryad.sh5cs.
Stapley, J. *et al.* (2017) "Variation in recombination frequency and distribution across eukaryotes: Patterns and processes," *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1736). doi:10.1098/rstb.2016.0455.

Timusk, E.R. *et al.* (2011) “Genome evolution in the fish family salmonidae: Generation of a brook charr genetic map and comparisons among charrs (Arctic charr and brook charr) with rainbow trout,” *BMC Genetics*, 12. doi:10.1186/1471-2156-12-68.
Tortereau, F. *et al.* (2012) “A high density recombination map of the pig reveals a correlation between sex-specific recombination and GC content,” *BMC Genomics*, 13(1). doi:10.1186/1471-2164-13-586.
Turbek, S.P. *et al.* (2021) “Variable Signatures of Selection despite Conserved Recombination Landscapes Early in Speciation,” *Journal of Heredity*, 112(6), pp. 485–496. doi:10.1093/jhered/esab054.
Veller, C., Kleckner, N. and Nowak, M.A. (2019) “A rigorous measure of genome-wide genetic shuffling that takes into account crossover positions and Mendel’s second law,” *Proceedings of the National Academy of Sciences*, 116(5), pp. 1659–1668. doi:10.1073/PNAS.1817482116.
Wang, R.J. *et al.* (2017) “Recombination rate variation in mice from an isolated island,” *Molecular Ecology*, 26(2), pp. 457–470. doi:10.1111/mec.13932.
Wickham H (2016) “ggplot2: Elegant Graphics for Data Analysis .” New York: Springer-Verlag.
Yang, J. *et al.* (2011) “GCTA: A tool for genome-wide complex trait analysis,” *American Journal of Human Genetics*, 88(1), pp. 76–82. doi:10.1016/j.ajhg.2010.11.011.
Zelkowski, M. *et al.* (2019) “Diversity and Determinants of Meiotic Recombination Landscapes,” *Trends in Genetics*. Elsevier Ltd, pp. 359–370. doi:10.1016/j.tig.2019.02.002.

Supplementary

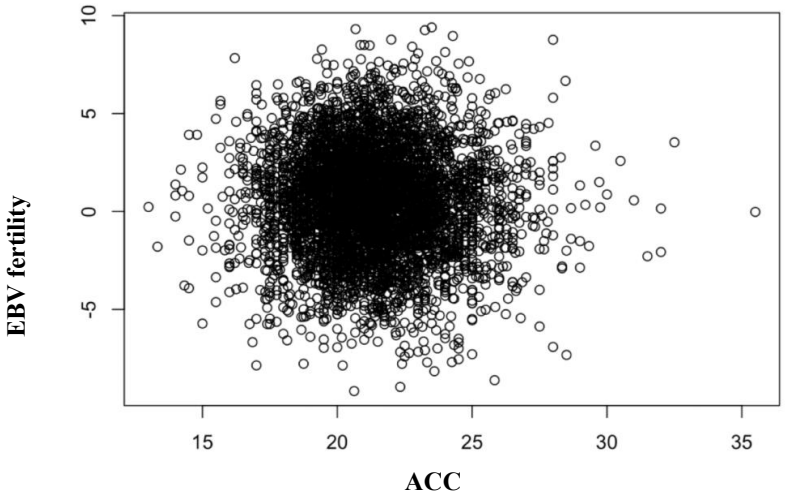


Figure S1. Scatterplot with EBV for fertility on the Y axis and crossover count on the x-axis for all female pigs.

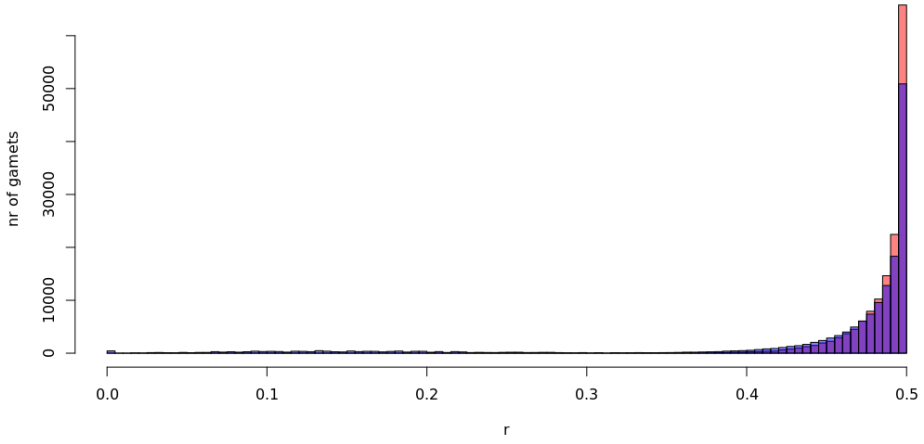


Figure S2. Distribution of total genetic shuffling in maternal gametes in red and paternal gametes in blue. Total genetic shuffling is the intrachromosomal and interchromosomal shuffling combined.

ISBN: 978-82-575-1892-9
ISSN: 1894-6402

Norwegian University
of Life Sciences

Postboks 5003
NO-1432 Ås, Norway
+47 67 23 00 00
www.nmbu.no