

**The recombination rate variation and
genomic distribution in plants using
barley as a model: assessment,
associated genomic features, and
manipulation.**

Inaugural dissertation

for the attainment of the title of doctor
in the Faculty of Mathematics and Natural Sciences
at the Heinrich Heine University Düsseldorf

presented by

Federico A. Casale

from Buenos Aires, Argentina

Düsseldorf, August 2024

From the institute for Quantitative Genetics
and Genomics of Plants
at the Heinrich Heine University Düsseldorf

Published by the permission of the
Faculty of Mathematics and Natural Sciences at
Heinrich Heine University Düsseldorf

Contributors:

1. Prof. Dr. Benjamin Stich
2. Prof. Dr. Korbinian Schneeberger

Date of the oral examination:

Declaration of the Doctoral Dissertation

I herewith declare under oath that this dissertation was the result of my own work without any unauthorized help in compliance with the “Principles for the Safeguarding of Good Scientific Practice at Heinrich Heine University Düsseldorf”. This dissertation has never been submitted in this or similar format to any other institution. I have not previously failed a doctoral examination procedure. The experimental work for this dissertation has been conducted at Heinrich Heine University Düsseldorf.

Düsseldorf, 09.08.2024

A handwritten signature in black ink, consisting of several overlapping loops and strokes, positioned above a horizontal line.

Federico A. Casale

Contents

1	Summary	1
2	General Introduction	3
3	Genomic prediction of the recombination rate variation in barley – A route to highly recombinogenic genotypes ¹	86
4	The role of methylation and structural variants in shaping the recombination landscape of barley ²	120
5	Accurate recombination estimation from pooled genotyping and sequencing: a case study on barley ³	204
7	List of publications	226
8	Acknowledgements	227

¹ **Casale, F. A.**, van Inghelandt, D., Weisweiler, M., Li J., and Stich, B. (2022). Genomic prediction of the recombination rate variation in barley – A route to highly recombinogenic genotypes. *Plant Biotechnology Journal*, 20: 676–690.

² **Casale, F. A.**, Arlt, C., Kühn, M., Li, J., Engelhorn, J., Hartwig, T., and Stich, B. (2024). The role of methylation and structural variants in shaping the recombination landscape of barley. Submitted for publication.

³ Schneider, M., **Casale, F. A.**, and Stich, B. (2022). Accurate recombination estimation from pooled genotyping and sequencing: a case study on barley. *BMC Genomics*, 23:468.

1 Summary

Meiotic recombination is a fundamental mechanism for the adaptation of sexually reproducing eukaryotes. Furthermore, it is also crucial for accumulating favorable alleles in plant breeding populations. However, the effective manipulation of the recombination rate still requires a better understanding of the mechanisms regulating the rate and distribution of recombination events in plant genomes. The present paper accumulation thesis aims to pave the road in such direction using barley (*Hordeum vulgare*) as a model species. The core plant material throughout the work is a set of 45 segregating populations derived from crosses that followed a double round-robin design (DRR populations) among 23 inbreds with origins worldwide. Firstly, the recombination rate variation among the DRR populations has been assessed using genetic maps, revealing extensive variation genome-wide and locally in the genome among populations. A mixed-model approach (Best linear unbiased prediction, BLUP) has been used to quantify the importance of the general recombination effects (GRE) of individual parental inbreds from the specific recombination effects (SRE) caused by the combinations of parental inbreds. The variance of the genome-wide GRE was found to be several times the variance of the SRE, indicating that parental inbreds differ in the efficiency of their recombination machinery. Genomic selection (GS) using BLUP was shown to provide a high ability to predict the recombination rate of an inbred line. This demonstrated the possibility to screen large genetic materials for their recombining effect in their progeny and to manipulate the recombination rate using natural variation. Secondly, the genomic features that better explain the recombination variation among the DRR populations were identified at a resolution of 1 Mbp. The genetic effects (GREs not assigned to methylation) were found to be the most important factor explaining differences in recombination rates among populations along with the methylation and the parental sequence divergence. The parental sequence divergence had a sigmoidal correlation with recombination, indicating an upper limit of mismatch among homologous chromosomes for crossover (CO) formation. In addition, the occurrence of hotspots and coldspots for recombination was detected at 10 kb genomic windows, and how methylation and structural variants (SVs) determine such regions was investigated. The inheritance of a highly methylated genomic fragment from one parent only was enough to generate a coldspot but both parents must be equally low methylated at a genomic segment to allow a hotspot. Our findings suggest that recombination in barley is highly

predictable, occurring mostly in multiple short sections located in proximity to genes and being modulated by local levels of methylation and SV load. Lastly, the reliability of a new approach that uses the allele frequency differences and physical distance of neighboring polymorphisms to estimate the recombination rate from pool sequencing was demonstrated with computer simulations and experimentally on the DRR populations. This approach implies a reduction in the cost compared to recombination rate estimations based on genotyping single individuals.

2 General Introduction

2.1 What is meiotic recombination?

The majority of multicellular eukaryotes reproduce sexually, involving a cycle where the fertilization among cells with a half set of chromosomes (i.e., gametes) generates a zygote with a full set of chromosomes that can develop into an independent individual capable of generating new gametes to start the cycle again (Urry et al., 2017). The gametes are produced by meiosis: a type of cell division where the chromosomes after being replicated once (in the S-phase during pre-meiosis) follow two rounds of division —the segregation of homologous chromosomes in Meiosis I and the separation of sister chromatids in Meiosis II—, thus generating cells with half of the parental ploidy (for a review see Mercier et al. 2015). During the prophase of Meiosis I, every chromosome composed of two replicated sister chromatids identifies and physically links to each homolog forming a bivalent prior to their random segregation to opposite cell nucleus poles. Such link is granted by the occurrence of at least one crossover (CO), visualized at the cytological level as chiasmata, through which meiotic recombination —the reciprocal exchange between two homologous non-sister chromatids— takes place (Morgan, 1916; Muller, 1916; De Massy, 2013). The resulting recombination of homologs followed by their random assortment generates new combinations of alleles that can be transmitted to the next generation (Burt, 2000; Barton and Charlesworth, 1998).

2.2 Significance of recombination in nature

2.2.1 Advantage of recombination in natural populations

Sexual reproduction with recombination is widely conserved across eukaryotes since it produces a rapid source of genetic variability upon which natural selection can act, thus being a crucial process for adaptation that facilitates through multiple mechanisms (Weismann, 1886; Smith, 1978). The most accepted explanation for the prevalence of sex and recombination in eukaryotes is the maintenance of high additive genetic variance in fitness (Barton, 2009), i.e., natural selection without recombination is expected to narrow down genetic variation (Charlesworth et al., 1993). Back in the 30s, Fisher (1930) and Muller (1932) hypothesized that sexual reproduction and recombination could enhance the probability of fixation of ben-

eficial mutations because they allow two or more beneficial mutations that arise in different individuals to be united in the same genome when otherwise would outcompete each other and only one could be fixed (Peck, 1994). This would reduce one form of the so-called ‘Hill-Robertson interference’: when advantageous alleles at different loci arise in different backgrounds they would outcompete each other, being the unique possibility to co-exist when a given beneficial allele is close to fixation and the other arises spontaneously by mutation which is very unlikely, thus both ways producing a slower adaptation compared to the allelic combination driven by recombination (Hill and Robertson, 1966). In addition, recombination increases the chances of fixation of new beneficial mutations because frees such from a typically random and negative association with their genetic background that would drive new mutations to be lost, thus maintaining the power of selection versus random drift (Muller, 1932; Barton, 2009; Ritz et al., 2017). Moreover, the re-shuffling of alleles produced by recombination breaks the linkage between beneficial and deleterious mutations, avoiding genetic hitchhiking (i.e., genetic draft), when an allele changes its frequency due to linkage with a locus under selection (Smith and Haigh, 1974), and thus decreasing another form of the ‘Hill–Robertson interference’ that is when a deleterious allele is linked to an advantageous allele reducing the selection efficiency on the beneficial one (Ritz et al., 2017). For example, it has been observed, especially in asexually reproducing species, the accumulation of deleterious mutations in populations in the absence of recombination (Andersson and Hughes, 1996; Söderberg and Berg, 2011), a phenomena known as Muller’s ratchet effect (Muller, 1932; Felsenstein, 1974). In this sense, sexual reproduction with recombination facilitates the accumulation of favorable mutations in a context where detrimental mutations occur much more often, thus providing a faster adaptation to sexual species over asexual ones (Smith, 1978; Peck, 1994; Ritz et al., 2017).

2.2.2 Evolutionary constraints on recombination rate variation

Despite the benefits of recombination for adaptation, the observed limit of CO per chromosome per meiosis in nature suggests that an upper limit for the recombination rate might be evolutionary beneficial (Ritz et al., 2017). Recombination can break apart beneficial alleles stacked on the same haplotype, potentially decreasing the fitness of a population. Thus, positive epistasis would drive selection to limit recombination across associated adaptive loci (Smith, 1978; Otto and Lenormand, 2002). In relation to this, previous studies have shown that chromosomal inversions suppressing recombination between genes, preserve the combination of

alleles that independently increase fitness (Hoffmann and Rieseberg, 2008; Stevenson et al., 2011). Moreover, large structural variations may disrupt the course of meiosis leading to a decrease in hybrid fertility. Therefore, inducing reproductive isolation and leading to speciation (Stevenson et al., 1998; Fuller et al., 2018; Boideau et al., 2022). In this way, during speciation, the coadaptation of beneficial alleles favoring isolation is expected to indirectly promote the natural selection against recombination (Ortiz-Barrientos et al., 2016), e.g., the reduced or even suppressed recombination in animal sex chromosomes which is believed to be a mechanism to preserve the association of loci with sex-specific functions (Rice, 1987). On the reverse, in the absence of geographic barriers, recombination is considered to be the main obstacle to speciation by homogenizing the genetic variation generated by divergent natural selection (Ortiz-Barrientos et al., 2016).

Consequently, these two evolutionary effects of recombination oppose the evolution of sex with the formation of new species, as the limit for one promotes the other, providing grounds for the existence of an optimal recombination rate level for a given natural population under selection (Otto and Lenormand, 2002; Ortiz-Barrientos et al., 2016). Studies reporting intra-species recombination rate variation across taxa indicate that such optimal is close to one CO per chromosome, suggesting species would share a common balance of cost and benefits establishing the upper and lower limits of the recombination rate (Mercier et al., 2015; Ritz et al., 2017). Also important, an optimal level of the recombination rate implies that the recombination rate is an adaptive trait itself and can evolve to increase the efficacy of selection (Burt, 2000). For example, high recombination rates may be favored under strong selection such as fluctuating environmental conditions (Sasaki and Iwasa, 1987). In this respect, because COs are the main factor determining linkage disequilibrium (LD) breakdown, modifying levels of LD must have fitness consequences that alter the action of selection on the recombination rate (Kim et al., 2007; Drouaud et al., 2013; Ortiz-Barrientos et al., 2016).

2.3 The molecular basis of meiotic recombination in plants

Meiotic recombination initiates with the occurrence of a large number of programmed DNA double-strand breaks (DSBs) induced by the SPO11 protein which is highly conserved among eukaryotes (Szostak et al., 1983; Keeney et al., 1997; Grelon, 2001; Edlinger and Schlögelhofer, 2011). In most eukaryotes, SPO11 is encoded by a single gene, but in plant genomes, several *SPO11* homologs exist (Stacey et al., 2006; Hartung et al., 2007). Other additional proteins needed for

the formation of DSBs are less conserved across kingdoms either at the sequence or the functional level, indicating variation in the recombination machinery across species, e.g., RAD50, MRE11, and XRS2 are required proteins for DSB formation in budding yeast (*Saccharomyces cerevisiae*) while their orthologs are instead required for DSB processing in the plant model species *Arabidopsis thaliana* (Mercier et al., 2015). Across mammal species, DSB induction by SPO11 is triggered by the binding of the histone methyltransferase PRDM9 to a specific DNA sequence motif (De Massy, 2013). Similarly, proteins required for DSB formation in *Arabidopsis* as such as DFO and PRD3 -and its homolog in rice (*Oryza sativa*), PAIR1- appear to be plant-specific (Nonomura et al., 2004a; Muyt et al., 2009; Zhang et al., 2012a). In addition, some protein functions might vary between plant species such as the CRC1 protein (Miao et al., 2013; Mercier et al., 2015).

After cleavage, SPO11 remains covalently bound to the 5' ends of DNA through a tyrosine, resulting in the formation of SPO11-oligos (Neale et al., 2005; Keeney and Neale, 2006). Such are recognized and processed by the MRN/MRX complex leading to SPO11-oligo removal after which the 5' ends of the DSBs are resected to produce 3'-OH single-stranded DNA (ssDNA) tails on either side of the break (Rothenberg et al., 2009; Garcia et al., 2011). The ssDNA tails are subsequently invaded by the recombinases RAD51 and DMC1 to form nucleoprotein filaments which can invade either its own intact sister chromatid (i.e., inter-sister repair) or one of the two non-sister homologous chromatids to generate an inter-homologous invasion (D-loop) (Kasamatsu et al., 1971; Keeney and Neale, 2006; Hunter, 2015). The D-loops can be dissolved to form noncrossovers (NCOs) —the unidirectional copy of a small fragment (kilobases or less) from any of the intact homologous (non-sister) chromatids to the broken chromatid without affecting the template—, or further resolved as COs —reciprocal exchanges of large DNA regions (usually megabases) between homologous non-sister chromatids— (Szostak et al., 1983; Hunter, 2015).

Different pathways lead to COs and NCOs. In most eukaryotes, COs are generated through two pathways producing either class I or class II COs, respectively (Mercier et al., 2015). In a single meiosis, successive Class I COs along a given chromosome arm occur more widely spaced than expected by chance, a phenomenon referred to as CO interference (firstly described in *Drosophila melanogaster* by Muller 1916, 1932), while Class II COs occur distributed independently of one another (Berchowitz and Copenhagen, 2010; Youds and Boulton, 2011). Both CO types have been shown to cohabit in plants where Class I being the big majority (85–90%) of the generated COs (Copenhagen et al., 2002; Mercier et al., 2005; Basu-Roy et al., 2013). The biological importance of CO interference, a widespread phenomenon across eukaryotes, is to prevent subsequent COs from overlapping in

the same meiosis (i.e., double COs) (Copenhaver et al., 2002; Mancera et al., 2009; Crismani et al., 2013; Ritz et al., 2017). The Class I COs are generated through a pathway dependent on a group of proteins called ZMMs (firstly, identified in budding yeast by Börner et al. 2004) which in plants include MSH4/MSH5 MutS-related heterodimers, MER3 DNA helicase, PARTING DANCERS (PTD), ZIP4/SPO22, and the SHORTAGE OF CROSSOVERS1 (SHOC1) XPF nuclease (Higgins et al., 2004; Mercier et al., 2005; Wijeratne et al., 2006; Chelysheva et al., 2007; Macaisne et al., 2008; Higgins et al., 2008a; Shen et al., 2012; Luo et al., 2013). Such pathway generates many double Holliday junction (dHJ) intermediates (Holliday, 1964; Wyatt and West, 2014) a few of which are marked by HEI10 E3 ligase and the MLH1/MLH3 MutL-related heterodimer to mature into COs (Chelysheva et al., 2012; Wang et al., 2012; Ziolkowski et al., 2017; Serra et al., 2018b). Most ZMM pathway-related proteins have been found to be strong modifiers of the recombination rate in plants (Mercier et al., 2015). For example, the introduction of additional HEI10 coding gene copies in *Arabidopsis* was found to elevate euchromatic COs genome-wide (Ziolkowski et al., 2017; Serra et al., 2018b). In this way, *Arabidopsis* and rice mutants silencing ZMM proteins eliminate close to 85% of COs, but none could decrease beyond 90% (Chelysheva et al., 2007; Higgins et al., 2008a; Chelysheva et al., 2012; Shen et al., 2012; Wang et al., 2012; Luo et al., 2013). The remaining class II COs are produced by a ZMM-independent pathway which is less studied (Mercier et al., 2015). The only related protein characterized in plants is MUS81 whose mutants show a decreased recombination of 10% (Berchowitz et al., 2007; Higgins et al., 2008b).

The repair of inter-homologous intermediates resulting in NCOs can be mediated by multiple pathways including synthesis-dependent strand annealing (SDSA) and dHJ dissolution, among other less studied (Allers and Lichten, 2001a; McMahon et al., 2007). These pathways are referred to as anti-crossover and their combination produces the majority of DSBs to result in NCOs (close to 90%), resulting in the low ratio CO/DSB per meiosis observed across kingdoms, including plants (Mercier et al., 2015; Ziolkowski et al., 2017). Several proteins involved in such processes were experimentally shown to prevent CO formation in *Arabidopsis* such as the homolog of the human Fanconi anemia complementation group M helicase (FANCM) and its two cofactors MHF1 and MHF2, the DNA helicases RECQ4A and RECQ4B, the TOPOISOMERASE3 α (TOP3 α) and its co-factor RMI1, and the AAA-ATPase FIDGETIN-LIKE-1 (FIGL1) (Hartung et al., 2008; Knoll et al., 2012; Crismani et al., 2012; Girard et al., 2014, 2015; Séguéla-Arnaud et al., 2017; Fernandes et al., 2018a).

At the site of the strand invasion where both CO and NCO occur, the annealing generates heteroduplex DNA which produces mismatches between the homologous

chromosomes if sequence polymorphisms exist among them (Allers and Lichten, 2001b). Such mismatches are recognized by the Mismatch Repair (MMR) system that mends the mismatch either in favor of the sister chromatid restoring the original allelic state (i.e., inter-sister repair) (Borts et al., 2000; Goldfarb and Lichten, 2010; Spies and Fishel, 2015) or in favor of the homologous allele resulting in gene conversion (GC): the non-reciprocal exchange of alleles between homologous non-sister chromatids (Zickler, 1934; Holliday, 1964). The inter-sister repair maintains the expected Mendelian 2:2 ratio of alleles at the repaired locus, but GC generates a segregation bias in favor of the non-inducing allele, thus altering the allele frequency to a non-Mendelian 3:1 ratio and generating new alleles in an otherwise unchanged genetic background (Sun et al., 2012; Wijnker et al., 2013). In addition, in heterozygous individuals, CO-associated conversion tract (COCTs) are long enough to assume that SNPs are likely to occur at the site of CO initiation and thus each CO is expected to generate a GC (Burt, 2000; Lu et al., 2012). In contrast, most NCOs do not generate GCs because the short fragments that NCOs encompass are unlikely to overlap with polymorphisms. Nevertheless, GCs are the only way to detect NCOs in the resulting gametes (Wijnker et al., 2013). It is worth noting that GC is a phenomenon that occurs beyond meiotic recombination as it can also occur during mitosis as well as among non-allelic sequences that share homology (i.e., ectopic), previously shown in maize (*Zea mays*) (Shalev and Levy, 1997; Chen et al., 2007).

Meiotic chromosomes are organized in loops, being sister chromatids attached to a common protein axis (i.e., the axial element) formed of cohesins and other meiosis-specific proteins (Zickler and Kleckner, 1999; Panizza et al., 2011). There are two axial-composing proteins identified in plants: the HORMA domain-containing protein ASY1 (PAIR2 in rice) and ASY3 (PAIR3 in rice) which were suggested to regulate the choice between inter-homolog and inter-sister recombination (Nonomura et al., 2004b; Sanchez-Moran et al., 2007; Yuan et al., 2009; Ferdous et al., 2012). Concurrently with the rise of DSBs on the chromatin loops in mid-prophase, the two homologous chromosome axial elements are held through a central element, forming a structure called the synaptonemal complex (SC) from which the loops of chromatin are projected laterally (Padmore et al., 1991; Kleckner, 2006). A central protein of the SC is the ZIPPER1 (ZYP1), firstly characterized in budding yeast (Börner et al., 2004), which homologs have been found in plants (Higgins et al., 2005; Wang et al., 2010; Barakate et al., 2014). Plant ZYP1-deficient mutants were reported to show recombination defects (Higgins et al., 2005; Wang et al., 2010; Barakate et al., 2014; Wang et al., 2015; Capilla-Pérez et al., 2021), thus indicating a central role in meiotic recombination of the SC which exact function is yet not fully understood (Mercier et al., 2015).

Moreover, the length of the SC has been found to be positively correlated with CO rate in several organisms and has been associated with the capacity to pack greater chromatin loops (Lynn et al., 2002; Kleckner et al., 2003; Giraut et al., 2011; Lu et al., 2012; Li et al., 2015).

Meiosis is driven by cyclins and cyclin-dependent kinases (CDKs) complexes like mitosis but fine-regulated enough to ensure ploidy reduction by avoiding the alternation between replication and division of mitosis (Carlile and Amon, 2008; Bulankova et al., 2010). This implies cyclin-CDK activity to be low enough to exit meiosis I but not too low as to generate a mitosis-like cell division and enter the second division round without replication (Futcher, 2008; D’Erfurth et al., 2009). In addition, to ensure a balanced chromosome distribution until the second division, the two rounds of chromosome segregation need tight control of both the release of sister chromatid cohesion and the change in the kinetochore orientation (Crismani et al., 2013; Mercier et al., 2015). The mechanism providing cohesion at mitosis and meiosis –cohesin bonded by adherin and hydrolyzed by separase– was found to be conserved across eukaryotes (Liu and Makaroff, 2006; Sebastian et al., 2009; Singh et al., 2013). Related proteins identified in plants involve the mitosis-meiosis-related cohesin sub-unit SCC3 and the meiosis-specific cohesin sub-unit REC8 (Bhatt et al., 1999; Chelysheva et al., 2005; Golubovskaya et al., 2006; Shao et al., 2011; Shen et al., 2012; Lambing et al., 2020; Dreissig et al., 2020). In plants, the absence of REC8 leads to DSB repair defects, loss of sister chromatid cohesion, and wrong orientation of kinetochores (Bhatt et al., 1999; Chelysheva et al., 2005; Golubovskaya et al., 2006; Shao et al., 2011).

2.4 The assessment of recombination events in plants

Research on recombination and, thus, the detection of recombination events, is a long-standing topic in science given the importance of the field in biology. In this way, phenotypic, cytological, and molecular approaches have been implemented either to assess meiotic recombination or to evaluate its properties in numerous organisms (Toyota et al., 2011).

2.4.1 Crossovers (CO)

Among the different recombination events, the occurrence of COs has been by far the most studied one because it is the easiest to detect and the most determining on the generation of genetic variation. The presence of COs can be detected either

during meiosis or later in both the meiotic products and the offspring.

Firstly, phenotypic screens of qualitative traits have been used to calculate the frequency of recombinant offspring as a measure of the linkage between the loci determining the phenotype under study which is a function of the CO frequency among such loci. Thomas Hunt Morgan did that for the first time by analyzing offspring from a double heterozygous parent of *Drosophila melanogaster* and realized that CO frequency among linked loci was related to their intervening distance, i.e., genetic distance (that later gave rise to the centiMorgan (cM) unit to quantify genetic distance among loci) (Morgan, 1911, 1912). Morgan’s student, Alfred Henry Sturtevant, realized that such calculation allowed to map linked loci linearly on a chromosome, i.e., the genetic map also called the linkage map (Sturtevant, 1913, 1915; Morgan, 1916). Genetic maps have been widely used to map quantitative trait loci (QTL) across organisms (Nordborg and Weigel, 2008) while they are also useful to both compare the recombination rate among populations of the same (or genetically close) species by assuming individuals to have similar genome physical lengths (Williams et al., 1995; Cai et al., 2014; Phillips et al., 2015) or to estimate the recombination rate per se by dividing genetic distance over physical genomic distance at any scale (most typically, cM/Mbp) (Pedersen et al., 1995; Künzel et al., 2000; Wu et al., 2003). It is noteworthy that a mapping function that converts recombinant fraction to genetic distance is needed because with increasing genetic distance the probability of double COs also increases, and thus, the frequency of recombinants does not reflect the true probability of COs (i.e., odd COs are detected but even COs maintain the allelic combinations of the parents) (Sturtevant, 1913; Castle, 1919; Sturtevant et al., 1919). In addition to Morgan’s, different genetic mapping functions exist to account for different patterns of CO interference (Wei et al., 2020). The most used are Haldane’s which assumes non-interference, and Kosambi’s and Carter-Falconer’s, both assuming interference (Haldane, 1919; Kosambi, 1943; Carter and Falconer, 1951). Many other mapping functions have been designed in later years to consider the different recombination patterns across species (Tan and Fornage, 2008).

Phenotypic screens to assess recombination have been used in plants for a long time, for example, to analyze the recombination in reduced (i.e., no genome-wide) regions of maize, such as the bronze gene (*bz*) and its neighboring loci (e.g., the waxy locus, *wx*) (Dooner, 1986; Dooner and Martínez-Férez, 1997; Okagaki and Weil, 1997; Fu et al., 2001; He and Dooner, 2009) and the *anthocyaninless1* gene (*a1*) and its neighboring region (Brown and Sundaresan, 1991; Yao et al., 2002; Yao and Schnable, 2005; Yandeau-Nelson et al., 2005, 2006). More contemporary studies in plants used transgenic-fluorescent-tagged seeds to spot recombinants in *Arabidopsis* (Melamed-Bessudo et al., 2005; Melamed-Bessudo and Levy, 2012;

Yelina et al., 2012; Ziolkowski et al., 2017).

In addition, the direct visualization of chiasmata by observing meiotically active cells with microscopy has been used to assess COs for years (Stack and Soulliere, 1984; Säll et al., 1990; Herickhoff et al., 1993). In plants, this approach has been combined with techniques for fluorescent microscopy –such as fluorescence *in situ* hybridization (FISH), 4',6-diamidino-2-phenylindole (DAPI), or immuno-staining– in different species such as *Arabidopsis* (Stevenson et al., 1998; Sanchez-Moran et al., 2002; Chelysheva et al., 2007; López et al., 2012; Varas et al., 2015), rice (Wang et al., 2010), barley (*Hordeum vulgare*) (Leitch and Heslop-Harrison, 1993; Higgins et al., 2012; Phillips et al., 2015), *Brassica* spp. (Leflon et al., 2010; Higgins et al., 2020), among other species (King et al., 2002; Anderson et al., 2003). Furthermore, the direct visualization of either late recombination nodules (RNs or LNs) or synaptonemal complexes was used in maize, wheat (*Triticum aestivum*), and tomato (*Solanum lycopersicum*) to assess recombination (Herickhoff et al., 1993; Gill et al., 1996; Stack and Anderson, 2002; Anderson et al., 2003; Lhuissier et al., 2007).

More recently, molecular genetic approaches have been implemented to assess recombination in plants and other organisms. The immuno-staining (i.e., labeling) of meiosis-related proteins at leptotene-stage meiotic chromosomes can be analyzed with fluorescence microscopy to detect COs and other related processes (Choi et al., 2013; De Massy, 2013), e.g., to track MLH1 activity in *Arabidopsis* (Pawlowski et al., 2003; Esch et al., 2007; Chelysheva et al., 2012), *Brassica* spp. (Leflon et al., 2010), and tomato (Lhuissier et al., 2007). In addition, in *Arabidopsis*, chromatin immuno-precipitation assays (ChIP) of meiotic proteins associated with CO have been sequenced (ChIP-Seq) to map COs at high-resolution (Choi et al., 2013).

The above-mentioned phenotypic and cytogenetic measurements have been used for genetic distance calculation for years (Dooner, 1986; Anderson et al., 2003), however, in the molecular marker era, the ordering of polymorphic genomic markers based on the log-likelihood (e.g., LOD) of the recombination fraction between markers became the standard for genetic map construction (Lander and Botstein, 1989; Esch et al., 2007; Ganai et al., 2011; Bauer et al., 2013). In this way, the resolution and reliability of genetic maps have evolved along with the increasing availability of molecular markers, from being of low resolution and reduced genomic representation to genome-wide high-resolution maps. For example, previously, due to the length and polyploid nature of the wheat genome, recombination analyses in wheat had been conducted on single chromosomes with low-resolution (Gill et al., 1996; Stein et al., 2000; Saintenac et al., 2009, 2011; Darrier et al., 2017) while more recent studies in wheat are genome-wide high-resolution assessments (Jordan et al., 2018; Gutierrez-Gonzalez et al., 2019). Different al-

gorithms for genetic map construction have been implemented in software such as, in order of appearance: MAPMAKER (Lander et al., 1987), JoinMap (Stam, 1993), CARHTAGENE (Schiex and Gaspin, 1997), Map Manager QTX (Manly et al., 2001), R/QTL (Broman et al., 2003), AntMap (Iwata and Ninomiya, 2006), TMAP (Cartwright et al., 2007), MSTMAP (Wu et al., 2008a), R/mpMap (Huang and George, 2011), and Multipoint-UDM (MUDM) (Ronin et al., 2017). Moreover, previously, reported genetic maps had low species-wide representation because were based on population-specific offspring from a few parental genotypes, e.g., recombinant inbred lines (RILs) derived from the maize IBM (Lee et al., 2002) or the wheat SynOpRIL (W7984 x Opata) populations (Sorrells et al., 2011; Gutierrez-Gonzalez et al., 2019). In recent years, the advent of multi-parent populations developed for high-precision QTL mapping for complex traits enabled high-resolution species-wide representative genetic maps (Cavanagh et al., 2008; Sannemann et al., 2015). The most remarkable of this kind in plants are the American and Chinese NAM maize populations (Yu et al., 2008; McMullen et al., 2009; Rodgers-Melnick et al., 2015), Dent and Flint European maize genetic pools (Bauer et al., 2013), wheat NAM population (Jordan et al., 2018), and the barley HEB-KI NAM population (Dreissig et al., 2020). The integration of several genetic maps into consensus genetic maps has also become possible through algorithms implemented in software such as CARHTAGENE (de Givry et al., 2005) and MergeMap (Wu et al., 2008b).

In addition to the construction of genetic maps, the genotyping of segregating populations (i.e., sequencing of parental lines and offspring at polymorphic positions) allows to directly detect the trace of COs that occurred from the F1 to the sequenced generation (COs are detected as switches of parental alleles between two markers of known physical order) (Darrier et al., 2017). In this way, the genotyping of F2 or backcross (BC) populations reflects the precise CO distribution in F1 meiosis without the existence of double-crossover if interference is assumed (Mirouze et al., 2012). Several studies in *Arabidopsis* used this approach by genotyping either the F2 or BC coming from an F1 between the accessions Columbia and Landsberg erecta (Drouaud et al., 2006; Giraut et al., 2011; Toyota et al., 2011; Salomé et al., 2012; Yang et al., 2012; Mirouze et al., 2012; Rowan et al., 2015; Ziolkowski et al., 2017; Rowan et al., 2019; Blackwell et al., 2020). In grasses, this approach has been used in rice (Si et al., 2015) and wheat, the last by sequencing F2s derived from the cross between the cultivars Chinese Spring (Cs) and Renan (Re), and Cs and Courtot (Ct) (Saintenac et al., 2009, 2011). A similar option is the genotyping of double haploids (DHs) coming from an F1 as previously performed in *Arabidopsis* (Wijnker et al., 2013) and wheat (Saintenac et al., 2009; Gutierrez-Gonzalez et al., 2019).

Alternatively, the genotyping of recombinant inbred lines (RILs) allows to detect CO position but not precisely estimate the rate of CO per generation because the multiple crossover events occurring between the same two markers in subsequent selfing generations (departing from the F1) will be underscored while approaching homozygosity as well as the overlap of multiple COs also makes difficult to discriminate between NCOs and close CO events that generated double-allele swaps (Esch et al., 2007; Darrier et al., 2017). However, because RILs are useful for several genetic studies in research and breeding programs such as QTL analysis and genomic selection (GS), RILs are a widely available resource to estimate recombination in plants. RILs have been used to estimate recombination, by either allele switches or genetic maps, in *Arabidopsis* (Esch et al., 2007), maize (Esch et al., 2007; McMullen et al., 2009), wheat (Esch et al., 2007; Darrier et al., 2017; Jordan et al., 2018; Gardiner et al., 2019), barley (Dreissig et al., 2020), among other species. In *Arabidopsis*, this approach has been combined with distinct methylated parental mutants to generate epiRILs which can be used to analyze the impact of the segregation of methylation states on the frequency and distribution of recombination events along chromosomes (Mirouze et al., 2012; Colomé-Tatché et al., 2012). Notably, the same as occurred with genetic maps, the precision in the assessment of alleles switches in segregating populations increased notably with the availability of high-throughput genotyping technologies that allowed high-resolution analyses that capture the recombination variation along the genome at the hotspot level (< 100 kb). Up to now, high-resolution analyses have been performed in *Arabidopsis* (Sun et al., 2012; Lu et al., 2012; Yang et al., 2012; Wijnker et al., 2013; Rowan et al., 2019) and crop species such as maize (Rodgers-Melnick et al., 2015), rice (Si et al., 2015), and wheat (Gardiner et al., 2019).

The analysis of haplotype markers in population studies can only detect fixed recombination events that occurred at several meioses either in one or many contiguous generations (Lu et al., 2012). Alternatively, tetrad analysis, the assessment of all four daughter cells from a single meiosis, allows not only to detect the changes generated in one meiosis process but also to distinguish from two to four-strand CO events, enabling GC detection by observing 3:1 inheritance between sister gametes (Choi and Henderson, 2015). This method was first applied in budding yeast because, like other fungi, meiotic products are kept together as spores in an ascus forming a tetrad, thus facilitating the analysis of the gametes coming from the same meiosis (Nicolas et al., 1989). In contrast, in most flowering plant species gametes do not remain together after meiosis. In this sense, in maize, microspores have been separated manually under the microscope when they were still united as a tetrad during pollen grain formation and subsequently sequenced at high-resolution (Li et al., 2015). In *Arabidopsis*, tetrad analyses have been

performed with *qrt1* mutants that produce the four products of male meiosis to remain united after meiosis (i.e., attached spores that develop into attached functional pollen grains) (Preuss et al., 1994; Francis et al., 2006). Such *Arabidopsis* pollen tetrads have been combined with linked heterozygous transgenes that express different fluorescent-colored proteins (Fluorescent Tagged Lines, FTLs) to facilitate the visual CO count Francis et al. (2007); Berchowitz and Copenhaver (2008); Sun et al. (2012); Yelina et al. (2012, 2013), or crossed to a common genetic background for further sequencing of the offspring (Lu et al., 2012; Wijnker et al., 2013).

It is worth noting that the above-mentioned visual reporters (FTLs) on pollen grains allowed for the high-throughput analysis of recombination in *Arabidopsis* (Francis et al., 2007; Yelina et al., 2013) which is not possible by other previously-described methods because of the time and cost associated per sample. Similarly, an assay based on seed fluorescent markers has been developed for the rapid screening of recombinants in *Arabidopsis* (Melamed-Bessudo et al., 2005; Yelina et al., 2012). In addition, pollen typing (i.e., sperm typing) allows the high-throughput CO determination of bulk recombinants by the amplification of allele-specific oligonucleotides (ASOs) paired to polymorphic sites in gametes from heterozygous individuals. This method allows the detection of a large number of COs at a given hotspot which is barely possible with the CO detection methods that rely on plant populations as hotspots have relatively short genetic distances (0.1–0.5 cM), thus the screen of many meioses is required to catch several COs in the same hotspot (Choi and Henderson, 2015). This is typically done with male meiotic products (pollen or sperm) because these are easier to isolate but the analysis with female gametes is possible (Choi and Henderson, 2015). In plants, this method has been used in *Arabidopsis* (Yelina et al., 2012; Choi et al., 2013; Drouaud et al., 2013) and barley (Dreissig et al., 2015). The described high-throughput analysis methods can only assess COs in limited regions of the genome. Alternatively, the linked-read bulk sequencing (i.e., pool sequencing) of gamete samples from *Arabidopsis* recombinant plants has been used to efficiently generate a genome-wide CO map with a single sequencing experiment and without growing plants (Sun et al., 2019). In relation to this, genome-wide recombination rates were estimated from pooled sequencing samples of *Drosophila melanogaster* based on the segregation distortion decay from loci under selection (i.e., loci whose alleles have a fitness differential) to distant loci as such attenuation is a function of the genetic distance between loci (i.e., depends on the CO frequency among the locus under selection and the rest) (Wei et al., 2020).

Finally, historical recombination (i.e., ancestral recombination) can be estimated for a sample of unrelated individuals using the observed LD among a set

of polymorphic markers which is considered to be a product of the decay in LD generated by the CO occurred over all generations involving the ancestors of the sample (assuming such departed from a common ancestor of the species) (Lewontin and Kojima, 1960; Mcvean et al., 2002). Based on this, several available software –such as LDhat (Mcvean et al., 2002), Hotspotter (Li and Stephens, 2003), LDhot (Myers et al., 2005), and SequenceLDhot (Fearnhead, 2006)– employ coalescent theory to calculate backward in time the number of CO events occurred among the used marked loci departing from the respective pattern of non-random association of alleles, a technology that was firstly applied to human genetics (Mcvean et al., 2002; McVean et al., 2004; Li and Stephens, 2003; Myers et al., 2005). The CO rate calculated using this approach in a given genomic segment is a quantity called population-scaled recombination rate (ρ), defined as $\rho = 4Ner$, where r is the recombination rate per generation occurring in the segment and N_e is the effective population size (Mcvean et al., 2002). In plants, this approach has been used in *Arabidopsis* (Horton et al., 2012; Choi et al., 2013), maize (Hufford et al., 2012; Rodgers-Melnick et al., 2015), rice (Marand et al., 2019), wheat (Darrier et al., 2017), barley (Dreissig et al., 2019), sorghum (*Sorghum bicolor*) (Morris et al., 2013), rye (*Secale cereale*) (Schreiber et al., 2022), cotton (*Gossypium hirsutum* and *G. arboreum*) (Shen et al., 2019), *Medicago truncatula* (Branca et al., 2011; Paape et al., 2012), *Eucalyptus grandis* (Silva-Junior and Grattapaglia, 2015), *Populus* spp. (Slavov et al., 2012; Wang et al., 2016; Apuli et al., 2020), and monkeyflower *Mimulus guttatus* (Hellsten et al., 2013). Importantly, when sampling the historical number of meiosis occurring among a group of individuals, the existing population structure -generated by population genetic forces, such as mutation, selection, and drift- might obscure the observed non-random associations between polymorphic markers of the sample. In addition, structural variants among the compared individuals might overestimate the real genetic variation generated by recombination, thus the necessity to combine historical with mapping populations-based calculations has been suggested (Choi and Henderson, 2015; Darrier et al., 2017; Marand et al., 2019; Apuli et al., 2020).

2.4.2 Noncrossovers (NCOs) and gene conversions (GCs)

Molecularly, NCOs only leave a genetic trace when led to GCs. Thus, to be detectable, NCOs must both occur between polymorphic segments and not be restored to the parental genotypes via sister-chromatid repair (Lu et al., 2012; Sun et al., 2012; Wijnker et al., 2013). Inter sister-chromatid repair has been widely observed in yeast (Hyppa and Smith, 2010; Kim et al., 2010; Goldfarb and Lichten, 2010) and it is expected to be a common DSB output in plants, being

already reported in *Arabidopsis* (Cifuentes et al., 2013). A related complication is that GC events are of very short length –reported to be less than 2 kb in plants (Drouaud et al., 2013; Wijnker et al., 2013; Li et al., 2015)– thus making their detection very sensitive to level of polymorphism among homologous chromosomes, used marker density, and tract length of the repair intermediate (Wijnker et al., 2013). For a long time, short shifts in genotyping data were typically disregarded as potential genotyping errors or wrong marker order as markers were on average at least hundreds of kb apart (Yang et al., 2012; Gardiner et al., 2019). Even in the next-generation-sequencing era, the detection of GCs is highly dependent on the reads’ quality alignment, the sequencing coverage, and the quality control of the genotyping data to distinguish between artifactual, heterozygous, and homozygous calls (Wijnker et al., 2013; Qi et al., 2014). For example, it was shown that SVs can lead to false-positive NCO-GC calls in *Arabidopsis* (Wijnker et al., 2013; Qi et al., 2014). Moreover, because tetrads naturally do not remain grouped after meiosis in flowering plants, it is difficult to screen the GC expected 3:1 segregation ratio in plants (Yang et al., 2012; Sun et al., 2012; Wijnker et al., 2013). Another related issue is the difficulty associated to differentiate between CO and NCO-related GCs which makes most authors report their combined frequency or make assumptions to estimate both separate values (Lu et al., 2012). Such causes lead NCO occurrence to be typically under-estimated and NCO rates to be poorly documented in plants (Yang et al., 2012; Mercier et al., 2015). For example, genome-wide studies in *Arabidopsis* reported low frequency of NCOs per meiosis that differ from such estimated from the observed DSB rate (Lu et al., 2012; Sun et al., 2012; Wijnker et al., 2013).

Beyond the above-mentioned constraints, efforts have been made to assess NCO in several organisms by different approaches. The amplification or genotyping of polymorphisms have been used to detect GCs in reduced genomic segments such as the maize *bz* and *a1* loci (Dooner and Martínez-Férez, 1997; Okagaki and Weil, 1997; Dooner, 2002). Similarly, pollen typing has been used to detect NCOs at two recombination hotspots in *Arabidopsis* (Drouaud et al., 2013) and ChIP-Seq has been used to map NCO in maize centromeres (Shi et al., 2010). Among genome-wide analyses, the tetrad analysis of *Arabidopsis* and maize has been the most precise method employed to detect GCs and NCOs (Sun et al., 2012; Lu et al., 2012; Wijnker et al., 2013; Varas et al., 2015; Li et al., 2015). In a few genome-wide genotyping analyses, GC events have been assumed when a polymorphic marker block, typically shorter than 10 kb, switched parental phase but not the flanking markers in close vicinity (Yang et al., 2012; Si et al., 2015; Gardiner et al., 2019). Although such coarse observation can be generated by closely spaced double COs, interference is expected to minimize such situation in many species if

meiotic products from only one meiosis generation are analyzed (F2) as well as the probability of a given recombination event to occur with matching breakpoints in both male and female meiosis is very low at high marker resolutions (Yang et al., 2012). Historical recombination approaches can also be used to estimate GCs in organisms (Gay et al., 2007) as successfully applied in *Brassica* spp. (Xiong et al., 2011; Chalhoub et al., 2014; Cai et al., 2014) and maize (Shi et al., 2010; Rodgers-Melnick et al., 2015).

2.4.3 Double-strand breaks (DSBs)

The effective detection of DSBs presents a challenge because such are temporary events of which only a few will be detectable in meiotic products. In this sense, in several organisms, the genome-wide landscape of DSBs has been revealed indirectly by the identification of COs and NCOs. Such extrapolation is unreliable without establishing a hypothetical system accounting for the respective bias as a result of the recombination regulation (i.e., while a CO indicate the presence of a DSB, the genome-wide landscape of COs in a given species does not represent completely such of the DSBs) (De Massy, 2013). Fortunately, direct assessments of DSBs have been developed for several species providing a reliable notion of the extent of DSBs in genomes. Firstly, based on the DNA molecular size differential caused by DSBs, gel electrophoresis has been used to detect DSBs in budding yeast (Nicolas et al., 1989). The occurrence of DSBs has been also detected by immuno-staining of meiotic proteins –such as RAD51, DMC1, ASY1, ZYP1, and HEI10– at the leptotene-stage of meiotic chromosomes which can then be analyzed with fluorescence microscopy (Choi et al., 2013; De Massy, 2013). The method was applied extensively in *Arabidopsis* (Mercier et al., 2005; Chelysheva et al., 2007; Sanchez-Moran et al., 2007; Esch et al., 2007; Varas et al., 2015; Yelina et al., 2012; Ziolkowski et al., 2017; Blackwell et al., 2020) and other plants species such as maize (Franklin et al., 1999; Pawlowski et al., 2003) and wheat (Gardiner et al., 2019). Importantly, the combined immuno-localization of DSB and CO-related proteins allowed to estimate the rate of DSB that effectively undergone CO in *Arabidopsis* (Varas et al., 2015). In addition, because the SPO11 is removed from the DSBs as SPO11-oligo complexes, the immuno-precipitation of SPO11 and further gel electrophoresis or sequencing (e.g., Chip-Seq and ssDNA- sequencing) can be used to map DSBs at high-resolution (Keeney et al., 1997; Neale et al., 2005; Keeney and Neale, 2006; Pan et al., 2011), and, thus, this is the most sensitive and resolute method currently available for DSB detection (De Massy, 2013). A similar procedure can be done by the immuno-precipitation of recombinases (e.g., DMC1 and RAD51) bound to the meiotic ssDNA produced after DSB resection

(Smagulova et al., 2011; Brick et al., 2012) which has been applied in maize (He et al., 2013).

2.4.4 CO interference

The CO interference is typically modeled as the superposition of COs with no interference that follows a Poisson distribution and is denoted as p , and the remainder (proportion = $1-p$) that generates COs with an inter-CO distance probability that follows a Chi-square distribution (Mortimer and Fogel, 1974; Cobbs, 1978; Stam, 1979; Foss et al., 1993; Lange et al., 1997; Copenhaver et al., 2002). The Chi-square model has a single parameter m which is a non-negative integer and controls the strength of interference, i.e., $m = 0$ corresponds to no interference (Broman et al., 2002). That was later expanded to the Gamma model that allows non-integer values of m (McPeck and Speed, 1995; Broman and Weber, 2000). Recombination studies in different plant species such as *Arabidopsis* (Copenhaver et al., 2002; Berchowitz et al., 2007; Lam et al., 2005; Toyota et al., 2011; Salomé et al., 2012; Lu et al., 2012; Basu-Roy et al., 2013; Rowan et al., 2019), maize (Bauer et al., 2013; Li et al., 2015), wheat (Saintenac et al., 2009), and tomato (Lhuissier et al., 2007), tested if the CO distribution pattern followed the Gamma model, showing that interfering and non-interfering COs cohabit in most plant species (Mercier et al., 2015). In addition, these models have been extensively applied by offspring simulation software such as simcross (Broman et al., 2002), Plabsoft (Maurer et al., 2008), AlphaSimR (Faux et al., 2016), MoBPS (Pook et al., 2020), and genomicSimulation (Villiers et al., 2022), to emulate recombination in hypothetical populations (e.g., simulate populations under different plant breeding schemes).

2.5 Recombination rate variation inter- and intra-species

At least one CO per bivalent is required for the correct segregation of homologous chromosomes (i.e., obligate CO) and, thus, for meiosis to be viable (Hall, 1972; Jones, 1984). Accordingly, it was observed across eukaryotes that the absence of CO can generate chromosomal nondisjunction in meiosis I, generating the unbalanced segregation of homologs that results in aneuploid gametes (Hall, 1972; Koehler et al., 1996). Aneuploid gametes lead to sterility, embryo-lethality, or developmental problems, for example, human trisomies (e.g., Down’s syndrome) are typically related to chromosome nondisjunction as a product of insufficient recom-

ination (Koehler et al., 1996; Ferguson et al., 2007; Nagaoka et al., 2012). Therefore, the obligate crossover per chromosome per meiosis establishes a minimum recombination rate across species. In addition, the observed recombination rate across species revealed an upper limit for the recombination rate (Ritz et al., 2017). Excessive recombination may generate chromosome instability through structural rearrangements causing deleterious phenotypes (Robbins et al., 1991; Sasaki et al., 2010), e.g., some human cancers (Mao et al., 2009; Pal et al., 2011). Indirect evidence of such upper bound in eucaryotes is the observed number of DSBs generated during meiosis which greatly exceeds the number of observed COs (by 10–50-fold in plants), implying most DSBs are resolved as NCOs (Martini et al., 2006; Baudat and Massy, 2007). In *Arabidopsis*, 100 to 250 DSBs estimated by immuno-stained recombination-related proteins lead to about 10 COs per meiosis (Mercier et al., 2005; Chelysheva et al., 2007; Sanchez-Moran et al., 2007), and a mutant with increased DSBs did not lead to a higher rate of CO but GC (Varas et al., 2015). The maintenance of the CO frequency beyond such of DSB is known as CO homeostasis and it has been related to CO interference (Martini et al., 2006; Berchowitz and Copenhaver, 2010). In this way, the observed CO rate in most species is typically between 1–2 and rarely more than 3 COs per chromosome per meiosis independently of the genome size (Baudat and Massy, 2007; Mercier et al., 2015).

Remarkably, the rate of DSBs has been observed to be quite stable across sexually reproducing species, varying between 150–250 DSBs per meiosis (Moens et al., 1997; Buhler et al., 2007; Chelysheva et al., 2007; Sanchez-Moran et al., 2007; Varas et al., 2015) while the CO rate varies dramatically across species (Nachman, 2002; Henderson, 2012). For example, to cite model organisms the CO rate is close to 90 COs per meiosis in budding yeast (Mancera et al., 2009), 5–29 COs per meiosis in the nematode *Caenorhabditis elegans* (Barnes et al., 1995; Rockman and Kruglyak, 2009), 6–10 COs per meiosis in *Arabidopsis* (Giraut et al., 2011; Salomé et al., 2012; Wijnker et al., 2013; Rowan et al., 2019), 5–20 COs per meiosis in *Drosophila* spp. (Kulathinal et al., 2008; Comeron et al., 2012), 8–24 COs per meiosis in mouse (*Mus musculus*) (Anderson et al., 1999; Cox et al., 2009; Dumont et al., 2009), and 23–56 COs per meiosis in humans (Kong et al., 2002; Matise et al., 2007; Coop et al., 2008). In the plant kingdom, the reported genome-wide recombination rates varied from 0.74 to 16 cM/Mb in maize and *Populus tremula*, respectively (Bauer et al., 2013; Silva-Junior and Grattapaglia, 2015).

The magnitude of intra-species recombination variation is quite stable across taxa as reported CO frequencies in eucaryotes do not vary more than 1.1 to 2-fold between the highest- and lowest-recombining genotypes (Ritz et al., 2017). In *Arabidopsis*, previous studies reported CO rates that range from 1.1 to 2.3 COs

per chromosome per meiosis, resulting in genome-wide rates that range from 6 to 10 COs per generation across populations (Sanchez-Moran et al., 2002; Chelysheva et al., 2007; Giraut et al., 2011; Toyota et al., 2011; Salomé et al., 2012; Lu et al., 2012; Yang et al., 2012; Sun et al., 2012; Wijnker et al., 2013; Ziolkowski et al., 2017; Rowan et al., 2019; Blackwell et al., 2020). The recombination rate variation in major crop species was also investigated to some extent. Reports in maize provided rates between 0.4 to 0.8 cM/Mbp on average genome-wide (Anderson et al., 2003; Esch et al., 2007; McMullen et al., 2009; Bauer et al., 2013), 3–12 CO per chromosome per meiosis (Li et al., 2015), and 20–112 COs per RIL (Esch et al., 2007; Rodgers-Melnick et al., 2015). Recent comparisons among the American maize NAM populations revealed extensive intra-species recombination rate variation on a genome-wide level but also along the genome (McMullen et al., 2009), thus confirming earlier observations among individuals of the Iowa Stiff Stalk Synthetic maize population (Fatmi et al., 1993; Hadad et al., 1996; Li and Pfeiffer, 2009). However, both the American and Chinese NAM maize populations were shown to have a similar recombination rate distribution pattern on average among populations (Rodgers-Melnick et al., 2015). Similarly, the European Dent and Flint maize genetic pools were shown to have significant intra-pool but not inter-pool recombination variation (Bauer et al., 2013). Previous studies in rice reported recombination rates from 3.90 to 4.53 cM/Mbp on average genome-wide and from 1.5 to 2.8 COs per chromosome per generation (Wu et al., 2003; Shen et al., 2012; Si et al., 2015). Interestingly, the recently diverged sub-species *indica* and *japonica* correlated in the broad-scale recombination rates but around 80% of their CO hotspots were unique to each subspecies (Marand et al., 2019). In wheat, previous studies reported rates that range from 0.17 to 0.9 cM/Mbp on average genome-wide (Stein et al., 2000; Akhunov et al., 2003; Saintenac et al., 2009, 2011; Jordan et al., 2018; Gutierrez-Gonzalez et al., 2019) and from 1.2 to 2.8 COs per chromosome per RIL (Esch et al., 2007; Darrier et al., 2017; Jordan et al., 2018; Gardiner et al., 2019). Earlier studies in barley reported recombination rates ranging from 0.19 to 0.56 cM/Mbp on average genome-wide (Künzel et al., 2000; Phillips et al., 2015; Dreissig et al., 2020), 2.12 COs per meiosis (Higgins et al., 2012), and an accumulation in RILs that goes from 18.7 to 25.6 genome-wide (Dreissig et al., 2020). In sorghum, the recombination rate was reported to range from 1.8 to 4 cM/Mbp on average genome-wide (Mace et al., 2009; Bouchet et al., 2017). In non-herbaceous plant species, genome-wide recombination assessments ranged from 1.94 to 16 cM/Mbp in *Eucalyptus grandis* and *Populus tremula*, respectively (Petroli et al., 2012; Silva-Junior and Grattapaglia, 2015; Gion et al., 2016).

The above-mentioned difficulties in detecting GCs and, in particular, NCO-

related GCs, make such events to be poorly documented in plants. In *Arabidopsis*, different tetrad analyses reported combined NCO-GC and CO-GC rates to be between 1.1 and 8.8×10^{-5} per site per meiosis (Sun et al., 2012; Wijnker et al., 2013; Qi et al., 2014) while NCO-GC rates range from 1×10^{-6} to 2.5×10^{-5} per site per meiosis (Lu et al., 2012; Wijnker et al., 2013). Similarly, the GC rate in rice was reported to be 3×10^{-5} (Si et al., 2015). It is worth mentioning that the above-cited studies estimated NCOs at an unlikely low rate (NCO/CO ratios equal to 1 or even below) considering the number of detected COs and the expected number of DSBs per meiotic cell (above 100) (Sanchez-Moran et al., 2007; Sun et al., 2012; Wijnker et al., 2013; Si et al., 2015). Such discrepancy is typically attributed to the short length of NCO-related GC tracts (25-50 bp) and to higher occurrence probability of inter-sister chromatid repair compared with conversion (Sun et al., 2012; Lu et al., 2012; Drouaud et al., 2013; Wijnker et al., 2013; De Massy, 2013). In this respect, higher GC rates in *Arabidopsis* were reported when considering marker blocks shorter than 10 kb as GCs (Yang et al., 2012), however, such calculations were demonstrated to not be accurately validated (Qi et al., 2014). A strong validated similar approach in wheat revealed GCs to be more prevalent in wheat than in other studied species, being the rate of 4×10^{-3} in RILs (Gardiner et al., 2019).

In addition, for a given population, differences in recombination rate were observed among sexes (i.e., heterochiasmy). This phenomenon has been proposed to arose to suppress recombination in the heterogametic sex chromosomes –known as Haldane-Huxley rule (Haldane, 1922; Huxley, 1928)– based on early observations in *Drosophila* spp. (Morgan, 1912, 1914) and the silkworm (*Bombyx mori*) (Tanaka, 1913) which heterogametic sex lacks meiotic recombination entirely (i.e., achiasmy) (Ritz et al., 2017). In line with such observations, many studied animal species have shown a lower genome-wide recombination rate in the heterogametic sex compared with the homogametic sex (Dunn and Bennett, 1967; Ritz et al., 2017). However, on a finer scale, such differences were shown to be more complex, for example, the recombination rate in women is higher than in men in specific genomic regions (Kong et al., 2010). In plants, heterochiasmy was reported in several species such as *Arabidopsis* (Armstrong and Jones, 2001; Drouaud et al., 2006; Toyota et al., 2011; Giraut et al., 2011), maize (Rhoades, 1941; Robertson, 1984), wheat (Saintenac et al., 2009), barley (Devaux et al., 1995; Phillips et al., 2015), *Brassica* spp. (Lagercrantz and Lydiate, 1995; Kearsey et al., 1996), oriental garlic (*Allium tuberosum*) (Gohil and Kaul, 1981), and others (Fogwill, 1957; Korzun et al., 1996; Lenormand and Dutheil, 2005). It is worth noting that Haldane’s rule does not explain the observed heterochiasmy in plants as such do not have sex chromosomes (Phillips et al., 2015). Moreover, several plant species

showed a higher recombination rate in male meiosis (i.e., pollen formation) than in female meiosis (i.e., megaspore formation). For example, male meiosis recombines at a rate of 50% higher than such of female in all five chromosomes of *Arabidopsis* (Drouaud et al., 2006; Giraut et al., 2011; Toyota et al., 2011). Such observations were explained as the product of haploid selection by which female meiosis might recombine less than male meiosis when inbreeding increases the selection intensity on female meiosis or decreases the selection intensity over male gametophytes (Lenormand and Dutheil, 2005). Accordingly, in barley, a selfing species in the Triticeae tribe, the recombination rate in male meiosis was found to be higher than such of female, and those differences were increased at high temperatures (Devaux et al., 1995; Phillips et al., 2015), while in rye, the only out-crossing species in the same tribe, more recombination in the female meiosis compared to male was observed (Korzun et al., 1996).

2.6 Recombination rate variation along the genome

Across species, the distribution of recombination events along the chromosomes is neither homogeneous nor random at any observation scale (De Massy, 2013; Mercier et al., 2015; Melamed-Bessudo et al., 2016). On the broad scale (Mbp scale), the recombination occurs mostly in chromosome arms and is depleted in extensive regions (i.e., recombination deserts) (Gill et al., 1996; Akhunov et al., 2003; Wu et al., 2003; De Massy, 2013; Rodgers-Melnick et al., 2015). The centromeres and telomeres, in particular, are universally suppressed for crossover events across eukaryotes, although empirical evidence of recombination in centromeres mostly mediated by GCs does exist (Ma and Bennetzen, 2006; Shi et al., 2010; Fernandes et al., 2024). In addition, in several species, the depleted recombination rate in the centromere is also extended throughout the proximal region (i.e., pericentromeric region), determining most of the CO to occur in the distal region of the chromosomes (Mercier et al., 2015). Consequently, many earlier studies reported a positive correlation between the recombination rate and the distance from the centromere in various species (Akhunov et al., 2003; Wu et al., 2003; Anderson, 2004; Paape et al., 2012). Moreover, in several species such as most grasses, such pericentromeric region can be large, thus being recombination depleted in a long fraction of the chromosomes. That pattern was shown in maize (McMullen et al., 2009; Bauer et al., 2013; Rodgers-Melnick et al., 2015), rice (Si et al., 2015; Marand et al., 2019), wheat (Gill et al., 1996; Akhunov et al., 2003; Saintenac et al., 2009; Jordan et al., 2018; Gutierrez-Gonzalez et al., 2019; Gardiner et al., 2019), barley

(Pedersen et al., 1995; Künzel et al., 2000; Stephens et al., 2004; Higgins et al., 2012; Mayer et al., 2012; Dreissig et al., 2019, 2020), sorghum (Mace et al., 2009; Morris et al., 2013; Bouchet et al., 2017), and rye (Schreiber et al., 2022). Interestingly, GCs were shown to occur with high frequency in the pericentromeric regions of wheat (Gardiner et al., 2019), and even in the centromere cores of maize (Shi et al., 2010), suggesting that DSBs in these regions occur but NCOs prevail over COs (Gardiner et al., 2019). In contrast to grasses, CO rate was found to be possibly high near centromeres in most studied dicots such as *Arabidopsis* (Kim et al., 2007; Giraut et al., 2011; Yang et al., 2012; Choi et al., 2013; Rowan et al., 2019; Fernandes et al., 2024), *Medicago trunculata* (Branca et al., 2011; Paape et al., 2012), cucumber (*Cucumis sativus*) (Zhang et al., 2012), and Welsh onion (*Allium fistulosum*) (Jones, 1984), with some exceptions such as tomato (*Solanum lycopersicum*, *S. pennellii*, and *S. pimpinellifolium*) (Sim et al., 2012), and bananas (*Musa* spp.) (Martin et al., 2020).

On the fine scale (nucleotide to kb scale), broad chromosomal regions with low or even depleted recombination rate (i.e., coldspots) are patched with narrow sections (< 10 kb and even < 1 kb) that accumulate most of the recombination events (i.e., hotspots) and indicate the location of meiotic DSBs (i.e., COs and NCOs cluster around the DSBs sites in hotspot regions) (Henderson, 2012; Drouaud et al., 2013; Melamed-Bessudo et al., 2016). This pattern implies that more than 80 % of the recombination to be accumulated in less than 25 % of the genome in most eukaryotes (Mercier et al., 2015). In the literature, there is no unified criteria to define recombination hotspots which makes comparisons among studies difficult, but typically considered hotspots are such regions where the CO rate is higher than either the chromosome/genome average or the neighboring region, both depending on the genomic resolution being employed in the analysis (Choi and Henderson, 2015; Rowan et al., 2019). For example, studies in maize have shown the *bz* locus, which is 1.5 kb in length, to have a recombination rate 100 times higher than its surrounding region (Dooner and Martínez-Férez, 1997; Dooner, 2002; He and Dooner, 2009). Differently, by splitting the *Arabidopsis* genome in 50 kb genomic windows, the 5 % most recombining windows were found to have a mean recombination rate of 9.3 cM/Mb while the genome average window was 3.2 cM/Mb (Rowan et al., 2019). The occurrence of CO hotspots has been reported in several species, including, budding yeast (Mancera et al., 2009), the nematode *Caenorhabditis elegans* (Bernstein and Rockman, 2016), *Drosophila* spp. (Cirulli et al., 2007; Comeron et al., 2012), mice (Lindahl, 1991), and humans (McVean et al., 2004; Coop et al., 2008). In plants, hotspots have been reported in *Arabidopsis* (Drouaud et al., 2006; Kim et al., 2007; Horton et al., 2012; Yang et al., 2012; Drouaud et al., 2013; Wijnker et al., 2013; Basu-Roy et al., 2013;

Rowan et al., 2019), maize (Dooner, 1986; Brown and Sundaresan, 1991; Okagaki and Weil, 1997; Dooner and Martínez-Férez, 1997; Rodgers-Melnick et al., 2015), rice (Si et al., 2015; Marand et al., 2019), wheat (Saintenac et al., 2011; Darrier et al., 2017; Coulton et al., 2020), barley (Künzel et al., 2000), *Medicago truncatula* (Paape et al., 2012), peanut (*Arachis hypogaea*) (Wang et al., 2020), *Populus* spp (Slavov et al., 2012; Apuli et al., 2020), *Eucalyptus grandis* (Silva-Junior and Grattapaglia, 2015), and monkeyflower (Hellsten et al., 2013).

NCOs have been shown to have a much more even distribution along the chromosomes than do COs, thus representing the distribution of DSBs with higher fidelity than do such of COs. Previous studies across kingdoms have shown that the CO/NCO ratio varies across the genome (Mancera et al., 2009; Drouaud et al., 2013), e.g., in two different CO hotspots in *Arabidopsis*, the CO/NCO ratio was 1/1 and 30/1, respectively (Drouaud et al., 2013). As mentioned above, NCOs were detected in regions such as heterochromatin, centromeres, and telomeres where COs are scarce or even null (De Massy, 2013), e.g., a study among maize diverse inbreds revealed that GC rate in maize could be higher at centromeres, explaining part of the genetic diversity found at centromeres while no CO occur at them (Shi et al., 2010). In this respect, GC frequencies might vary depending on DSB distribution, CO rate-CO/NCO balance, CO interference, bias in restoration versus conversion by MMR, or through the indirect effects of sequence context or epigenetic influences (Henderson, 2012).

2.6.1 Genomic features associated with the recombination rate variation along the genome

The above-described distribution of recombination events is a complex interplay between various genomic features that defines the spatio-temporal progression of meiosis and recombination which ultimately define CO distribution (Wijnker et al., 2013; Higgins et al., 2014). Such complexity generates recombination rate and distribution variation among and within species (e.g., recombination hotspots were found to be cross-specific in *Arabidopsis* and maize), however, the mechanisms by which such features operate on the occurrence of recombination events were found to be highly conserved across species establishing general patterns and making the occurrence of recombination to be highly predictable along the most studied species (Rodgers-Melnick et al., 2015). In this way, most COs in plants occur in euchromatic regions where chromatin is accessible (hypomethylated DNA and nucleosome-free), associated with chromatin marks and sequence motifs that favor RNA polymerase II transcription, and close to gene promoters and to a lesser extent to gene terminators. Consequently, COs are typically suppressed in

the heterochromatin and repetitive regions as a product of the increased methylation, high nucleosome density, suppressed RNA Pol-II transcription, and late-replication (Henderson, 2012; Mercier et al., 2015). The most important findings on the association between the occurrence of COs and different genomic features are summarized below.

2.6.1.0.1 Epigenetic-related features and sequence motifs. The progress of prophase I is associated with programmed cycles of distinct chromatin contraction and expansion periods where nucleosome binding to DNA plays a central role (Kleckner et al., 2004; Daenen et al., 2008; Higgins et al., 2012). In this context, earlier studies in *Arabidopsis* have shown that COs tend to occur in nucleosome-depleted regions (NDRs, also called LNDs: low-nucleosome density regions) (Yelina et al., 2012; Drouaud et al., 2013; Choi et al., 2013; Wijnker et al., 2013), specially at gene promoters, in line with previous observations in mouse (Buard et al., 2009) and budding yeast (Pan et al., 2011). Nucleosomes can bind to any DNA sequence but the binding probability depends on the specific sequence and other factors such as epigenetic profiles (Daenen et al., 2008), e.g., COs in *Arabidopsis* were found to preferentially occur at nucleosome-free regions associated with Poly-A and GC rich motifs (Wijnker et al., 2013). In this way, several studies have shown that recombination sites are determined by sequence motifs and epigenetic marks targeting SPO11 activity, thus favoring DSB formation (Drouaud et al., 2013).

Different sequence motifs associated with open chromatin and cis-regulatory regions were found to be positively associated with recombination rate in plants: the Poly-A homopolymers located upstream of gene TSS sites, which are known to prevent nucleosomes from binding DNA (Horton et al., 2012; Wijnker et al., 2013; Choi et al., 2013; Shilo et al., 2015; Darrier et al., 2017; Choi et al., 2017; Rowan et al., 2019), GAA/CTT-repeats located immediately downstream of TSSs (Choi et al., 2013; Wijnker et al., 2013; Yelina et al., 2015; Rowan et al., 2019), CCN repeats (Shilo et al., 2015; Darrier et al., 2017; Rowan et al., 2019; Marand et al., 2019), AT-rich sequences known to exclude nucleosomes and allows SPO11 access (Slavov et al., 2012; Choi et al., 2017; Marand et al., 2019), and ATAC-seq sites (Rowan et al., 2019). The absence of such repeats in genomic regions bordering the centromeres has been proposed to partially explain the low occurrence of recombination in that region (Shilo et al., 2015). The positive association between the mentioned motifs and recombination was also found in non-plant organisms, including *Drosophila* spp. (Comeron et al., 2012), mouse (Baudat et al., 2010; Smagulova et al., 2011), primates (Myers et al., 2010), and humans (Kong et al., 2010; Baudat et al., 2010).

In addition, recombination sites in plants were associated with regions enriched in histone H3 trimethylated on lysine 4 (H3K4), a histone modification associated with active transcription that occurs at the 5' end of genes (Zilberman et al., 2007; Henderson, 2012; De Massy, 2013). This was observed in *Arabidopsis* (Drouaud2013), maize (Liu2009), rice (Choi et al., 2013; Marand et al., 2019), and barley (Higgins et al., 2012), as observed in other non-plant species like budding yeast (Borde et al., 2009; Sommermeyer et al., 2013), mouse (Buard et al., 2009; Baudat et al., 2010; Grey et al., 2011), and humans (Baudat et al., 2010; Acquaviva et al., 2013). Another histone variant found to promote recombination hotspots in *Arabidopsis* was H2A.Z which is known to facilitate transcriptional regulation at TSSs (Choi et al., 2013). Furthermore, DNase I hypersensitive sites (DHSs), an indicator of chromatin accessibility was found to be positively associated with COs in *Arabidopsis* (Rowan et al., 2019), maize (Rodgers-Melnick et al., 2015), rice (Marand et al., 2019), and potato (*Solanum tuberosum*) (Marand et al., 2017). On the reverse, epigenetic marks repressing recombination have been reported in previous studies. For example, elevated histone H3 acetylation was associated with defects in meiosis and changes in CO number (Perrella et al., 2010). Moreover, heterochromatin is enriched with histones such as the H3 methylated at lysine 9 (H3K9me, me2, me3) which were observed to repress CO in *Arabidopsis* (Yelina et al., 2015) as first observed in *Caenorhabditis elegans* (Reddy and Villeneuve, 2004). Similarly, CO hotspots in rice were shown to be depleted of H3K36me3, a modified histone negatively correlated with DSB formation in yeast (Marand et al., 2019). The direct role on recombination of many other histone modifications known to repress RNA polymerase II remains to be confirmed (Yelina et al., 2012).

Besides histone modifications, the methylation of the DNA nucleotides per se, typically cytosines, is an indicator of repressed and packed chromatin state which implies the inhibition of RNA Pol II transcriptional initiation (Choi et al., 2013). For example, methylation of heterochromatic DNA is responsible for transcriptional silencing of repeats and transposons that are extensive in the centromeric and pericentromeric regions along with the stabilization of the chromatin compaction needed for chromosome pairing and synapsis formation in meiosis (Reinders and Paszkowski, 2009; Mirouze et al., 2012). In this way, previous studies in *Arabidopsis* have shown a strong association between repressed recombination rate and methylated cytosines (Salomé et al., 2012; Colomé-Tatché et al., 2012; Yelina et al., 2012; Mirouze et al., 2012; Melamed-Bessudo and Levy, 2012; Choi et al., 2013; Wijnker et al., 2013; Yelina et al., 2015; Ziolkowski et al., 2017) similar to firstly observed in the fungus *Ascobolus immersus* (Maloisel and Rossignol, 1998). Similar patterns were observed in other plant species such as maize (He

and Dooner, 2009; Rodgers-Melnick et al., 2015), rice (Marand et al., 2017, 2019), *Brassica* spp. (Boideau et al., 2022), and *Populus* spp. (Slavov et al., 2012; Apuli et al., 2020). The hypothesis behind such observation include an increase in the replication time length, decreased accessibility of the recombination machinery (including the direct inhibition of SPO11 access), and decreased processing of recombination intermediates (Henderson, 2012). In plants, depending on the neighboring nucleotides, methylated cytosines can be classified as CpG and non-CG sequence contexts, being the last also divided into CHG and CHH methylation contexts (Law and Jacobsen, 2010). In most studied plants, the strong association between recombination and cytosine methylation was found to be mostly driven by the CpG and CHG methylated sequence contexts while little or no association was found linked to CHH sites (Rodgers-Melnick et al., 2015). Interestingly, the methylation at CHH contexts was found positively correlated with CO hotspots in rice (Marand et al., 2019; Peñuela et al., 2022). Moreover, in maize, it was suggested that increased recombination was associated with increased CHH methylation in regions with high CpG-related methylation levels, but with decreased recombination where CpG methylation was low (Rodgers-Melnick et al., 2015).

DNA cytosine methylation is an epigenetic modification that can occur spontaneously (Becker et al., 2011) as well as be heritably maintained through DNA replication (Yelina et al., 2012). Mutants for genes related to the methylation in plants such as *MET1* and the *DDM1* have been used to evaluate the effect of methylation on recombination (Melamed-Bessudo and Levy, 2012; Mirouze et al., 2012; Colomé-Tatché et al., 2012; Yelina et al., 2012). In most studies, de-methylated mutants showed increased local recombination in the euchromatin (chromosome arms), but recombination suppression in the heterochromatic pericentromeric regions was maintained. This suggests that cytosine methylation per se does not determine the level of meiotic recombination in heterochromatin (Mirouze et al., 2012; Melamed-Bessudo and Levy, 2012; Colomé-Tatché et al., 2012). Moreover, mutants did not alter the total number of COs genome-wide, but rather CO frequency was remodeled along the chromosomes (Melamed-Bessudo and Levy, 2012; Mirouze et al., 2012; Yelina et al., 2012; Colomé-Tatché et al., 2012). Mechanisms such as homeostasis, interference, and specific states of chromatin modification were proposed to be responsible for the observed patterns (Melamed-Bessudo and Levy, 2012; Yelina et al., 2012).

2.6.1.0.2 Proximity to genes. On the broad genomic scale, although a few studies alleged a negative correlation between recombination and gene density (Kim et al., 2007; Giraut et al., 2011; Yang et al., 2012; Rodgers-Melnick et al., 2015), most studies in plants reported a positive correlation between recombination

and gene density given that most recombination events occur in the sub-telomeric regions where most genes are located. This was observed in several species such as *Arabidopsis* (Yelina et al., 2012; Choi et al., 2013; Wijnker et al., 2013), maize (Anderson, 2004; Li et al., 2015), rice (Marand et al., 2019), wheat (Gill et al., 1996; Saintenac et al., 2009, 2011; Darrier et al., 2017; Gardiner et al., 2019), barley (Künzel et al., 2000; Dreissig et al., 2020), sorghum (Bouchet et al., 2017), *Medicago trunculata* (Paape et al., 2012), *Eucalyptus* spp. (Silva-Junior and Gratapaglia, 2015; Gion et al., 2016), and *Populus tremula* (Apuli et al., 2020). In addition, because in several species, such as wheat, genomic regions with high recombination tend to carry genes that are likely to be expressed in meiosis (Darrier et al., 2017), it has been proposed that the location of recombination events is driven by the recruitment of the recombination machinery (Henderson, 2012). However, the distance among recombination hotspots and genes may vary among eukaryotes. For example, the *Arabidopsis* genome is similar to that of budding yeast in which COs occur in close vicinity to genes but different to such of fission yeast (*Schizosaccharomyces pombe*) in which most COs occur away of genes (Buhler et al., 2007; Cromie et al., 2007; Pan et al., 2011; Mercier et al., 2015). Part of such variation has been attributed to differentials in genomic structures among species, e.g., *Arabidopsis* and budding yeast have both a compact genome producing a high gene density and short intergenic regions (Drouaud et al., 2013). In this context, it was observed in *Arabidopsis* that the recombination rate in a given intergenic region depends on the size of the intergenic segment as it was suppressed for intergenic regions whose size is less than 1.5 kb (Hsu et al., 2022).

At a finer scale, the location of recombination events towards genes is more complex and depends on the interplay of the genomic features mentioned addressed in this section. For example, in *Arabidopsis*, CO frequency around genes was found to follow the inverse distribution gradient of DNA methylation which is low at transcription starting and termination sites (TSSs and TTSs), increasing along gene bodies with the distance from the TSS, and being higher in exons compared to introns (Choi et al., 2013). In this way, gene body methylation has been proposed as a protector of genome integrity by inhibiting supernumerary COs within exons (Mirouze et al., 2012), determining CO location to be mostly in intergenic regions next to the promoters or terminators of genes but depleted in gene bodies as observed in several species including *Arabidopsis* (Horton et al., 2012; Yelina et al., 2012; Choi et al., 2013; Drouaud et al., 2013; Wijnker et al., 2013; Sun et al., 2019), maize (Yao et al., 2002; He and Dooner, 2009; Li et al., 2015), rice (Marand et al., 2019), *Populus trichocarpa* (Slavov et al., 2012), and monkeyflower (Hellsten et al., 2013). Another example of this complexity was described in the mouse genome in which meiotic DSBs are positioned close to

intergenic H3K4 marks away from gene promoters guided by the PRDM9 zinc finger while in the absence of PRDM9 hot spots position revert close to the TSSs (Smagulova et al., 2011; Brick et al., 2012).

Importantly, while recombination in plants occurs mostly outside of genes, COs may still occur inside gene bodies leading to the formation of new allelic variants of genes (Wijnker et al., 2013; Bouchet et al., 2017). This might facilitate adaptive evolution by increasing the mutational target size, modifying the genes copy-number, or generating chimerical genes, thus modulating, turning off, or creating new genetic functionalities (Horton et al., 2012). In this respect, COs were shown to occur in loci involved in different adaptive processes in plants such as disease resistance (R-genes) -in species such as *Arabidopsis* (Horton et al., 2012; Choi et al., 2016; Rowan et al., 2019), wheat (Wicker et al., 2007), rice (Wicker et al., 2007; Si et al., 2015), barley (Mayer et al., 2012), and *Medicago trunculata* (Paape et al., 2012)-, and the response to environmental stimuli in peanuts (*Arachis hypogaea*) (Wang et al., 2020), in line with observations on human immune loci (McVean et al., 2004; Winckler et al., 2005).

2.6.1.0.3 Sequence divergence and structural variations. The relation between polymorphisms and recombination rate is complex and studies in plants reported contrasting results (Henderson, 2012). A few studies in plants reported negative or non association between CO rate and sequence diversity in species such as *Arabidopsis* (Yelina et al., 2012; Wijnker et al., 2013; Choi et al., 2016; Serra et al., 2018), maize (Bauer et al., 2013), wheat (Saintenac et al., 2011; Jordan et al., 2018; Gutierrez-Gonzalez et al., 2019), sorghum (Bouchet et al., 2017), and *Eucalyptus* spp. (Gion et al., 2016). Because sequence homology is required for homolog identification and pairing during recombination, sequence polymorphisms that generate base pair mismatches, may suppress CO due to defective strand invasion and homology pairing, thus favoring NCOs at the expense of COs as observed in budding yeast (Borts et al., 2000; Henderson, 2012). In humans and budding yeast, inhibition of recombination via sequence mismatches is dependent on MutS homolog 2 (MSH2), a DNA mismatch repair (MMR) protein that binds mismatches in D-loop structures and generates its disassembly (Evans et al., 2000; Borts et al., 2000). In *Arabidopsis*, however, MSH2 has not been found related to recombination suppression but with COs redistribution in the pericentromeric regions which have relatively high sequence divergence (Blackwell et al., 2020). In addition, from an evolutionary standpoint, it can be hypothesized that the increased linkage break generated by recombination would decrease allelic diversity as a by-product of the selection against detrimental alleles (Charlesworth et al., 1993; Rodgers-Melnick et al., 2015; Apuli et al., 2020). That would imply

the reduction of deleterious polymorphisms in recombination hotspots and the accumulation of deleterious mutations in areas of low recombination rates due to extensive linkage disequilibrium as reported in maize and wheat (Rodgers-Melnick et al., 2015; Jordan et al., 2018; Gutierrez-Gonzalez et al., 2019).

However, it has also been hypothesized from an evolutionary standpoint that increased linkage break generated by recombination would conserve diversity because of the fewer sites being dragged in the vicinity of positively selected sites (Begun and Aquadro, 1992). This is supported by the observed positive correlation between local recombination rates and neutral genetic diversity across a wide range of organisms (Nachman, 2002; Wang et al., 2016; Apuli et al., 2020). Moreover, genomic regions with a high rate of DSBs are expected to have a historically increased mutation rate produced by COs and NCOs which elevated the allelic diversity at such regions among genotypes as demonstrated in humans (Arbeithuber et al., 2015; Halldorsson et al., 2019) as well as in *Arabidopsis* (Choi et al., 2016; Wijnker et al., 2013; Rowan et al., 2019). In relation to that, due to observations in *Arabidopsis*, it has been proposed that recombination would have a bias towards variable regions in the genome that maximize the diversifying effects of meiosis (Blackwell et al., 2020), e.g., large heterozygous regions attract crossovers at the expense of linked homozygous regions when both are adjacent (i.e., heterozygous block effect) (Barth et al., 2001; Ziolkowski et al., 2015; Blackwell et al., 2020; Hsu et al., 2022). Accordingly, despite the inhibitory effect of sequence divergence on recombination, different studies in plants found a positive association between recombination and both nucleotide diversity and genetic divergence among homologous chromosomes. Such has been observed in *Arabidopsis* (Yang et al., 2012; Rowan et al., 2019; Blackwell et al., 2020), rice (Si et al., 2015; Marand et al., 2019), barley (Dreissig et al., 2020), *Medicago trunculata* (Paape et al., 2012), and *Eucalyptus grandis* (Silva-Junior and Grattapaglia, 2015).

The contradiction might be explained by a parabolic relationship between both variables meaning that recombination has a positive correlation with genetic divergence until a level after which the high polymorphism among homologs suppresses COs due to the increase of mismatches as recently reported in *Arabidopsis* (Blackwell et al., 2020; Hsu et al., 2022). This explanation is in line not only with the above-mentioned studies reporting crossover suppression at high SNP densities, but also agrees with a long-standing observation across species that structural variations suppress COs (proposed by Sturtevant, 1921) which was first observed by Dobzhansky (1931) in *Drosophila melanogaster* as well as in other model organisms later such as *Caenorhabditis elegans* (McKim et al., 1988; Hammarlund et al., 2005), mouse (Hsu et al., 2000), and humans (Jeffreys and Neumann, 2005). Accordingly, in plants, the presence of structural variants (SVs) - insertions and

deletions (indels), translocations, and inversions- was shown to suppress COs in previous studies in *Chlamydomonas reinhardtii* (Ferris and Goodenough, 1994), *Arabidopsis* (Salomé et al., 2012; Fransz et al., 2016; Rowan et al., 2019), maize (Dooner, 1986; He and Dooner, 2009; Rodgers-Melnick et al., 2015), rice (Shen et al., 2019), barley (Jayakodi et al., 2020), *Brassica* spp. (Boyes et al., 1997; Boideau et al., 2022), tomato (Herickhoff et al., 1993; Parniske et al., 1997), bananas (Martin et al., 2020), and *Eucalyptus* spp. (Gion et al., 2016). Importantly, in *Arabidopsis*, DSB formation was shown to not be affected by the presence of SVs, indicating that the SV’s suppressing effect on COs is likely to occur downstream of DSB formation (Rowan et al., 2019). Also interesting, in *Arabidopsis*, the type and size of the SVs were not found to be related to such effect, and CO suppression was mostly comprised within the SV borders (Rowan et al., 2019). In addition, a few studies in *Arabidopsis* (Rowan et al., 2019) and rice (Marand et al., 2019) showed that CO rates were slightly elevated in the flanking regions of the SV, a phenomenon first observed in rye (Sybenga, 1970). Possible explanations for this phenomenon are the CO re-distribution from the region spanned by SVs to the SV-flanking regions and the high occurrence of SVs in regions of high recombination rate resulting from the historical generation of chromosomal rearrangements by recombination (Rowan et al., 2019).

2.6.1.0.4 Repetitive regions. About the mutagenic effect of recombination, COs are suppressed in repetitive regions which are extensive in the genome of several plant species (Fu et al., 2001; Wu et al., 2003; Saintenac et al., 2009; Apuli et al., 2020). For example, in the regions surrounding the centromeres in *Arabidopsis* (Giraut et al., 2011; Salomé et al., 2012). From an evolutionary standpoint, this would help prevent the generation of detrimental recombination products such as non-allelic COs, tandem repeat duplications, or deletion by slippage, thus ensuring the proper repair of DSBs while minimizing the generation of deleterious events that might compromise genome integrity (Henderson, 2012; De Massy, 2013). The repressed recombination over repetitive regions was proposed to be due to the combined presence of heterochromatin and structural polymorphism in such regions as these generate small RNAs (siRNAs) that cause RNA-directed DNA methylation (RdDM) as well as are known to harbor extensive sequence polymorphisms among plant genotypes (Henderson, 2012).

2.6.1.0.5 Transposable elements (TEs). On the broad scale, TEs were not found to be associated with COs in plants as most are located in repeats-rich heterochromatin regions (Wijnker et al., 2013; Shilo et al., 2015). However, several studies revealed associations of TEs with local recombination rate variation in

plants with variable effect depending on the TE family (Shalev and Levy, 1997; Yao and Schnable, 2005; Yandea-Nelson et al., 2006; He and Dooner, 2009; Horton et al., 2012; Sun et al., 2012; Choi et al., 2017; Sun et al., 2019). For example, Mutator like transposons were shown to be enriched in CO hotspots in *Arabidopsis* (Horton et al., 2012; Choi et al., 2017; Sun et al., 2019) and maize (Yandea-Nelson et al., 2005, 2006), while LTR retrotransposons (such as Gypsy-like) were found to be associated with decreased CO rates in *Arabidopsis* (Horton et al., 2012; Choi et al., 2017; Sun et al., 2019), maize (He and Dooner, 2009), wheat (Darrier et al., 2017; Gutierrez-Gonzalez et al., 2019), barley (Wei et al., 1999, 2002), and *Populus trichocarpa* (Slavov et al., 2012). However, contrasting effects were also found, e.g., Helitrons have been found positively associated with recombination in *Arabidopsis* (Choi et al., 2017; Sun et al., 2019) but negatively in maize (He and Dooner, 2009). Interestingly, it has been proposed that transpose via a cut-and-paste mechanism might generate recombination through increased DSBs and the direction of the effect of a given TE on recombination might be dependent on the differential association of CO-associated epigenetic marks with such TE (Melamed-Bessudo et al., 2016).

2.6.1.0.6 GC-bias. The GC-content, a sequence feature non-related to gene expression, was also shown to be positively associated with recombination rate in studies across eukaryotes, presumably as a product of mismatch repairs of divergent homologs biased towards G/C over A/T (Duret and Galtier, 2009; Pan et al., 2011; Pessia et al., 2012; Glémin et al., 2014). In plants, contrasting results have been reported in *Arabidopsis* (Kim et al., 2007; Giraut et al., 2011; Wijnker et al., 2013) and a weak negative correlation was found in *Medicago trunculata* (Paape et al., 2012), while in grasses, a positive correlation has been observed in maize and rice (Rodgers-Melnick et al., 2015; Marand et al., 2019), but no correlation has been found in wheat (Saintenac et al., 2011). Similarly, in non-herbaceous species, the GC content was found positively correlated with the recombination rate across *Eucalyptus* spp. species (Silva-Junior and Grattapaglia, 2015; Gion et al., 2016), but lowly correlated with recombination hotspots in *Populus trichocarpa* (Slavov et al., 2012).

2.6.2 Genetic control of the recombination rate variation

Firstly observed in *Hypochoeri* spp. (Parker, 1975), genomic factors generating differences in recombination frequencies can be classified in trans-acting factors that affect recombination genome-wide and cis-acting factors that affect recombination locally, e.g., SVs and sequence motifs are considered cis-acting recombina-

tion controllers (for a review see Lawrence et al. 2017). Furthermore, in addition to the interplay of the above-discussed genomic features, the recombination rate is genetically controlled [for review read (Lawrence et al., 2017)]. In this way, since the recombination rate has been genetically analyzed as any other quantitative trait (Rasmusson, 1927; Esch et al., 2007; Dole and Weber, 2007), several studies have reported the identification of QTLs linked to recombination modifiers in natural populations of several plant species such as *Arabidopsis* (Esch et al., 2007; Ziolkowski et al., 2017), maize (McMullen et al., 2009), wheat (Jordan et al., 2018; Gutierrez-Gonzalez et al., 2019), barley (Dreissig et al., 2020), and *Brassica napus* (Higgins et al., 2020). In total, more than 80 genes have been identified and characterized that are involved in recombination rate variation in plants (Mercier et al., 2015; Lawrence et al., 2017; Gardiner et al., 2019). Classic examples of cis-genetic recombination modulators in plants were found in maize in which polymorphisms in both the *bz* gene and in the *a1-sh2* interval were found to modulate recombination variation at individual neighboring CO hotspots (Dooner, 2002; Yao and Schnable, 2005). In addition, several trans-acting genetic factors have been found in plant species, those mostly related to the CO and NCO generation pathways. The HEI10 meiotic E3 ligase that promotes CO in the ZMM-interfering pathway throughout eukaryotes (Kong et al., 2008; Fledel-Alon et al., 2011; Chelysheva et al., 2012; Sandor et al., 2012) was also found to be a major recombination QTL for CO natural variation in *Arabidopsis*, rice, and wheat (Ziolkowski et al., 2017; Wang et al., 2012; Jordan et al., 2018). Similarly, anti-CO factors such as AtFANCM, the AAA-ATPase FIDGETIN-LIKE-1 (FIGL1), and subunits of the BTR complex (RECQ4 -TOP3 α -RMI1) were found to be genome-wide CO modulators in plants (Crismani et al., 2012; Girard et al., 2014, 2015; Séguéla-Arnaud et al., 2015, 2017; Jordan et al., 2018).

Importantly, many trans-acting genetic factors are not conserved across plant species (Mercier et al., 2015), e.g., a particular ATP-dependent RNA helicase RecQ gene found to be a controller of the GC frequency in wheat is absent in *Arabidopsis* (Gardiner et al., 2019). Furthermore, genetic factors explaining a large fraction of the recombination rate variability among samples were found in individual plant populations (Esch et al., 2007; Ziolkowski et al., 2017; Gardiner et al., 2019). Moreover, studies that analyzed multiple plant populations detected controllers of recombination to have a strong effect in individual families rather than across families (McMullen et al., 2009; Jordan et al., 2018; Dreissig et al., 2020). For example, the cohesin subunit REC8 is a trans-acting recombination rate modifier in barley that explained 42% of the effect size of a given allele, but only 3% of the recombination variance across the analyzed populations in the study (Dreissig et al., 2020). In addition, most of the variation in recombination rate among

populations in maize, wheat, and other species was not explained by individual loci of large effects but rather by multiple loci of small effects (McMullen et al., 2009; Jordan et al., 2018; Gutierrez-Gonzalez et al., 2019).

2.7 Recombination in plant breeding

2.7.1 Importance of recombination in plant breeding

Among the fundamental evolutionary forces, recombination has the largest impact on plant breeding (Rodgers-Melnick et al., 2015). In the first place, recombination defines the efficiency of selection in breeding programs. The genetic variation in the progeny depends on the amount of recombination before selection. Thus, breeders typically look for high recombination at the beginning of the breeding cycle to increase the genetic variation and thus the possibilities for selection, e.g., high diversity among F2 individuals from a common F1 (Li and Pfeiffer, 2009; Mieulet et al., 2018). Therefore, highly recombinant parental genotypes could have the potential to accelerate selection and decrease the need for very large populations (Mirouze et al., 2012). In contrast, low recombination is desired later in breeding cycles to accelerate the fixation of favorable alleles in the offspring of selected individuals (i.e., decrease allele segregation) to then be propagated to an industrial scale (Crismani et al., 2013; Mieulet et al., 2018), e.g., selected F2 family individuals undergo several generations of selection and selfing until the remaining genotypes become mostly homozygous.

In second place, recombination is a key factor determining the precision of favorable allele introgression into the elite genetic material of breeding programs because new genetic variation –typically exotic alleles obtained from poorly adapted or distantly genetic-related material– is often accompanied by undesired linkage drag (Bauer et al., 2013; Gardiner et al., 2019). For example, the offspring from a cross between an elite cultivar with a disease-resistant poor-yield-performer genetic background needs to be backcrossed with an elite cultivar several times to remove the low-performing alleles, thus increasing COs at meiosis can accelerate plant breeding by reducing linkage drag (Bauer et al., 2013; Gardiner et al., 2019).

Also important, recombination determines the resolution power of marker-based analysis such as quantitative trait mapping (e.g., QTL analysis and GWAS), marker assisted selection (MAS), genomic selection (GS), and positional cloning. On the one hand, recombination generates allelic combinations that increase the QTL mapping resolution and the accuracy of the marker ordering needed in the construction of genetic maps used to map such QTLs. On the other hand, QTL mapping, as any marker-based approach such as MAS and GS, relies on high LD

between the utilized markers and the loci conforming at least partly the genetic architecture of the trait. Recombination may break such LD, decreasing the mapping resolution or the prediction accuracy of marker effects and breeding values for a given set of markers (i.e., regions with high recombination rates require higher marker densities) (Bauer et al., 2013).

It is worth noting that the low to null recombination in extensive regions of plant genomes results in a substantial loss of breeding efficiency (Bauer et al., 2013; Jordan et al., 2018). In particular, this applies to the long low-recombining pericentromeric regions in cereal genomes that contain over 20% of functional genes and that accumulate deleterious mutations, being the increase in recombination in such regions for making them accessible for selection a long-standing goal in plant breeding (Mayer et al., 2012; Bauer et al., 2013; Rodgers-Melnick et al., 2015; Darrier et al., 2017).

Based on the above-mentioned factors, it is clear that the effective application of different plant breeding methods –such as crossing strategies, QTL mapping, and marker-based predictions– requires the understanding of the recombination landscape at the species or even at the population level (Bauer et al., 2013). Moreover, the manipulation of the recombination rate and distribution in plants has the potential to accelerate plant breeding by fixing desirable haplotypes in fewer generations and smaller populations, being such a long-standing goal for breeders (Choi and Henderson, 2015; Taagen et al., 2020). In this way, the better understanding of the meiotic process in recent years has allowed the emergence of applied methods for the effective manipulation of the recombination rate and distribution in the genome of plants (i.e., controlled recombination) (Crismani et al., 2013; Mercier et al., 2015; Taagen et al., 2020). However, their practical applications in crops have shown slow progress due to several factors including genomic differences in recombination control among the model species and specific crops (e.g., ploidy level), difficulties to introduce new technologies in actual breeding programs, and government regulation constraints (Wang et al., 2015; Jordan et al., 2018; Taagen et al., 2020).

2.7.2 Manipulation of the recombination rate and distribution

2.7.2.0.1 Leveraging natural genetic variation. Primarily, the recombination rate in a population may be modified by taking advantage of the natural genetic variation of the species, e.g., choosing highly-recombining recurrent back-cross parents may speed up the introgression process (Bauer et al., 2013). In this respect, the existence of intra-population variation for recombination has been

reported in crop species such as the above-mentioned recombination variation in several maize populations (Fatmi et al., 1993; Hadad et al., 1996; Li and Pfeiffer, 2009; McMullen et al., 2009; Bauer et al., 2013; Rodgers-Melnick et al., 2015). Moreover, selection for increased and decreased recombination has been probed to modify the recombination frequency on different regions and chromosomes in maize and wheat (Li and Pfeiffer, 2009; Gardiner et al., 2019), in line with early reports in insects (Kidwell, 1972; Lander and Ryman, 1974).

Furthermore, the optimization of parental selection to modify the recombination rate in breeding programs might be driven by the identification of genotypes carrying specific alleles of the above-mentioned recombination rate modifiers genes (i.e., cis- and trans-acting CO modulators). Interestingly, a trans-acting QTL related to the recombination rate variation in the low-recombining pericentromere was reported in wheat (Jordan et al., 2018). In addition, because most DSBs are resolved as NCOs in pericentromeric regions, altering GC rates might be a route to increase recombination in such regions. In wheat, the *RECQ7* gene appears to be a pro-recombination factor that controls GCs, thus increasing its expression might be a route to generate GCs at targeted locations or along the chromosomes (Gardiner et al., 2019).

Similarly, natural epigenetic diversity also contributes to recombination variation, thus selection for epigenetic variants may contribute to alter recombination. In this way, different studies utilized methylation mutants as a tool to modify recombination in the *Arabidopsis* genome, either genome-wide or in specific regions (Melamed-Bessudo and Levy, 2012; Mirouze et al., 2012; Colomé-Tatché et al., 2012). However, such studies suggested an upper limit for the amount of demethylation in genomes (Colomé-Tatché et al., 2012). In addition, a study with *Arabidopsis* epiRILs indicated that the magnitude of the effect of hypomethylation on recombination may decline over generations (Mirouze et al., 2012).

2.7.2.0.2 Artificial introduction of genetic variation to the plant genome. Altered CO frequency by the repression or over-expression of genes related to CO factors -such as ZEP1, HEI10, RECQ4, FANCM, and FIGL1- has been extensively reported in *Arabidopsis* (Ziolkowski et al., 2017; Fernandes et al., 2018b; Serra et al., 2018b) but also in crop species (Wang et al., 2015; Mieulet et al., 2018). For example, the knockdown of RECQ4 produced over three-fold increase of COs in *Arabidopsis*, rice, pea (*Pisum sativum*), and tomato (Mieulet et al., 2018) while knocking down FANCM was shown to elevate CO frequency in rice and pea but not in *Arabidopsis* (Fernandes et al., 2018a; Serra et al., 2018b; Mieulet et al., 2018). A limitation to such application is that anti-CO genes have been shown to have pleiotropic effects (Hu et al., 2017) or to be essential for

normal meiotic recombination. For example, knock down of FIGL1 altered fertility in rice and pea (Zhang et al., 2017; Mieulet et al., 2018). In this line, the *Pairing homoeologous 1* (*Ph1*) locus in hexaploid wheat prevents recombination between homeologous (i.e., non-homologous) chromosomes, producing a diploid-like behavior at meiosis (Griffiths et al., 2006) and its deletion in wheat allowed some recombination between non-homologous regions in hybridizations with rye (Anugrahwati et al., 2008; Rey et al., 2018). However, *Ph1* is essential for fertility and, thus, its function must be restored after obtaining recombinants (Crismani et al., 2013).

Moreover, the genome-wide manipulation of epigenetic marks might be also a promising area to expand CO distribution (Corem et al., 2018; Taagen et al., 2020). In *Arabidopsis*, *ddm1* mutants induced demethylation genome-wide but increased recombination in euchromatic regions while not in the pericentromeric regions (Melamed-Bessudo and Levy, 2012). In contrast, the combined loss of H3K9me2 and non-CG DNA methylation did induce meiotic recombination near centromeres (Underwood et al., 2018). It is important to consider that the sensitivity to genome-wide modifications of methylation patterns that varies across plant species appears as a limitation for different techniques across species, e.g., DDM1 mutants in maize, rice, and tomato resulted in developmental abnormalities (Corem et al., 2018).

In addition to classical methods to generate plant mutants such as EMS or conventional plant transformations, precise methods such as TILLING (Targeting Induced Local Lesions IN Genome), zinc fingers proteins (ZFs), transcription activator-like effector nucleases (TALENs), and CRISPR-Cas9, have been used to generate recombinant mutants (Mieulet et al., 2018; Rey et al., 2018). Importantly, genome editing tools not only have the potential to perturb pro- or anti-CO pathway genes but also to precisely target DSBs or epigenetic modifications to sites of desired recombination (Taagen et al., 2020). In this way, DSBs have been successfully targeted in specific regions of the yeast's genome by means of ZFs, TALENs, and CRISPR-Cas9 (Sadhu et al., 2016; Sarno et al., 2017). Remarkably, such techniques allowed to increase the CO rate by tethering SPO11 protein at low-recombining regions in yeast (Sarno et al., 2017), being a viable strategy to increase recombination in heterochromatic regions (Taagen et al., 2020). In plants, CRISPR-Cas9 has been used to induce DSBs in wild tomato (*Solanum pimpinellifolium*) (Hayut et al., 2017). Furthermore, genome editing techniques such as CRISPR/Cas9 have been shown to successfully modify gene activity by targeting epigenetic modifiers (Vojta et al., 2016), e.g., in tomato, short-interfering RNAs (siRNAs) mediated methylation has been altered by CRISPR/Cas9-induced DMM1 mutants (Corem et al., 2018).

Moreover, manipulating the karyotype composition might be applied to increase CO frequency, e.g., polyploids have been associated with an increase in the recombination rate compared to diploid relatives in *Arabidopsis* (Pecinka et al., 2011) and *Brassica* spp. (Leflon et al., 2010; Boideau et al., 2021). Remarkably, *Brassica* triploids were found to be more efficient than diploid or allotetraploid hybrids to introduce recombination in low-recombining regions, increasing the breeding efficiency for agronomic traits (Boideau et al., 2021). Also interesting, re-synthesised *Brassica napus* lines, created by crossing *Brassica oleracea* with *Brassica rapa* and doubling the chromosome complement, showed higher homoeologous recombination rates compared to established lines (Higgins et al., 2020).

2.7.2.0.3 Environmental manipulation of the recombination rate.

Temperature has been first reported to alter chiasma frequency in *Drosophila* spp. (Plough, 1917) while in plants in *Vicia faba* (Maeda, 1930). Later experiments in ornamental flowering plants such as *Endymion non-scriptus*, *Hyacinthus orientalis*, *Tradescantia bracteata*, and *Uvularia perfoliat* confirmed that chiasma frequency in plants increases with temperature being maximized at given species-specific temperature after which higher temperatures decreases COs towards zero (Elliott, 1955; Dowrick, 1957). This was further confirmed in *Arabidopsis* (Francis et al., 2007; Modliszewski et al., 2018; Lloyd et al., 2018) and major crops species such as rice (Si et al., 2015), wheat (Coulton et al., 2020), and barley (Higgins et al., 2012), in which temperature variation typically modifies recombination rate by 10–30% (Mieulet et al., 2018). Remarkably, studies in barley reported an increase of the CO rate in the low-recombining pericentromeric region of the chromosomes by moderate increments in growth temperature (Higgins et al., 2012; Phillips et al., 2015). In addition, in *Arabidopsis*, it was reported an increase of the CO rate at temperatures below the moderate temperature range that previous studies had investigated (Lloyd et al., 2018). The effect of temperature in the rate and distribution of COs has been associated with the chromatin unwrapping cycles that define the meiotic progression along the chromosomes while the suppression of recombination at extreme temperatures is related to structural disruptions of the chromosomes (Kumar and Wigge, 2010; Higgins et al., 2012).

In addition, several other abiotic and biotic stress factors such as UV-light, high salinity, and pathogen attack were shown to increase homologous recombination in plants such as *Arabidopsis*, tobacco (*Nicotiana tabacum*), and rice (Puchta et al., 1995; Kovalchuk et al., 2003; Lucht et al., 2002; Francis et al., 2007; Si et al., 2015), as shown previously in non-plant organisms (Simchen and Stamberg, 1969). Moreover, the recombination rate in plants was shown to be altered artificially by the application of chemical agents such as methyl methanesulfonata (MMS), ethyl

methane sulfonate (EMS), and N-nitroso-N-methyl urea (NMU) (Jayabalan and Rao, 1987; Puchta et al., 1995). Interestingly, it has been proposed that stress-induced recombination might be a mechanism fueling evolution (Zhong and Priest, 2011), and accordingly, the historical recombination variation among domesticated landraces and wild barley accessions was shown to be shaped by environmental conditions during domestication such as temperature, solar radiation, and precipitation (Dreissig et al., 2020).

2.7.2.0.4 Genome-wide recombination suppression. When generating lines in breeding programs, recombination needs to be avoided after selection to generate homozygous plants which is typically done by selfing over several generations. A shortcut for those many generations can be done by generating double-haploids (DHs), firstly demonstrated in jimsonweed (*Datura stramonium*) (Blakeslee et al., 1922), which can be done in vitro by gynogenesis or androgenesis, or in vivo by parthenogenesis, pseudogamy, or chromosome elimination after wide crossing (Barnabás et al., 1999; Sood et al., 2013; Ahmadi and Ebrahimzadeh, 2020; Seguí-Simarro, 2021). A classical example of gynogenesis is the generation of wheat DHs by the pollination with maize pollen (Eliby et al., 2022). Wide crossing consists of crossing a genotype (typically an F1 or F2) with an inducer that generates several abnormal fertilization products among which a fraction are haploid seeds (haploid induction rate, HIR) which ploidy is then doubled with colchicine. This has been applied in crop species such as maize in which the inducers are intraspecific lines (Prigge et al., 2012; Chen et al., 2023) and barley where, differently, the inducer is usually the wild relative *Hordeum bulbosum* (KASHA and KAO, 1970; Powell et al., 1985). Interestingly, high temperatures were shown to increase the haploid induction rate in *Arabidopsis* (Wang et al., 2023).

In addition, in breeding programs which final output is an F1 hybrid (e.g., maize) that must be recreated continually because genetic segregation would dilute heterosis in the F2, suppressing recombination in the hybrid might be the way to avoid the laborious work of crossing parental lines at a commercial scale each time the hybrid must be created (Crismani et al., 2013). Firstly developed in *Arabidopsis*, a method called reverse breeding consists of regenerating homozygous parental lines from hybrid plants by producing doubled haploid plants from the hybrid gametes free of COs (Wijnker et al., 2012). Importantly, this method would allow to first select the best hybrid and then recreate the parental lines as well as the production of substitution lines (e.g., heterozygous for only one chromosome) (Crismani et al., 2013). Furthermore, apomixis, the clonal reproduction through seeds, can be used to maintain heterosis through generations, however, in most major crops this is not applicable naturally (Crismani et al., 2013; Mercier et al.,

2015). In this way, a study in *Arabidopsis* artificially enabled asexual seed formation by crossing mutants that produce diploid clonal gametes with a strain whose chromosomes are engineered to be eliminated after fertilization, then generating seeds but avoiding recombination in the parent (Marimuthu et al., 2011).

2.8 Recombination in barley

Barley was domesticated in the Fertile Crescent about 10,000 years ago, being an important crop from the annals of agriculture (Comadran et al., 2011; Pasam et al., 2012). Nowadays, barley is the fourth most produced crop worldwide (only after maize, rice, and wheat) because its importance in animal feeding, food, and beverage production (Pasam et al., 2012). In addition, the barley diploid genome ($2n = 2x = 14$) and its self-pollinated reproduction converted barley into a suitable model species for the Triticeae tribe, which includes other important crops such as wheat and rye (Phillips et al., 2015). This condition has been favored with the existence of a large germplasm collection (Knüpffer and van Hintum, 2003; Haseneyer et al., 2010; Pasam et al., 2012), genotyping technologies and resources (Stephens et al., 2004; Stein et al., 2007; Sato et al., 2009; Comadran et al., 2011; Sato et al., 2011; Sannemann et al., 2015; Maurer et al., 2015; Bayer et al., 2017), among the last being remarkable the whole-genome sequence of the variety Morex that has been used as the reference physical genome for barley analysis (Mayer et al., 2012; Mascher et al., 2017; Monat et al., 2019; Mascher et al., 2021) and that opened the door for barley into the pan-genome era (Jayakodi et al., 2020).

The existence of genome-wide intra-species recombination variation has been demonstrated in barley but most studies have been performed in a few cultivated varieties (Nilsson and Pelger, 1991; Dreissig et al., 2019) except studies on the HEB-25 NAM population of barley which involves crosses of a cultivated variety (Barke) as common parent with 25 wild barley (*Hordeum vulgare subsp. spontaneum*) accessions (Dreissig et al., 2020). In such study, the observed recombination rate variation among populations was possibly obscured by extensive SVs between parental genomes. In addition, heterochiasmy has been observed in barley, being male CO rate higher than such of females (Devaux et al., 1995), and this pattern can be altered by temperature (Phillips et al., 2015). Moreover, the recombination rate in barley varies along the genome. On the broad scale, recombination occurs in the distal regions of the chromosomes while is suppressed in the long pericentromeric regions around the centromeres as also observed in other grasses (Pedersen et al., 1995; Comadran et al., 2011), thus making it inaccessible for breeding at least 20% of the functional genes (Mayer et al., 2012; Baker et al.,

2014; Phillips et al., 2015), and being such recombination landscape pattern highly conserved between wild and domesticated barley (Dreissig et al., 2019). At a finer scale, COs are confined to small hotspots that alternate with regions of highly suppressed recombination (Künzel et al., 2000). Furthermore, the recombination rate in barley has been observed to be positively associated with the genetic load and sequence diversity (Künzel et al., 2000; Higgins et al., 2012; Dreissig et al., 2020). Additionally, it was shown that meiotic spatiotemporal progression determines CO distribution in the barley chromosomes, being the synapsis and recombination firstly initiated in distal chromosomal regions, in coordination with the chromatin cycles, both determining chiasmata to occur in distal regions while recombination initiation starts later in proximal regions and rarely progresses to form chiasmata (Higgins et al., 2012, 2014; Phillips et al., 2015). Importantly, such polarization of the CO distribution can be manipulated by modest temperature shifts producing a re-distribution of COs in favor of the proximal regions (Higgins et al., 2012; Phillips et al., 2015). In relation to this, natural variation in the recombination rate among subpopulations of wild barley was found to be correlated with environmental variables in their domestication origins such as temperature, isothermality, solar radiation, and annual precipitation (Dreissig et al., 2019).

Despite the above-mentioned efforts, the recombination rate in barley has been less studied than in other important crops such as maize, rice, and wheat, and much information is lacking for its effective manipulation.

2.9 Objectives

1. Assess the extent and distribution of recombination rate variation in cultivated barley.
2. Evaluate the ability of a genomic selection approach to predict the recombination rate variation.
3. Identify the genomic features that best explain the recombination variation among populations of cultivated barley.
4. Detect recombination events in the barley genome at high-resolution.
5. Analyze the effect of genomic features in determining the location of recombination hotspots and coldspots along the genome.
6. Assess the accuracy of estimated genetic maps and recombination rate based on pool genotyping and sequencing.

References

- Acquaviva, L., Székvölgyi, L., Dichtl, B., Dichtl, B. S., de La Roche Saint André, C., Nicolas, A. and Géli, V. (2013), ‘The COMPASS Subunit Spp1 Links Histone Methylation to Initiation of Meiotic Recombination’, *Science* **339**, 215–218.
- Ahmadi, B. and Ebrahimzadeh, H. (2020), ‘In vitro androgenesis: spontaneous vs. artificial genome doubling and characterization of regenerants’, *Plant Cell Reports* **39**, 299–316.
- Akhunov, E. D., Goodyear, A. W., Geng, S., Qi, L. L., Echalier, B., Gill, B. S., Miftahudin, A., Gustafson, J. P., Lazo, G., Chao, S., Anderson, O. D., Linkiewicz, A. M., Dubcovsky, J., Rota, M. L., Sorrells, M. E., Zhang, D., Nguyen, H. T., Kalavacharla, V., Hossain, K., Kianian, S. F., Peng, J., Lapitan, N. L., Gonzalez-Hernandez, J. L., Anderson, J. A., Choi, D. W., Close, T. J., Dilbirligi, M., Gill, K. S., Walker-Simmons, M. K., Steber, C., McGuire, P. E., Qualset, C. O. and Dvorak, J. (2003), ‘The organization and rate of evolution of wheat genomes are correlated with recombination rates along chromosomes arms’, *Genome Research* **13**, 753–763.
- Allers, T. and Lichten, M. (2001*a*), ‘Differential Timing and Control of Non-crossover and Crossover Recombination during Meiosis’, *Cell* **106**, 47–57.
- Allers, T. and Lichten, M. (2001*b*), ‘Intermediates of Yeast Meiotic Recombination Contain Heteroduplex DNA’, *Molecular Cell* **8**, 225–231.
- Anderson, L. K. (2004), ‘Integrating Genetic Linkage Maps With Pachytene Chromosome Structure in Maize’, *Genetics* **166**, 1923–1933.
- Anderson, L. K., Doyle, G. G., Brigham, B., Carter, J., Hooker, K. D., Lai, A., Rice, M. and Stack, S. M. (2003), ‘High-Resolution Crossover Maps for Each Bivalent of Zea mays Using Recombination Nodules’, *Genetics* **165**, 849–865.
- Anderson, L. K., Reeves, A., Webb, L. M. and Ashley, T. (1999), ‘Distribution of Crossing Over on Mouse Synaptonemal Complexes Using Immunofluorescent Localization of MLH1 Protein’, *Genetics* **151**, 1569–1579.
- Andersson, D. I. and Hughes, D. (1996), ‘Muller’s ratchet decreases fitness of a DNA-based microbe.’, *Proceedings of the National Academy of Sciences* **93**, 906–907.

- Anugrahwati, D. R., Shepherd, K. W., Verlin, D. C., Zhang, P., Ghader Mirzagheri, Walker, E., Francki, M. G. and Dundas, I. S. (2008), ‘Isolation of wheat-rye 1RS recombinants that break the linkage between the stem rust resistance gene SrR and secalin’, *Genome* **51**, 341–349.
- Apuli, R. P., Bernhardsson, C., Schiffthaler, B., Robinson, K. M., Jansson, S., Street, N. R. and Ingvarsson, P. K. (2020), ‘Inferring the genomic landscape of recombination rate variation in European Aspen (*Populus tremula*)’, *G3: Genes, Genomes, Genetics* **10**, 299–309.
- Arbeithuber, B., Betancourt, A. J., Ebner, T. and 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**, 2109–2114.
- Armstrong, S. J. and Jones, G. H. (2001), ‘Female meiosis in wild-type *Arabidopsis thaliana* and in two meiotic mutants’, *Sexual Plant Reproduction* **13**, 177–183.
- Baker, K., Bayer, M., Cook, N., Dreißig, S., Dhillon, T., Russell, J., Hedley, P. E., Morris, J., Ramsay, L., Colas, I., Waugh, R., Steffenson, B., Milne, I., Stephen, G., Marshall, D. and Flavell, A. J. (2014), ‘The low-recombining pericentromeric region of barley restricts gene diversity and evolution but not gene expression’, *Plant Journal* **79**, 981–992.
- Barakate, A., Higgins, J. D., Vivera, S., Stephens, J., Perry, R. M., Ramsay, L., Colas, I., Oakey, H., Waugh, R., Franklin, F. C. H., Armstrong, S. J. and Halpin, C. (2014), ‘The synaptonemal complex protein ZYP1 is required for imposition of meiotic crossovers in barley’, *Plant Cell* **26**, 729–740.
- Barnabás, B., Obert, B. and Kovács, G. (1999), ‘Colchicine, an efficient genome-doubling agent for maize (*Zea mays* L.) microspores cultured in anthero’, *Plant Cell Reports* **18**, 858–862.
- Barnes, T. M., Kohara, Y., Coulson, A. and Hekimi, S. (1995), ‘Meiotic recombination, noncoding DNA and genomic organization in *Caenorhabditis elegans*.’, *Genetics* **141**, 159–79.
- Barth, S., Melchinger, A. E., Devezi-Savula, B. and Lübberstedt, T. (2001), ‘Influence of genetic background and heterozygosity on meiotic recombination in *Arabidopsis thaliana*’, *Genome* **44**, 971–978.
- Barton, N. H. (2009), ‘Why Sex and Recombination?’, *Cold Spring Harbor Symposium on Quantitative Biology* **74**, 187–195.

- Barton, N. H. and Charlesworth, B. (1998), ‘Why Sex and Recombination?’, *Science* **281**, 1986–1990.
- Basu-Roy, S., Gauthier, F., Giraut, L., Mézard, C., Falque, M. and Martin, O. C. (2013), ‘Hot regions of noninterfering crossovers coexist with a Nonuniformly interfering pathway in *Arabidopsis thaliana*’, *Genetics* **195**, 769–779.
- Baudat, F., Buard, J., Grey, C., Fledel-Alon, A., Ober, C., Przeworski, M., Coop, G. and Massy, B. D. (2010), ‘PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice’, *Science* **327**, 836–840.
- Baudat, F. and Massy, B. D. (2007), ‘Regulating double-stranded DNA break repair towards crossover or non-crossover during mammalian meiosis’, *Chromosome Research* **15**, 565–577.
- Bauer, E., Falque, M., Walter, H., Bauland, C., Camisan, C., Campo, L., Meyer, N., Ranc, N., Rincet, R., Schipprack, W., Altmann, T., Flament, P., Melchinger, A. E., Menz, M., Moreno-González, J., Ouzunova, M., Revilla, P., Charcosset, A., Martin, O. C. and Schön, C. C. (2013), ‘Intraspecific variation of recombination rate in maize’, *Genome Biology* **14**.
- Bayer, M. M., Rapazote-Flores, P., Ganai, M., Hedley, P. E., Macaulay, M., Plieske, J., Ramsay, L., Russell, J., Shaw, P. D., Thomas, W. and Waugh, R. (2017), ‘Development and evaluation of a barley 50k iSelect SNP array’, *Frontiers in Plant Science* **8**, 1–10.
- Becker, C., Hagmann, J., Müller, J., Koenig, D., Stegle, O., Borgwardt, K. and Weigel, D. (2011), ‘Spontaneous epigenetic variation in the *Arabidopsis thaliana* methylome’, *Nature* **480**, 245–249.
- Begun, D. J. and Aquadro, C. F. (1992), ‘Levels of naturally occurring DNA polymorphism correlate with recombination rates in *D. melanogaster*’, *Nature* **356**, 519–520.
- Berchowitz, L. E. and Copenhaver, G. P. (2008), ‘Fluorescent *Arabidopsis* tetrads: A visual assay for quickly developing large crossover and crossover interference data sets’, *Nature Protocols* **3**, 41–50.
- Berchowitz, L. E. and Copenhaver, G. P. (2010), ‘Genetic Interference: Don’t Stand So Close to Me’, *Current Genomics* **11**, 91–102.
- Berchowitz, L. E., Francis, K. E., Bey, A. L. and Copenhaver, G. P. (2007), ‘The Role of AtMUS81 in Interference-Insensitive Crossovers in *A. thaliana*’, *PLoS Genetics* **3**, e132.

- Bernstein, M. R. and Rockman, M. V. (2016), ‘Fine-scale crossover rate variation on the *Caenorhabditis elegans* X chromosome’, *G3: Genes, Genomes, Genetics* **6**, 1767–1776.
- Bhatt, A. M., Lister, C., Page, T., Fransz, P., Findlay, K., Jones, G. H., Dickinson, H. G. and Dean, C. (1999), ‘The DIF1 gene of *Arabidopsis* is required for meiotic chromosome segregation and belongs to the REC8/RAD21 cohesin gene family’, *Plant Journal* **19**, 463–472.
- Blackwell, A. R., Dluzewska, J., Szymanska-Lejman, M., Desjardins, S., Tock, A. J., Kbiri, N., Lambing, C., Lawrence, E. J., Bieluszewski, T., Rowan, B., Higgins, J. D., Ziolkowski, P. A. and Henderson, I. R. (2020), ‘MSH 2 shapes the meiotic crossover landscape in relation to interhomolog polymorphism in *Arabidopsis*’, *The EMBO Journal* **39**, 1–22.
- Blakeslee, A. F., Belling, J., Farnham, M. E. and Bergner, A. D. (1922), ‘A Haploid Mutant in the Jimson Weed, *Datura stramonium*’, *Science* **55**, 646–647.
- Boideau, F., Pelé, A., Tanguy, C., Trotoux, G., Eber, F., Maillet, L., Gilet, M., Lodé-Taburel, M., Huteau, V., Morice, J., Coriton, O., Falentin, C., Delourme, R., Rousseau-Gueutin, M. and Chèvre, A. M. (2021), ‘A modified meiotic recombination in *brassica napus* largely improves its breeding efficiency’, *Biology* **10**.
- Boideau, F., Richard, G., Coriton, O., Huteau, V., Belser, C., Deniot, G., Eber, F., Falentin, C., de Carvalho, J. F., Gilet, M., Lodé-Taburel, M., Maillet, L., Morice, J., Trotoux, G., Aury, J. M., Chèvre, A. M. and Rousseau-Gueutin, M. (2022), ‘Epigenomic and structural events preclude recombination in *Brassica napus*’, *New Phytologist* **234**, 545–559.
- Borde, V., Robine, N., Lin, W., Bonfils, S., Géli, V. and Nicolas, A. (2009), ‘Histone H3 lysine 4 trimethylation marks meiotic recombination initiation sites’, *EMBO Journal* **28**, 99–111.
- Borts, R. H., Chambers, S. and Abdullah, M. (2000), ‘The many faces of mismatch repair in meiosis’, *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **451**, 129–150.
- Bouchet, S., Olatoye, M. O., Marla, S. R., Perumal, R., Tesso, T., Yu, J., Tuinstra, M. and Morris, G. P. (2017), ‘Increased Power To Dissect Adaptive Traits in Global Sorghum Diversity Using a Nested Association Mapping Population’, *Genetics* **206**, 573–585.

- Boyes, D. C., Nasrallah, M. E., Vrebalov, J. and Nasrallah, J. B. (1997), ‘The Self-Incompatibility (S) Haplotypes of Brassica Contain Highly Divergent and Rearranged Sequences of Ancient Origin’, *The Plant Cell* **9**, 237–247.
- Branca, A., Paape, T. D., Zhou, P., Briskine, R., Farmer, A. D., Mudge, J., Bharti, A. K., Woodward, J. E., May, G. D., Gentzbittel, L., Ben, C., Denny, R., Sadowsky, M. J., Ronfort, J., Bataillon, T., Young, N. D. and Tiffin, P. (2011), ‘Whole-genome nucleotide diversity, recombination, and linkage disequilibrium in the model legume *Medicago truncatula*’, *Proceedings of the National Academy of Sciences of the United States of America* **108**, E864–E870.
- Brick, K., Smagulova, F., Khil, P., Camerini-Otero, R. D. and Petukhova, G. V. (2012), ‘Genetic recombination is directed away from functional genomic elements in mice’, *Nature* **485**, 642–645.
- Broman, K. W., Rowe, L. B., Churchill, G. A. and Paigen, K. (2002), ‘Crossover Interference in the Mouse’, *Genetics* **160**, 1123–1131.
- Broman, K. W. and Weber, J. L. (2000), ‘Characterization of Human Crossover Interference’, *Am. J. Hum. Genet* **66**.
- Broman, K. W., Wu, H., Šaunak Sen and Churchill, G. A. (2003), ‘R/qt1: QTL mapping in experimental crosses’, *Bioinformatics* **19**, 889–890.
- Brown, J. and Sundareshan, V. (1991), ‘A recombination hotspot in the maize A1 intragenic region’, *Theoretical and Applied Genetics* **81**, 185–188.
- Buard, J., Barthès, P., Grey, C. and de Massy, B. (2009), ‘Distinct histone modifications define initiation and repair of meiotic recombination in the mouse’, *The EMBO Journal* **28**, 2616–2624.
- Buhler, C., Borde, V. and Lichten, M. (2007), ‘Mapping meiotic single-strand DNA reveals a new landscape of DNA double-strand breaks in *Saccharomyces cerevisiae*’, *PLoS Biology* **5**, 2797–2808.
- Bulankova, P., Riehs-Kearnan, N., Nowack, M. K., Schnittger, A. and Riha, K. (2010), ‘Meiotic progression in Arabidopsis is governed by complex regulatory interactions between SMG7, TDM1, and the meiosis I-specific cyclin TAM’, *Plant Cell* **22**, 3791–3803.
- Burt, A. (2000), ‘Perspective: Sex, recombination, and the efficacy of selection - Was Weismann right?’, *Evolution* **54**, 337–351.

- Börner, G., Kleckner, N. and Hunter, N. (2004), ‘Crossover/Noncrossover Differentiation, Synaptonemal Complex Formation, and Regulatory Surveillance at the Leptotene/Zygotene Transition of Meiosis’, *Cell* **117**, 29–45.
- Cai, G., Yang, Q., Yi, B., Fan, C., Edwards, D., Batley, J. and Zhou, Y. (2014), ‘A complex recombination pattern in the genome of allotetraploid *Brassica napus* as revealed by a high-density genetic map’, *PLoS ONE* **9**.
- Capilla-Pérez, L., Durand, S., Hurel, A., Lian, Q., Chambon, A., Taochy, C., Solier, V., Grelon, M. and Mercier, R. (2021), ‘The synaptonemal complex imposes crossover interference and heterochiasmy in *Arabidopsis*’, *Proceedings of the National Academy of Sciences of the United States of America* **118**, 1–11.
- Carlile, T. M. and Amon, A. (2008), ‘Meiosis I Is Established through Division-Specific Translational Control of a Cyclin’, *Cell* **133**, 280–291.
- Carter, T. C. and Falconer, D. S. (1951), ‘Stocks for detecting linkage in the mouse, and the theory of their design’, *Journal of Genetics* **50**, 307–323.
- Cartwright, D. A., Troggio, M., Velasco, R. and Gutin, A. (2007), ‘Genetic mapping in the presence of genotyping errors’, *Genetics* **176**, 2521–2527.
- Castle, W. E. (1919), ‘Is the arrangement of the genes in the chromosome linear?’, *Proceedings of the National Academy of Sciences* **5**, 25–32.
- Cavanagh, C., Morell, M., Mackay, I. and Powell, W. (2008), ‘From mutations to MAGIC: resources for gene discovery, validation and delivery in crop plants’, *Current Opinion in Plant Biology* **11**, 215–221.
- Chalhoub, B., Denoeud, F., Liu, S., Parkin, I. A. P., Tang, H., Wang, X., Chiquet, J., Belcram, H., Tong, C., Samans, B., Corr  a, M., Silva, C. D., Just, J., Falentin, C., Koh, C. S., Clainche, I. L., Bernard, M., Bento, P., Noel, B., Labadie, K., Alberti, A., Charles, M., Arnaud, D., Guo, H., Daviaud, C., Alamery, S., Jabbari, K., Zhao, M., Edger, P. P., Chelaifa, H., Tack, D., Lassalle, G., Mestiri, I., Schnel, N., Paslier, M.-C. L., Fan, G., Renault, V., Bayer, P. E., Golicz, A. A., Manoli, S., Lee, T.-H., Thi, V. H. D., Chalabi, S., Hu, Q., Fan, C., Tollenaere, R., Lu, Y., Battail, C., Shen, J., Sidebottom, C. H. D., Wang, X., Canaguier, A., Chauveau, A., B  rard, A., Deniot, G., Guan, M., Liu, Z., Sun, F., Lim, Y. P., Lyons, E., Town, C. D., Bancroft, I., Wang, X., Meng, J., Ma, J., Pires, J. C., King, G. J., Brunel, D., Delourme, R., Renard, M., Aury, J.-M., Adams, K. L., Batley, J., Snowdon, R. J., Tost, J., Edwards, D., Zhou, Y., Hua, W., Sharpe, A. G., Paterson, A. H., Guan, C. and Wincker,

- P. (2014), ‘Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome’, *Science* **345**, 950–953.
- Charlesworth, B., Morgan, M. T. and Charlesworth, D. (1993), ‘The effect of deleterious mutations on neutral molecular variation.’, *Genetics* **134**, 1289–1303.
- Chelysheva, L., Diallo, S., Vezon, D., Gendrot, G., Vrielynck, N., Belcram, K., Rocques, N., Márquez-Lema, A., Bhatt, A. M., Horlow, C., Mercier, R., Mézard, C. and Grelon, M. (2005), ‘AtREC8 and AtSCC3 are essential to the monopolar orientation of the kinetochores during meiosis’, *Journal of Cell Science* **118**, 4621–4632.
- Chelysheva, L., Gendrot, G., Vezon, D., Doutriaux, M. P., Mercier, R. and Grelon, M. (2007), ‘Zip4/Spo22 is required for class I CO formation but not for synapsis completion in *Arabidopsis thaliana*’, *PLoS Genetics* **3**, 802–813.
- Chelysheva, L., Vezon, D., Chambon, A., Gendrot, G., Pereira, L., Lemhemdi, A., Vrielynck, N., Guin, S. L., Novatchkova, M. and Grelon, M. (2012), ‘The *Arabidopsis* HEI10 is a new ZMM protein related to Zip3’, *PLoS Genetics* **8**.
- Chen, J. M., Cooper, D. N., Chuzhanova, N., Férec, C. and Patrinos, G. P. (2007), ‘Gene conversion: Mechanisms, evolution and human disease’, *Nature Reviews Genetics* **8**, 762–775.
- Chen, Y. R., Lübberstedt, T. and Frei, U. K. (2023), ‘Development of doubled haploid inducer lines facilitates selection of superior haploid inducers in maize’, *Frontiers in Plant Science* **14**.
- Choi, K. and Henderson, I. R. (2015), ‘Meiotic recombination hotspots - A comparative view’, *Plant Journal* **83**, 52–61.
- Choi, K., Reinhard, C., Serra, H., Ziolkowski, P. A., Underwood, C. J., Zhao, X., Hardcastle, T. J., Yelina, N. E., Griffin, C., Jackson, M., Mézard, C., McVean, G., Copenhaver, G. P. and Henderson, I. R. (2016), ‘Recombination Rate Heterogeneity within *Arabidopsis* Disease Resistance Genes’, *PLoS Genetics* **12**.
- Choi, K., Zhao, X., Kelly, K. A., Venn, O., Higgins, J. D., Yelina, N. E., Hardcastle, T. J., Ziolkowski, P. A., Copenhaver, G. P., Franklin, F. C. H., Mcvean, G. and Henderson, I. R. (2013), ‘*Arabidopsis* meiotic crossover hot spots overlap with H2A.Z nucleosomes at gene promoters’, *Nature Genetics* **45**, 1327–1338.
- Choi, K., Zhao, X., Lambing, C., Underwood, C. J., Hardcastle, T. J., Serra, H., Tock, A. J., Ziolkowski, P. A., Yelina, N. E., Martienssen, R. A. and Henderson,

- I. R. (2017), ‘Nucleosomes and DNA methylation shape meiotic DSB frequency in Arabidopsis transposons and gene regulatory regions’, *bioRxiv* pp. 532–546.
- Cifuentes, M., Rivard, M., Pereira, L., Chelysheva, L. and Mercier, R. (2013), ‘Haploid Meiosis in Arabidopsis: Double-Strand Breaks Are Formed and Repaired but Without Synapsis and Crossovers’, *PLoS ONE* **8**.
- Cirulli, E. T., Kliman, R. M. and Noor, M. A. (2007), ‘Fine-scale crossover rate heterogeneity in *Drosophila pseudoobscura*’, *Journal of Molecular Evolution* **64**, 129–135.
- Cobbs, G. (1978), ‘Renewal process approach to the theory of genetic linkage: Case of no chromatid interference’, *Genetics* **89**, 563–581.
- Colomé-Tatché, M., Cortijo, S., Wardenaar, R., Morgado, L., Lahouz, B., Sarazin, A., Etcheverry, M., Martin, A., Feng, S., Duvernois-Berthet, E., Labadie, K., Wincker, P., Jacobsen, S. E., Jansen, R. C., Colot, V. and Johannes, F. (2012), ‘Features of the Arabidopsis recombination landscape resulting from the combined loss of sequence variation and DNA methylation’, *Proceedings of the National Academy of Sciences of the United States of America* **109**, 16240–16245.
- Comadran, J., Ramsay, L., MacKenzie, K., Hayes, P., Close, T. J., Muehlbauer, G., Stein, N. and Waugh, R. (2011), ‘Patterns of polymorphism and linkage disequilibrium in cultivated barley’, *Theoretical and Applied Genetics* **122**, 523–531.
- Comeron, J. M., Ratnappan, R. and Bailin, S. (2012), ‘The Many Landscapes of Recombination in *Drosophila melanogaster*’, *PLoS Genetics* **8**.
- Coop, G., Wen, X., Ober, C., Pritchard, J. K. and Przeworski, M. (2008), ‘High-resolution mapping of crossovers reveals extensive variation in fine-scale recombination patterns among humans’, *Science* **319**, 1395–1398.
- Copenhaver, G. P., Housworth, E. A. and Stahl, F. W. (2002), ‘Crossover Interference in Arabidopsis’, *Genetics* **160**, 1631–1639.
- Corem, S., Doron-Faigenboim, A., Jouffroy, O., Maumus, F., Arazi, T. and Bouché, N. (2018), ‘Redistribution of CHH methylation and small interfering RNAs across the genome of tomato ddm1 mutants’, *Plant Cell* **30**, 1628–1644.
- Coulton, A., BurrIDGE, A. J. and Edwards, K. J. (2020), ‘Examining the Effects of Temperature on Recombination in Wheat’, *Frontiers in Plant Science* **11**.

- Cox, A., Ackert-Bicknell, C. L., Dumont, B. L., Yueming, D., Bell, J. T., Brockmann, G. A., Wergedal, J. E., Bult, C., Paigen, B., Flint, J., Tsaih, S. W., Churchill, G. A. and Broman, K. W. (2009), ‘A new standard genetic map for the laboratory mouse’, *Genetics* **182**, 1335–1344.
- Crismani, W., Girard, C., Froger, N., Pradillo, M., Santos, J. L., Chelysheva, L., Copenhaver, G. P., Horlow, C. and Mercier, R. (2012), ‘FANCM Limits Meiotic Crossovers’, *Science* **336**, 1588–1590.
- Crismani, W., Girard, C. and Mercier, R. (2013), ‘Tinkering with meiosis’, *Journal of Experimental Botany* **64**, 55–65.
- Cromie, G. A., Hyppa, R. W., Cam, H. P., Farah, J. A., Grewal, S. I. and Smith, G. R. (2007), ‘A discrete class of intergenic DNA dictates meiotic DNA break hotspots in fission yeast’, *PLoS Genetics* **3**, 1496–1507.
- Daenen, F., van Roy, F. and Bleser, P. J. D. (2008), ‘Low nucleosome occupancy is encoded around functional human transcription factor binding sites’, *BMC Genomics* **9**.
- Darrier, B., Rimbart, H., Balfourier, F., Pingault, L., Josselin, A. A., Servin, B., Navarro, J., Choulet, F., Paux, E. and Sourdille, P. (2017), ‘High-resolution mapping of crossover events in the hexaploid wheat genome suggests a universal recombination mechanism’, *Genetics* **206**, 1373–1388.
- de Givry, S., Bouchez, M., Chabrier, P., Milan, D. and Schiex, T. (2005), ‘CARTHAGENE: Multipopulation integrated genetic and radiation hybrid mapping’, *Bioinformatics* **21**, 1703–1704.
- De Massy, B. (2013), ‘Initiation of meiotic recombination: How and where? Conservation and specificities among eukaryotes’, *Annual Review of Genetics* **47**, 563–599.
- D’Erfurth, I., Jolivet, S., Froger, N., Catrice, O., Novatchkova, M. and Mercier, R. (2009), ‘Turning meiosis into mitosis’, *PLoS Biology* **7**.
- Devaux, P., Kilian, A. and Kleinhofs, A. (1995), ‘Comparative mapping of the barley genome with male and female recombination-derived, doubled haploid populations’, *Mol Gen Genet* **249**, 600–608.
- Dobzhansky, T. (1931), ‘Translocation involving the second and the fourth chromosomes of *Drosophila melanogaster*’, *Genetics* **16**, 629–658.

- Dole, J. and Weber, D. F. (2007), ‘Detection of quantitative trait loci influencing recombination using recombinant inbred lines’, *Genetics* **177**, 2309–2319.
- Dooner, H. K. (1986), ‘Genetic fine structure of the bronze locus in maize’, *Genetics* **113**, 1021–1036.
- Dooner, H. K. (2002), ‘Extensive interallelic polymorphisms drive meiotic recombination into a crossover pathway’, *Plant Cell* **14**, 1173–1183.
- Dooner, H. K. and Martínez-Férez, I. M. (1997), ‘Recombination Occurs Uniformly within the bronze Gene, a Meiotic Recombination Hotspot in the Maize Genome’, *The Plant Cell* **9**, 1633–1634.
- Dowrick, G. J. (1957), ‘The influence of temperature on meiosis’, *Heredity* **11**, 37–49.
- Dreissig, S., Fuchs, J., Cápál, P., Kettles, N., Byrne, E. and Houben, A. (2015), ‘Measuring meiotic crossovers via multi-locus genotyping of single pollen grains in barley’, *PLoS ONE* **10**, 1–10.
- Dreissig, S., Mascher, M., Heckmann, S. and Purugganan, M. (2019), ‘Variation in Recombination Rate Is Shaped by Domestication and Environmental Conditions in Barley’, *Molecular Biology and Evolution* **36**, 2029–2039.
- Dreissig, S., Maurer, A., Sharma, R., Milne, L., Flavell, A. J., Schmutzer, T. and Pillen, K. (2020), ‘Natural variation in meiotic recombination rate shapes introgression patterns in intraspecific hybrids between wild and domesticated barley’, *New Phytologist* **228**, 1852–1863.
- Drouaud, J., Camilleri, C., Bourguignon, P. Y., Canaguier, A., Bérard, A., Vezon, D., Giancola, S., Brunel, D., Colot, V., Prum, B., Quesneville, H. and Mézard, C. (2006), ‘Variation in crossing-over rates across chromosome 4 of *Arabidopsis thaliana* reveals the presence of meiotic recombination “hot spots”’, *Genome Research* **16**, 106–114.
- Drouaud, J., Khademian, H., Giraut, L., Zanni, V., Bellalou, S., Henderson, I. R., Falque, M. and Mézard, C. (2013), ‘Contrasted Patterns of Crossover and Non-crossover at *Arabidopsis thaliana* Meiotic Recombination Hotspots’, *PLoS Genetics* **9**.
- Dumont, B. L., Broman, K. W. and Payseur, B. A. (2009), ‘Variation in genomic recombination rates among heterogeneous stock mice’, *Genetics* **182**, 1345–1349.

- Dunn, L. C. and Bennett, D. (1967), ‘Sex differences in recombination of linked genes in animals’, *Genetical Research* **9**, 211–220.
- Duret, L. and Galtier, N. (2009), ‘Biased Gene Conversion and the Evolution of Mammalian Genomic Landscapes’, *Annual Review of Genomics and Human Genetics* **10**, 285–311.
- Edlinger, B. and Schlögelhofer, P. (2011), ‘Have a break: Determinants of meiotic DNA double strand break (DSB) formation and processing in plants’, *Journal of Experimental Botany* **62**, 1545–1563.
- Eliby, S., Bekkuzhina, S., Kishchenko, O., Iskakova, G., Kylyshbayeva, G., Jatayev, S., Soole, K., Langridge, P., Borisjuk, N. and Shavrukov, Y. (2022), ‘Developments and prospects for doubled haploid wheat’, *Biotechnology Advances* **60**, 108007.
- Elliott, C. G. (1955), ‘The effect of temperature on chiasma frequency’, *Heredity* **9**, 385–398.
- Esch, E., Szymaniak, J. M., Yates, H., Pawlowski, W. P. and Buckler, E. S. (2007), ‘Using crossover breakpoints in recombinant inbred lines to identify quantitative trait loci controlling the global recombination frequency’, *Genetics* **177**, 1851–1858.
- Evans, E., Sugawara, N., Haber, J. E. and Alani, E. (2000), ‘The *Saccharomyces cerevisiae* Msh2 Mismatch Repair Protein Localizes to Recombination Intermediates In Vivo 1983). Base pair exchange between the ss filament and the double-stranded DNA (dsDNA) stabilizes the interactions and promotes further pairing leading to the formation of heteroduplex DNA (hDNA). DNA synthesis primed from the 3 end of the invading strand results in the’, *Molecular Cell* **5**, 789–799.
- Fatmi, A., Poneleit, C. G. and Pfeiffer, T. W. (1993), ‘Variability of recombination frequencies in the Iowa Stiff Stalk Synthetic (*Zea mays* L.)’, *Theor Appl Genet* **86**, 859–866.
- Faux, A., Gorjanc, G., Gaynor, R. C., Battagin, M., Edwards, S. M., Wilson, D. L., Hearne, S. J., Gonen, S. and Hickey, J. M. (2016), ‘AlphaSim: Software for Breeding Program Simulation’, *The Plant Genome* **9**.
- Fearnhead, P. (2006), ‘SequenceLDhot: Detecting recombination hotspots’, *Bioinformatics* **22**, 3061–3066.

- Felsenstein, J. (1974), ‘The evolutionary advantage of recombination.’, *Genetics* **78**, 737–56.
- Ferdous, M., Higgins, J. D., Osman, K., Lambing, C., Roitinger, E., Mechtler, K., Armstrong, S. J., Perry, R., Pradillo, M., Cuñado, N. and Franklin, F. C. H. (2012), ‘Inter-homolog crossing-over and synapsis in Arabidopsis meiosis are dependent on the chromosome axis protein atasy3’, *PLoS Genetics* **8**.
- Ferguson, K. A., Wong, E. C., Chow, V., Nigro, M. and Ma, S. (2007), ‘Abnormal meiotic recombination in infertile men and its association with sperm aneuploidy’, *Human Molecular Genetics* **16**, 2870–2879.
- Fernandes, J. B., Duhamel, M., Seguéla-Arnaud, M., Froger, N., Girard, C., Choinard, S., Solier, V., Winne, N. D., Jaeger, G. D., Gevaert, K., Andrey, P., Grelon, M., Guerois, R., Kumar, R. and Mercier, R. (2018b), ‘FIGL1 and its novel partner FLIP form a conserved complex that regulates homologous recombination’, *PLoS Genetics* **14**.
- Fernandes, J. B., Naish, M., Lian, Q., Burns, R., Tock, A. J., Rabanal, F. A., Wlodzimierz, P., Habring, A., Nicholas, R. E., Weigel, D., Mercier, R. and Henderson, I. R. (2024), ‘Structural variation and DNA methylation shape the centromere-proximal meiotic crossover landscape in Arabidopsis’, *Genome Biology* **25**, 30.
- Fernandes, J. B., Séguéla-Arnaud, M., Larchevêque, C., Lloyd, A. H. and Mercier, R. (2018a), ‘Unleashing meiotic crossovers in hybrid plants’, *Proceedings of the National Academy of Sciences of the United States of America* **115**, 2431–2436.
- Ferris, P. J. and Goodenough, U. W. (1994), ‘The mating-type locus of *Chlamydomonas reinhardtii* contains highly rearranged DNA sequences’, *Cell* **76**, 1135–1145.
- Fisher, R. A. (1930), *The genetical theory of natural selection*, Clarendon Press.
- Fledel-Alon, A., Leffler, E. M., Guan, Y., Stephens, M., Coop, G. and Przeworski, M. (2011), ‘Variation in human recombination rates and its genetic determinants’, *PLoS ONE* **6**.
- Fogwill, M. (1957), ‘Differences in crossing-over and chromosome size in the sex cells of *Lilium* and *Fritillaria*’, *Chromosoma* **9**, 493–504.
- Foss, E., Lande, R., Stahl, F. W. and Steinberg, C. M. (1993), ‘Chiasma interference as a function of genetic distance.’, *Genetics* **133**, 681–691.

- Francis, K. E., Lam, S. Y. and Copenhaver, G. P. (2006), ‘Separation of Arabidopsis pollen tetrads is regulated by QUARTET1, a pectin methylesterase gene’, *Plant Physiology* **142**, 1004–1013.
- Francis, K. E., Lam, S. Y., Harrison, B. D., Bey, A. L., Berchowitz, L. E. and Copenhaver, G. P. (2007), ‘Pollen tetrad-based visual assay for meiotic recombination in Arabidopsis’, *Proceedings of the National Academy of Sciences of the United States of America* **104**, 3913–3918.
- Franklin, A. E., Mcelver, J., Sunjevaric, I., Rothstein, R., Bowen, B. and Cande, W. Z. (1999), ‘Three-Dimensional Microscopy of the Rad51 Recombination Protein during Meiotic Prophase’, *The Plant Cell* **11**, 809–824.
- Franz, P., Linc, G., Lee, C. R., Aflitos, S. A., Lasky, J. R., Toomajian, C., Ali, H., Peters, J., van Dam, P., Ji, X., Kuzak, M., Gerats, T., Schubert, I., Schneeberger, K., Colot, V., Martienssen, R., Koornneef, M., Nordborg, M., Juenger, T. E., de Jong, H. and Schranz, M. E. (2016), ‘Molecular, genetic and evolutionary analysis of a paracentric inversion in Arabidopsis thaliana’, *Plant Journal* **88**, 159–178.
- Fu, H., Zheng, Z. and Dooner, H. K. (2001), ‘Recombination rates between adjacent genic and retrotransposon regions in maize vary by 2 orders of magnitude’, *Proceedings of the National Academy of Sciences of the United States of America* **99**, 1082–1087.
- Fuller, Z. L., Leonard, C. J., Young, R. E., Schaeffer, S. W. and Phadnis, N. (2018), ‘Ancestral polymorphisms explain the role of chromosomal inversions in speciation’, *PLoS Genetics* **14**.
- Futcher, B. (2008), ‘Cyclins in Meiosis: Lost in Translation’, *Developmental Cell* **14**, 644–645.
- Ganal, M. W., Durstewitz, G., Polley, A., Bérard, A., Buckler, E. S., Charcosset, A., Clarke, J. D., Graner, E. M., Hansen, M., Joets, J., Paslier, M. C. L., McMullen, M. D., Montalent, P., Rose, M., Schön, C. C., Sun, Q., Walter, H., Martin, O. C. and Falque, M. (2011), ‘A large maize (*Zea mays* L.) SNP genotyping array: Development and germplasm genotyping, and genetic mapping to compare with the B73 reference genome’, *PLoS ONE* **6**, 1–7.
- Garcia, V., Phelps, S. E., Gray, S. and Neale, M. J. (2011), ‘Bidirectional resection of DNA double-strand breaks by Mre11 and Exo1’, *Nature* **479**, 241–244.

- Gardiner, L. J., Wingen, L. U., Bailey, P., Joynson, R., Brabbs, T., Wright, J., Higgins, J. D., Hall, N., Griffiths, S., Clavijo, B. J. and Hall, A. (2019), ‘Analysis of the recombination landscape of hexaploid bread wheat reveals genes controlling recombination and gene conversion frequency’, *Genome Biology* **20**, 69.
- Gay, J., Myers, S. and McVean, G. (2007), ‘Estimating meiotic gene conversion rates from population genetic data’, *Genetics* **177**, 881–894.
- Gill, K. S., Gill, B. S., Endot, T. R. and Boyko, E. V. (1996), ‘Identification and High-Density Mapping of Gene-Rich Regions in Chromosome Group 5 of Wheat’, *Genetics* **143**, 1001–1012.
- Gion, J. M., Hudson, C. J., Lesur, I., Vaillancourt, R. E., Potts, B. M. and Freeman, J. S. (2016), ‘Genome-wide variation in recombination rate in Eucalyptus’, *BMC Genomics* **17**, 1–12.
- Girard, C., Chelysheva, L., Choinard, S., Froger, N., Macaisne, N., Lehemdi, A., Mazel, J., Crismani, W. and Mercier, R. (2015), ‘AAA-ATPase FIDGETIN-LIKE 1 and Helicase FANCM Antagonize Meiotic Crossovers by Distinct Mechanisms’, *PLoS Genetics* **11**.
- Girard, C., Crismani, W., Froger, N., Mazel, J., Lemhemdi, A., Horlow, C. and Mercier, R. (2014), ‘FANCM-associated proteins MHF1 and MHF2, but not the other Fanconi anemia factors, limit meiotic crossovers’, *Nucleic Acids Research* **42**, 9087–9095.
- Giraut, L., Falque, M., Drouaud, J., Pereira, L., Martin, O. C. and Mézard, C. (2011), ‘Genome-wide crossover distribution in *Arabidopsis thaliana* meiosis reveals sex-specific patterns along chromosomes’, *PLoS Genetics* **7**.
- Glémin, S., Clément, Y., David, J. and Ressayre, A. (2014), ‘GC content evolution in coding regions of angiosperm genomes: a unifying hypothesis’, *Trends in Genetics* **30**, 263–270.
- Gohil, R. N. and Kaul, R. (1981), ‘Studies on Male and Female Meiosis in Indian Allium II. Autotetraploid *Allium tuberosum*’, *Chromosoma (Berl.)* **82**, 735–739.
- Goldfarb, T. and Lichten, M. (2010), ‘Frequent and efficient use of the sister chromatid for DNA double-strand break repair during budding yeast meiosis’, *PLoS Biology* **8**.
- Golubovskaya, I. N., Hamant, O., Timofejeva, L., Wang, C. J. R., Braun, D., Meeley, R. and Cande, W. Z. (2006), ‘Alleles of *afd1* dissect REC8 functions during meiotic prophase I’, *Journal of Cell Science* **119**, 3306–3315.

- Grelon, M. (2001), ‘AtSPO11-1 is necessary for efficient meiotic recombination in plants’, *The EMBO Journal* **20**, 589–600.
- Grey, C., Barthès, P., Friec, G., Langa, F., Baudat, F. and de Massy, B. (2011), ‘Mouse Prdm9 DNA-binding specificity determines sites of histone H3 lysine 4 trimethylation for initiation of meiotic recombination’, *PLoS Biology* **9**.
- Griffiths, S., Sharp, R., Foote, T. N., Bertin, I., Wanous, M., Reader, S., Colas, I. and Moore, G. (2006), ‘Molecular characterization of Ph1 as a major chromosome pairing locus in polyploid wheat’, *Nature* **439**, 749–752.
- Gutierrez-Gonzalez, J. J., Mascher, M., Poland, J. and Muehlbauer, G. J. (2019), ‘Dense genotyping-by-sequencing linkage maps of two Synthetic W7984×Opata reference populations provide insights into wheat structural diversity’, *Scientific Reports* **9**, 1–15.
- Hadad, R. G., Pfeiffer, T. W. and Poneleit, C. G. (1996), ‘Repeatability and heritability of divergent recombination frequencies in the Iowa Stiff Stalk Synthetic (*Zea mays* L.)’, *Theoretical and Applied Genetics* **93-93**, 990–996.
- Haldane, J. (1919), ‘The combination of linkage values, and the calculation of distances between the loci of linked factors’, *Journal of Genetics* **8**, 299–309.
- Haldane, J. B. S. (1922), ‘Sex ratio and unisexual sterility in hybrid animals’, *Journal of Genetics* **12**, 101–109.
- Hall, J. C. (1972), ‘Chromosome segregation influenced by two alleles of the meiotic mutant c(3)G in *Drosophila melanogaster*’, *Genetics* **71**, 367–400.
- 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., Thorsteinsdottir, U. and Stefansson, K. (2019), ‘Characterizing mutagenic effects of recombination through a sequence-level genetic map’, *Science* **363**.
- Hammarlund, M., Davis, M. W., Nguyen, H., Dayton, D. and Jorgensen, E. M. (2005), ‘Heterozygous insertions alter crossover distribution but allow crossover interference in *Caenorhabditis elegans*’, *Genetics* **171**, 1047–1056.
- Hartung, F., Suer, S., Knoll, A., Wurz-Wildersinn, R. and Puchta, H. (2008), ‘Topoisomerase 3 α and RMI1 suppress somatic crossovers and are essential for resolution of meiotic recombination intermediates in *Arabidopsis thaliana*’, *PLoS Genetics* **4**.

- Hartung, F., Wurz-Wildersinn, R., Fuchs, J., Schubert, I., Suer, S. and Puchta, H. (2007), ‘The catalytically active tyrosine residues of both SPO11-1 and SPO11-2 are required for meiotic double-strand break induction in Arabidopsis’, *Plant Cell* **19**, 3090–3099.
- Haseneyer, G., Stracke, S., Paul, C., Einfeldt, C., Broda, A., Piepho, H. P., Graner, A. and Geiger, H. H. (2010), ‘Population structure and phenotypic variation of a spring barley world collection set up for association studies’, *Plant Breeding* **129**, 271–279.
- Hayut, S. F., Bessudo, C. M. and Levy, A. A. (2017), ‘Targeted recombination between homologous chromosomes for precise breeding in tomato’, *Nature Communications* **8**, 1–9.
- He, L. and Dooner, H. K. (2009), ‘Haplotype structure strongly affects recombination in a maize genetic interval polymorphic for Helitron and retrotransposon insertions’, *Proceedings of the National Academy of Sciences of the United States of America* **106**, 8410–8416.
- He, Y., Sidhu, G. and Pawlowski, W. P. (2013), *Plant Meiosis*, Vol. 990, Humana Press.
- Hellsten, U., Wright, K. M., Jenkins, J., Shu, S., Yuan, Y., Wessler, S. R., Schmutz, J., Willis, J. H. and Rokhsar, D. S. (2013), ‘Fine-scale variation in meiotic recombination in *Mimulus* inferred from population shotgun sequencing’, *Proceedings of the National Academy of Sciences of the United States of America* **110**, 19478–19482.
- Henderson, I. R. (2012), ‘Control of meiotic recombination frequency in plant genomes’, *Current Opinion in Plant Biology* **15**, 556–561.
- Herickhoff, L., Stack, S. and Sherman, J. (1993), ‘The relationship between synapsis, recombination nodules and chiasmata in tomato translocation heterozygotes’, *Heredity* **71**, 373–385.
- Higgins, E. E., Howell, E. C., Armstrong, S. J. and Parkin, I. A. (2020), ‘A major quantitative trait locus on chromosome A9, BnaPh1, controls homoeologous recombination in *Brassica napus*’, *New Phytologist* pp. 3281–3293.
- Higgins, J. D., Armstrong, S. J., Franklin, F. C. H. and Jones, G. H. (2004), ‘The Arabidopsis MutS homolog AtMSH4 functions at an early step in recombination: Evidence for two classes of recombination in Arabidopsis’, *Genes and Development* **18**, 2557–2570.

- Higgins, J. D., Buckling, E. F., Franklin, F. C. H. and Jones, G. H. (2008b), ‘Expression and functional analysis of AtMUS81 in Arabidopsis meiosis reveals a role in the second pathway of crossing-over’, *Plant Journal* **54**, 152–162.
- Higgins, J. D., Osman, K., Jones, G. H. and Franklin, F. C. H. (2014), ‘Factors underlying restricted crossover localization in barley meiosis’, *Annual Review of Genetics* **48**, 29–47.
- Higgins, J. D., Perry, R. M., Barakate, A., Ramsay, L., Waugh, R., Halpin, C., Armstrong, S. J. and Franklin, F. C. H. (2012), ‘Spatiotemporal asymmetry of the meiotic program underlies the predominantly distal distribution of meiotic crossovers in barley’, *Plant Cell* **24**, 4096–4109.
- Higgins, J. D., Sanchez-Moran, E., Armstrong, S. J., Jones, G. H. and Franklin, F. C. H. (2005), ‘The Arabidopsis synaptonemal complex protein ZYP1 is required for chromosome synapsis and normal fidelity of crossing over’, *Genes and Development* **19**, 2488–2500.
- Higgins, J. D., Vignard, J., Mercier, R., Pugh, A. G., Franklin, F. C. H. and Jones, G. H. (2008a), ‘AtMSH5 partners AtMSH4 in the class I meiotic crossover pathway in Arabidopsis thaliana, but is not required for synapsis’, *Plant Journal* **55**, 28–39.
- Hill, W. G. and Robertson, A. (1966), ‘The effect of linkage on limits to artificial selection’, *Genetical Research* **8**, 269–294.
- Hoffmann, A. A. and Rieseberg, L. H. (2008), ‘Revisiting the impact of inversions in evolution: From population genetic markers to drivers of adaptive shifts and speciation?’, *Annual Review of Ecology, Evolution, and Systematics* **39**, 21–42.
- Holliday, R. (1964), ‘A mechanism for gene conversion in fungi’, *Genetics Research* **89**, 285–307.
- Horton, M. W., Hancock, A. M., Huang, Y. S., Toomajian, C., Atwell, S., Auton, A., Mulyati, N. W., Platt, A., Sperone, F. G., Vilhjálmsson, B. J., Nordborg, M., Borevitz, J. O. and Bergelson, J. (2012), ‘Genome-wide patterns of genetic variation in worldwide Arabidopsis thaliana accessions from the RegMap panel’, *Nature Genetics* **44**, 212–216.
- Hsu, S. J., Erickson, R. P., Zhang, J., Garver, W. S. and Heidenreich, R. A. (2000), ‘Fine linkage and physical mapping suggests cross-over suppression with a retroposon insertion at the npc1 mutation’, *Mammalian Genome* **11**, 774–778.

- Hsu, Y.-M., Falque, M. and Martin, O. C. (2022), ‘Quantitative modelling of fine-scale variations in the *Arabidopsis thaliana* crossover landscape’, *Quantitative Plant Biology* **3**, e3.
- Hu, Q., Li, Y., Wang, H., Shen, Y., Zhang, C., Du, G., Tang, D. and Cheng, Z. (2017), ‘Meiotic chromosome association 1 interacts with TOP3 α and regulates meiotic recombination in rice’, *Plant Cell* **29**, 1697–1708.
- Huang, B. E. and George, A. W. (2011), ‘R/mpMap: A computational platform for the genetic analysis of multiparent recombinant inbred lines’, *Bioinformatics* **27**, 727–729.
- Hufford, M. B., Xu, X., Heerwaarden, J. V., Pyhäjärvi, T., Chia, J. M., Cartwright, R. A., Elshire, R. J., Glaubitz, J. C., Guill, K. E., Kaeppeler, S. M., Lai, J., Morrell, P. L., Shannon, L. M., Song, C., Springer, N. M., Swanson-Wagner, R. A., Tiffin, P., Wang, J., Zhang, G., Doebley, J., McMullen, M. D., Ware, D., Buckler, E. S., Yang, S. and Ross-Ibarra, J. (2012), ‘Comparative population genomics of maize domestication and improvement’, *Nature Genetics* **44**, 808–811.
- Hunter, N. (2015), ‘Meiotic recombination: The essence of heredity’, *Cold Spring Harbor Perspectives in Biology* **7**, 1–36.
- Huxley, J. S. (1928), ‘Sexual difference of linkage in *Gammarus chevreuxi*’, *Journal of Genetics* **20**, 145–156.
- Hyppa, R. W. and Smith, G. R. (2010), ‘Crossover Invariance Determined by Partner Choice for Meiotic DNA Break Repair’, *Cell* **142**, 243–255.
- Iwata, H. and Ninomiya, S. (2006), ‘AntMap: Constructing Genetic Linkage Maps Using an Ant Colony Optimization Algorithm’, *Breeding Science* **56**, 371–377.
- Jayabalan, N. and Rao, G. R. (1987), ‘Effect of physical and chemical mutagens on chiasma frequency in *Lycopersicon esculentum* mill’, *Caryologia* **40**, 115–121.
- Jayakodi, M., Padmarasu, S., Haberer, G., Bonthala, V. S., Gundlach, H., Monat, C., Lux, T., Kamal, N., Lang, D., Himmelbach, A., Ens, J., Zhang, X. Q., Angessa, T. T., Zhou, G., Tan, C., Hill, C., Wang, P., Schreiber, M., Boston, L. B., Plott, C., Jenkins, J., Guo, Y., Fiebig, A., Budak, H., Xu, D., Zhang, J., Wang, C., Grimwood, J., Schmutz, J., Guo, G., Zhang, G., Mochida, K., Hirayama, T., Sato, K., Chalmers, K. J., Langridge, P., Waugh, R., Pozniak, C. J., Scholz, U., Mayer, K. F., Spannagl, M., Li, C., Mascher, M. and Stein, N.

- (2020), ‘The barley pan-genome reveals the hidden legacy of mutation breeding’, *Nature* **588**, 284–289.
- Jeffreys, A. J. and Neumann, R. (2005), ‘Factors influencing recombination frequency and distribution in a human meiotic crossover hotspot’, *Human Molecular Genetics* **14**, 2277–2287.
- Jones, G. H. (1984), ‘The control of chiasma distribution.’, *Symposia of the Society for Experimental Biology* **38**, 293–320.
- Jordan, K. W., Wang, S., He, F., Chao, S., Lun, Y., Paux, E., Sourdille, P., Sherman, J., Akhunova, A., Blake, N. K., Pumphrey, M. O., Glover, K., Dubcovsky, J., Talbert, L. and Akhunov, E. D. (2018), ‘The genetic architecture of genome-wide recombination rate variation in allopolyploid wheat revealed by nested association mapping’, *Plant Journal* **95**, 1039–1054.
- Kasamatsu, H., Robberson, D. L. and Vinograd, J. (1971), ‘A Novel Closed-Circular Mitochondrial DNA with Properties of a Replicating Intermediate’, *Proceedings of the National Academy of Sciences of the United States of America* **68**, 2252–2257.
- KASHA, K. J. and KAO, K. N. (1970), ‘High Frequency Haploid Production in Barley (*Hordeum vulgare* L.)’, *Nature* **225**, 874–876.
- Kearsey, M. J., Ramsay, L. D., Jennings, D. E., Lydiate, D. J., Bohuon, E. J. R. and Marshall, D. F. (1996), ‘Higher recombination frequencies in female compared to male meioses in *Brassica oleracea*’, *Theor Appl Genet* **92**, 363–367.
- Keeney, S., Giroux, C. N. and Kleckner, N. (1997), ‘Meiosis-Specific DNA Double-Strand Breaks Are Catalyzed by Spo11, a Member of a Widely Conserved Protein Family’, *Cell* **88**, 375–384.
- Keeney, S. and Neale, M. (2006), ‘Initiation of meiotic recombination by formation of DNA double-strand breaks: mechanism and regulation’, *Biochemical Society Transactions* **34**, 523–525.
- Kidwell, M. G. (1972), ‘Genetic change of recombination value in *Drosophila melanogaster*. I. Artificial Selection for high and low recombination and some properties of recombination-modifying genes’, *Genetics* **70**, 419–432.
- Kim, K. P., Weiner, B. M., Zhang, L., Jordan, A., Dekker, J. and Kleckner, N. (2010), ‘Sister cohesion and structural axis components mediate homolog bias of meiotic recombination’, *Cell* **143**, 924–937.

- Kim, S., Plagnol, V., Hu, T. T., Toomajian, C., Clark, R. M., Ossowski, S., Ecker, J. R., Weigel, D. and Nordborg, M. (2007), ‘Recombination and linkage disequilibrium in *Arabidopsis thaliana*’, *Nature Genetics* **39**, 1151–1155.
- King, J., Roberts, L. A., Kearsley, M. J., Thomas, H. M., Jones, R. N., Huang, L., Armstead, I. P., Morgan, W. G. and King, I. P. (2002), ‘A demonstration of a 1:1 correspondence between chiasma frequency and recombination using a *Lolium perenne*/*Festuca pratensis* substitution.’, *Genetics* **161**, 307–14.
- Kleckner, N. (2006), ‘Chiasma formation: Chromatin/axis interplay and the role(s) of the synaptonemal complex’, *Chromosoma* **115**, 175–194.
- Kleckner, N., Storlazzi, A. and Zickler, D. (2003), ‘Coordinate variation in meiotic pachytene SC length and total crossover/chiasma frequency under conditions of constant DNA length’, *Trends in Genetics* **19**, 623–628.
- Kleckner, N., Zickler, D., Jones, G. H., Dekker, J., Padmore, R., Henle, J. and Hutchinson, J. (2004), ‘A mechanical basis for chromosome function’, *Proceedings of the National Academy of Sciences* **101**, 12592–12597.
- Knoll, A., Higgins, J. D., Seeliger, K., Reha, S. J., Dangel, N. J., Bauknecht, M., Schröpfer, S., Franklin, F. C. H. and Puchta, H. (2012), ‘The fanconi anemia ortholog FANCM ensures ordered homologous recombination in both somatic and meiotic Cells in *Arabidopsis*’, *Plant Cell* **24**, 1448–1464.
- Knüpffer, H. and van Hintum, T. (2003), *Summarised diversity – the Barley Core Collection*, Elsevier Science, pp. 259–267.
- Koehler, K. E., Hawley, R. S., Sherman, S. and Hassold, T. (1996), ‘Recombination and nondisjunction in humans and flies’, *Human Molecular Genetics* **5**.
- Kong, A., Gudbjartsson, D. F., Sainz, J., Jonsdottir, G. M., Gudjonsson, S. A., Richardsson, B., Sigurdardottir, S., Barnard, J., Hallbeck, B., Masson, G., Shlien, A., Palsson, S. T., Frigge, M. L., Thorgeirsson, T. E., Gulcher, J. R. and Stefansson, K. (2002), ‘A high-resolution recombination map of the human genome’, *Nature Genetics* **31**, 241–247.
- Kong, A., Thorleifsson, G., Gudbjartsson, D. F., Masson, G., Sigurdsson, A., Jonasdottir, A., Walters, G. B., Jonasdottir, A., Gylfason, A., Kristinsson, K. T., Gudjonsson, S. A., Frigge, M. L., Helgason, A., Thorsteinsdottir, U. and Stefansson, K. (2010), ‘Fine-scale recombination rate differences between sexes, populations and individuals’, *Nature* **467**, 1099–1103.

- 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. and Stefansson, K. (2008), ‘Sequence Variants in the RNF212 Gene Associate with Genome-Wide Recombination Rate’, *Science* **319**, 1398–1401.
- Korzun, V., Plaschke, J., Börner, A. and Koebner, R. M. (1996), ‘Differences in recombination frequency during male and female gametogenesis in rye, *Secale cereale* L.’, *Plant Breeding* **115**, 422–424.
- Kosambi, D. D. (1943), ‘The estimation of map distances from recombination values’, *Annals of Eugenics* **12**, 172–175.
- Kovalchuk, I., Kovalchuk, O., Kalck, V., Boyko, V., Filkowski, J., Heinlein, M. and Hohn, B. (2003), ‘Pathogen-induced systemic plant signal triggers DNA rearrangements’, *Nature* **423**, 760–762.
- Kulathinal, R. J., Bennett, S. M., Fitzpatrick, C. L., Noor, M. A. F. and Schaeffer, S. W. (2008), ‘Fine-scale mapping of recombination rate in *Drosophila* refines its correlation to diversity and divergence’, *Proceedings of the National Academy of Sciences of the United States of America* **105**, 10051–10056.
- Kumar, S. V. and Wigge, P. A. (2010), ‘H2A.Z-Containing Nucleosomes Mediate the Thermosensory Response in *Arabidopsis*’, *Cell* **140**, 136–147.
- Künzel, G., Korzun, L. and Meister, A. (2000), ‘Cytologically integrated physical restriction fragment length polymorphism maps for the barley genome based on translocation breakpoints’, *Genetics* **154**, 397–412.
- Lagercrantz, U. and Lydiate, D. J. (1995), ‘RFLP mapping in *Brassica nigra* indicates differing recombination rates in male and female meioses’, *Genome* **38**, 255–264.
- Lam, S. Y., Horn, S. R., Radford, S. J., Housworth, E. A., Stahl, F. W. and Copenhaver, G. P. (2005), ‘Crossover interference on nucleolus organizing region-bearing chromosomes in *Arabidopsis*’, *Genetics* **170**, 807–812.
- Lambing, C., Tock, A. J., Topp, S. D., Choi, K., Kuo, P. C., Zhao, X., Osman, K., Higgins, J. D., Franklin, F. C. H. and Henderson, I. R. (2020), ‘Interacting Genomic Landscapes of REC8-Cohesin, Chromatin, and Meiotic Recombination in *Arabidopsis*’, *The Plant Cell* **32**, 1218–1239.

- Lander, E. S. and Botstein, D. (1989), ‘Mapping mendelian factors underlying quantitative traits using RFLP linkage maps.’, *Genetics* **121**, 185–199.
- Lander, E. S., Green, P., Abrahamson, J., Barlow, A., Daly, M. J., Lincoln, S. E. and Newburg, L. (1987), ‘MAPMAKER: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations’, *Genomics* **1**, 174–181.
- Lander, L. and Ryman, N. (1974), ‘Genetic control of recombination in *Neurospora crassa* II. Polygenic and maternal control in two proximal chromosome regions—a quantitative analysis’, *Hereditas* **78**, 169–184.
- Lange, K., Zhao, H. and Speed, T. P. (1997), ‘The Poisson-Skip Model of Crossing-Over’, *The Annals of Applied Probability* **7**, 299–313.
- Law, J. A. and Jacobsen, S. E. (2010), ‘Establishing, maintaining and modifying DNA methylation patterns in plants and animals’, *Nature Reviews Genetics* **11**, 204–220.
- Lawrence, E. J., Griffin, C. H. and Henderson, I. R. (2017), ‘Modification of meiotic recombination by natural variation in plants’, *Journal of Experimental Botany* **68**, 5471–5483.
- Lee, M., Sharopova, N., Beavis, W. D., Grant, D., Katt, M., Blair, D. and Hallauer, A. (2002), ‘Expanding the genetic map of maize with the intermated B73 x Mo17 (IBM) population’, *Plant Molecular Biology* **48**, 453–461.
- Leflon, M., Grandont, L., Eber, F., Huteau, V., Coriton, O., Chelysheva, L., Jenczewski, E. and Chèvre, A. M. (2010), ‘Crossovers get a boost in Brassica allotriploid and allotetraploid hybrids’, *Plant Cell* **22**, 2253–2264.
- Leitch, I. J. and Heslop-Harrison, J. S. (1993), ‘Physical mapping of four sites of 5S rDNA sequences and one site of the α -amylase-2 gene in barley (*Hordeum vulgare*)’, *Genome* **36**, 517–523.
- Lenormand, T. and Dutheil, J. (2005), ‘Recombination difference between sexes: A role for haploid selection’, *PLoS Biology* **3**, 0396–0403.
- Lewontin, R. C. and Kojima, K. (1960), ‘The evolutionary dynamics of complex polymorphisms’, *Evolution* **14**, 458–472.
- Lhuissier, F. G., Offenberg, H. H., Wittich, P. E., Vischer, N. O. and Heyting, C. (2007), ‘The mismatch repair protein MLH1 marks a subset of strongly interfering crossovers in tomato’, *Plant Cell* **19**, 862–876.

- Li, D. and Pfeiffer, T. W. (2009), ‘Three cycles of recurrent selection for altered recombination frequency in maize’, *Crop Science* **49**, 473–482.
- Li, N. and Stephens, M. (2003), ‘Modeling linkage disequilibrium and identifying recombination hotspots using single-nucleotide polymorphism data.’, *Genetics* **165**, 2213–33.
- Li, X., Li, L. and Yan, J. (2015), ‘Dissecting meiotic recombination based on tetrad analysis by single-microspore sequencing in maize’, *Nature Communications* **6**.
- Lindahl, K. F. (1991), ‘His and hers recombinational hotspots’, *Trends in Genetics* **7**, 273–276.
- Liu, Z. and Makaroff, C. A. (2006), ‘Arabidopsis separase AESP is essential for embryo development and the release of cohesin during meiosis’, *Plant Cell* **18**, 1213–1225.
- Lloyd, A., Morgan, C., Franklin, F. C. H. and Bomblies, K. (2018), ‘Plasticity of meiotic recombination rates in response to temperature in Arabidopsis’, *Genetics* **208**, 1409–1420.
- Lu, P., Han, X., Qi, J., Yang, J., Wijeratne, A. J., Li, T. and Ma, H. (2012), ‘Analysis of Arabidopsis genome-wide variations before and after meiosis and meiotic recombination by resequencing Landsberg erecta and all four products of a single meiosis’, *Genome Research* **22**, 508–518.
- Lucht, J. M., Mauch-Mani, B., Steiner, H. Y., Metraux, J. P., Ryals, J. and Hohn, B. (2002), ‘Pathogen stress increases somatic recombination frequency in Arabidopsis’, *Nature Genetics* **30**, 311–314.
- Luo, Q., Tang, D., Wang, M., Luo, W., Zhang, L., Qin, B., Shen, Y., Wang, K., Li, Y. and Cheng, Z. (2013), ‘The role of OsMSH5 in crossover formation during rice meiosis’, *Molecular Plant* **6**, 729–742.
- Lynn, A., Koehler, K. E., Judis, L., Chan, E. R., Cherry, J. P., Schwartz, S., Seftel, A., Hunt, P. A. and Hassold, T. J. (2002), ‘Covariation of Synaptonemal Complex Length and Mammalian Meiotic Exchange Rates’, *Science* **296**, 2222–2225.
- López, E., Pradillo, M., Oliver, C., Romero, C., Cuñado, N. and Santos, J. L. (2012), ‘Looking for natural variation in chiasma frequency in Arabidopsis thaliana’, *Journal of Experimental Botany* **63**, 887–894.

- Ma, J. and Bennetzen, J. L. (2006), ‘Recombination, rearrangement, reshuffling, and divergence in a centromeric region of rice’, *Proceedings of the National Academy of Sciences of the United States of America* **103**, 383–388.
- Macaisne, N., Novatchkova, M., Peirera, L., Vezon, D., Jolivet, S., Froger, N., Chelysheva, L., Grelon, M. and Mercier, R. (2008), ‘SHOC1, an XPF Endonuclease-Related Protein, Is Essential for the Formation of Class I Meiotic Crossovers’, *Current Biology* **18**, 1432–1437.
- Mace, E. S., Rami, J. F., Bouchet, S., Klein, P. E., Klein, R. R., Kilian, A., Wenzl, P., Xia, L., Halloran, K. and Jordan, D. R. (2009), ‘A consensus genetic map of sorghum that integrates multiple component maps and high-throughput Diversity Array Technology (DArT) markers’, *BMC Plant Biology* **9**.
- Maeda, T. (1930), *On the Configurations of Gemini in the Pollen Mother Cells of Vicia Faba, L.*, Botanical Institute, Science Department, Kyoto Imperial University.
- Maloisel, L. and Rossignol, J.-L. (1998), ‘Suppression of crossing-over by DNA methylation in *Ascomobolus*’, *Genes and Development* **12**, 1381–1389.
- Mancera, E., Bourgon, R., Brozzi, A., Huber, W. and Steinmetz, M. (2009), ‘High-resolution mapping of meiotic crossovers and noncrossovers in yeast’, *Nature* **454**, 479–485.
- Manly, K. F., Cudmore, R. H. and Meer, J. M. (2001), ‘Map Manager QTX, cross-platform software for genetic mapping’, *Mammalian Genome* **12**, 930–932.
- Mao, Z., Jiang, Y., Liu, X., Seluanov, A. and Gorbunova, V. (2009), ‘DNA repair by homologous recombination, but not by nonhomologous end joining, is elevated in breast cancer cells’, *Neoplasia* **11**, 683–691.
- Marand, A. P., Jansky, S. H., Zhao, H., Leisner, C. P., Zhu, X., Zeng, Z., Crisovan, E., Newton, L., Hamernik, A. J., Veilleux, R. E., Buell, C. R. and Jiang, J. (2017), ‘Meiotic crossovers are associated with open chromatin and enriched with Stowaway transposons in potato’, *Genome Biology* **18**.
- Marand, A. P., Zhao, H., Zhang, W., Zeng, Z., Fang, C. and Jianga, J. (2019), ‘Historical meiotic crossover hotspots fueled patterns of evolutionary divergence in rice’, *Plant Cell* **31**, 645–662.
- Marimuthu, M. P. A., Jolivet, S., Ravi, M., Pereira, L., Davda, J. N., Cromer, L., Wang, L., Nogu  , F., Chan, S. W. L., Siddiqi, I. and Mercier, R. (2011), ‘Synthetic Clonal Reproduction Through Seeds’, *Science* **331**, 876–876.

- Martin, G., Baurens, F. C., Hervouet, C., Salmon, F., Delos, J. M., Labadie, K., Perdereau, A., Mournet, P., Blois, L., Dupouy, M., Carreel, F., Ricci, S., Lemainque, A., Yahiaoui, N. and D'Hont, A. (2020), 'Chromosome reciprocal translocations have accompanied subspecies evolution in bananas', *Plant Journal* **104**, 1698–1711.
- Martini, E., Diaz, R. L., Hunter, N. and Keeney, S. (2006), 'Crossover Homeostasis in Yeast Meiosis', *Cell* **126**, 285–295.
- Mascher, M., Gundlach, H., Himmelbach, A., Beier, S., Twardziok, S. O., Wicker, T., Radchuk, V., Dockter, C., Hedley, P. E., Russell, J., Bayer, M., Ramsay, L., Liu, H., Haberer, G., Zhang, X. Q., Zhang, Q., Barrero, R. A., Li, L., Taudien, S., Groth, M., Felder, M., Hastie, A., Šimková, H., Stanková, H., Vrána, J., Chan, S., Munõz-Amatriaín, M., Ounit, R., Wanamaker, S., Bolser, D., Colmsee, C., Schmutzer, T., Aliyeva-Schnorr, L., Grasso, S., Tanskanen, J., Chailyan, A., Sampath, D., Heavens, D., Clissold, L., Cao, S., Chapman, B., Dai, F., Han, Y., Li, H., Li, X., Lin, C., McCooke, J. K., Tan, C., Wang, P., Wang, S., Yin, S., Zhou, G., Poland, J. A., Bellgard, M. I., Borisjuk, L., Houben, A., Doleael, J., Ayling, S., Lonardi, S., Kersey, P., Langridge, P., Muehlbauer, G. J., Clark, M. D., Caccamo, M., Schulman, A. H., Mayer, K. F., Platzer, M., Close, T. J., Scholz, U., Hansson, M., Zhang, G., Braumann, I., Spannagl, M., Li, C., Waugh, R. and Stein, N. (2017), 'A chromosome conformation capture ordered sequence of the barley genome', *Nature* **544**, 427–433.
- Mascher, M., Wicker, T., Jenkins, J., Plott, C., Lux, T., Koh, C. S., Ens, J., Gundlach, H., Boston, L. B., Tulpová, Z., Holden, S., Hernández-Pinzón, I., Scholz, U., Mayer, K. F., Spannagl, M., Pozniak, C. J., Sharpe, A. G., Šimková, H., Moscou, M. J., Grimwood, J., Schmutz, J. and Stein, N. (2021), 'Long-read sequence assembly: A technical evaluation in barley', *Plant Cell* **33**, 1888–1906.
- Matise, T. C., Chen, F., Chen, W., Vega, F. M. D. L., Hansen, M., He, C., Hyland, F. C., Kennedy, G. C., Kong, X., Murray, S. S., Ziegler, J. S., Stewart, W. C. and Buyske, S. (2007), 'A second-generation combined linkage-physical map of the human genome', *Genome Research* **17**, 1783–1786.
- Maurer, A., Draba, V., Jiang, Y., Schnaithmann, F., Sharma, R., Schumann, E., Kilian, B., Reif, J. C. and Pillen, K. (2015), 'Modelling the genetic architecture of flowering time control in barley through nested association mapping', *BMC Genomics* **16**.
- Maurer, H. P., Melchinger, A. E. and Frisch, M. (2008), 'Population genetic simulation and data analysis with Plabsoft', *Euphytica* **161**, 133–139.

- Mayer, K. F., Waugh, R., Langridge, P., Close, T. J., Wise, R. P., Graner, A., Matsumoto, T., Sato, K., Schulman, A., Ariyadasa, R., Schulte, D., Poursarebani, N., Zhou, R., Steuernagel, B., Mascher, M., Scholz, U., Shi, B., Madishetty, K., Svensson, J. T., Bhat, P., Moscou, M., Resnik, J., Muehlbauer, G. J., Hedley, P., Liu, H., Morris, J., Frenkel, Z., Korol, A., Bergès, H., Taudien, S., Felder, M., Groth, M., Platzner, M., Himmelbach, A., Lonardi, S., Duma, D., Alpert, M., Cordero, F., Beccuti, M., Ciardo, G., Ma, Y., Wanamaker, S., Cattonaro, F., Vendramin, V., Scalabrin, S., Radovic, S., Wing, R., Morgante, M., Nussbaumer, T., Gundlach, H., Martis, M., Poland, J., Pfeifer, M., Moisy, C., Tanskanen, J., Zuccolo, A., Spannagl, M., Russell, J., Druka, A., Marshall, D., Bayer, M., Swarbreck, D., Sampath, D., Ayling, S., Febrer, M., Caccamo, M., Tanaka, T., Wannamaker, S., Schmutzer, T., Brown, J. W., Fincher, G. B. and Stein, N. (2012), ‘A physical, genetic and functional sequence assembly of the barley genome’, *Nature* **491**, 711–716.
- McKim, K. S., Howell, A. M. and Rose, A. M. (1988), ‘The effects of translocations on recombination frequency in *Caenorhabditis elegans*.’, *Genetics* **120**, 987–1001.
- McMahill, M. S., Sham, C. W. and Bishop, D. K. (2007), ‘Synthesis-dependent strand annealing in meiosis’, *PLoS Biology* **5**, 2589–2601.
- McMullen, M. D., Kresovich, S., Villeda, H. S., Bradbury, P., Li, H., Sun, Q., Flint-Garcia, S., Thornsberry, J., Acharya, C., Bottoms, C., Brown, P., Browne, C., Eller, M., Guill, K., Harjes, C., Kroon, D., Lepak, N., Mitchell, S. E., Peterson, B., Pressoir, G., Romero, S., Rosas, M. O., Salvo, S., Yates, H., Hanson, M., Jones, E., Smith, S., Glaubitz, J. C., Goodman, M., Ware, D., Holland, J. B. and Buckler, E. S. (2009), ‘Genetic properties of the maize nested association mapping population’, *Science* **325**, 737–740.
- McPeck, M. S. and Speed, T. P. (1995), ‘Modeling interference in genetic recombination.’, *Genetics* **139**, 1031–1044.
- McVean, G. A. T., Myers, S. R., Hunt, S., Deloukas, P., Bentley, D. R. and Donnelly, P. (2004), ‘The Fine-Scale Structure of Recombination Rate Variation in the Human Genome’, *Science* **304**, 581–584.
- Mcvean, G., Awadalla, P. and Fearnhead, P. (2002), ‘A Coalescent-Based Method for Detecting and Estimating Recombination From Gene Sequences’, *Genetics* **160**(3), 1231–1241.

- Melamed-Bessudo, C. and Levy, A. A. (2012), ‘Deficiency in DNA methylation increases meiotic crossover rates in euchromatic but not in heterochromatic regions in *Arabidopsis*’, *Proceedings of the National Academy of Sciences of the United States of America* **109**.
- Melamed-Bessudo, C., Shilo, S. and Levy, A. A. (2016), ‘Meiotic recombination and genome evolution in plants’, *Current Opinion in Plant Biology* **30**, 82–87.
- Melamed-Bessudo, C., Yehuda, E., Stuitje, A. R. and Levy, A. A. (2005), ‘A new seed-based assay for meiotic recombination in *Arabidopsis thaliana*’, *Plant Journal* **43**, 458–466.
- Mercier, R., Jolivet, S., Vezon, D., Huppe, E., Chelysheva, L., Giovanni, M., Nogu  , F., Doutriaux, M. P., Horlow, C., Grelon, M. and M  zard, C. (2005), ‘Two meiotic crossover classes cohabit in *Arabidopsis*: One is dependent on MER3, whereas the other one is not’, *Current Biology* **15**, 692–701.
- Mercier, R., M  zard, C., Jenczewski, E., Macaisne, N. and Grelon, M. (2015), ‘The molecular biology of meiosis in plants’, *Annual Review of Plant Biology* **66**, 297–327.
- Miao, C., Tang, D., Zhang, H., Wang, M., Li, Y., Tang, S., Yu, H., Gu, M. and Cheng, Z. (2013), ‘Central region component1, a novel synaptonemal complex component, Is essential for meiotic recombination initiation in RiceC’, *Plant Cell* **25**, 2998–3009.
- Mieulet, D., Aubert, G., Bres, C., Klein, A., Droc, G., Vieille, E., Rond-Coissieux, C., Sanchez, M., Dalmais, M., Mauxion, J. P., Rothan, C., Guiderdoni, E. and Mercier, R. (2018), ‘Unleashing meiotic crossovers in crops’, *Nature Plants* **4**, 1010–1016.
- Mirouze, M., Lieberman-Lazarovich, M., Aversano, R., Bucher, E., Nicolet, J., Reinders, J. and Paszkowski, J. (2012), ‘Loss of DNA methylation affects the recombination landscape in *Arabidopsis*’, *Proceedings of the National Academy of Sciences of the United States of America* **109**, 5880–5885.
- Modliszewski, J. L., Wang, H., Albright, A. R., Lewis, S. M., Bennett, A. R., Huang, J., Ma, H., Wang, Y. and Copenhaver, G. P. (2018), ‘Elevated temperature increases meiotic crossover frequency via the interfering (Type I) pathway in *Arabidopsis thaliana*’, *PLoS Genetics* **14**, 1–15.
- Moens, P. B., Chen, D. J., Shen, Z., Kolas, N., Tarsounas, M., Heng, H. H. Q. and Spyropoulos, B. (1997), ‘Rad51 immunocytology in rat and mouse spermatocytes and oocytes’, *Chromosoma* **106**, 207–215.

- Monat, C., Padmarasu, S., Lux, T., Wicker, T., Gundlach, H., Himmelbach, A., Ens, J., Li, C., Muehlbauer, G. J., Schulman, A. H., Waugh, R., Braumann, I., Pozniak, C., Scholz, U., Mayer, K. F., Spannagl, M., Stein, N. and Mascher, M. (2019), ‘TRITEX: Chromosome-scale sequence assembly of Triticeae genomes with open-source tools’, *bioRxiv* pp. 1–18.
- Morgan, H. (1916), *A Critique of the Theory of Evolution*, Princeton University Press.
- Morgan, T. H. (1911), ‘Random segregation versus coupling in mendelian inheritance’, *Science* **34**, 384–384.
- Morgan, T. H. (1912), ‘Complete Linkage in the Second Chromosome of the Male of *Drosophila*’, *Science* **36**, 719–720.
- Morgan, T. H. (1914), ‘No crossing over in the male of *Drosophila* of genes in the second and third pairs of chromosomes’, *The Biological Bulletin* **26**, 195–204.
- Morris, G. P., Ramu, P., Deshpande, S. P., Hash, C. T., Shah, T., Upadhyaya, H. D., Riera-Lizarazu, O., Brown, P. J., Acharya, C. B., Mitchell, S. E., Hariman, J., Glaubitz, J. C., Buckler, E. S. and Kresovich, S. (2013), ‘Population genomic and genome-wide association studies of agroclimatic traits in sorghum’, *Proceedings of the National Academy of Sciences of the United States of America* **110**, 453–458.
- Mortimer, R. K. and Fogel, S. (1974), *Genetical Interference and Gene Conversion*, Springer US, pp. 263–275.
- Muller, H. J. (1916), ‘The Mechanism of Crossing-Over’, *The American Naturalist* **50**, 193–221.
- Muller, H. J. (1932), ‘Some Genetic Aspects of Sex’, *The American Naturalist* **66**, 118–138.
- Muyt, A. D., Pereira, L., Vezon, D., Chelysheva, L., Gendrot, G., Chambon, A., Lainé-Choinard, S., Pelletier, G., Mercier, R., Nogué, F. and Grelon, M. (2009), ‘A high throughput genetic screen identifies new early meiotic recombination functions in *Arabidopsis thaliana*’, *PLoS Genetics* **5**.
- Myers, S., Bottolo, L., Freeman, C., McVean, G. and Donnelly, P. (2005), ‘Genetics: A fine-scale map of recombination rates and hotspots across the human genome’, *Science* **310**, 321–324.

- Myers, S., Bowden, R., Tumian, A., Bontrop, R. E., Freeman, C., MacFie, T. S., McVean, G. and Donnelly, P. (2010), ‘Drive against hotspot motifs in primates implicates the PRDM9 gene in meiotic recombination’, *Science* **327**, 876–879.
- Nachman, M. W. (2002), ‘Variation in recombination rate across the genome: Evidence and implications’, *Current Opinion in Genetics and Development* **12**, 657–663.
- Nagaoka, S. I., Hassold, T. J. and Hunt, P. A. (2012), ‘Human aneuploidy: Mechanisms and new insights into an age-old problem’, *Nature Reviews Genetics* **13**, 493–504.
- Neale, M. J., Pan, J. and Keeney, S. (2005), ‘Endonucleolytic processing of covalent protein-linked double-strand breaks’, *Nature* **436**, 1053–1057.
- Nicolas, A., Treco, D., Schultes, N. P. and Szostak, J. W. (1989), ‘An initiation site for meiotic gene conversion in the yeast *Saccharomyces cerevisiae*’, *Nature* **338**, 35–39.
- Nilsson, N. and Pelger, S. (1991), ‘The relationship between natural variation in chiasma frequencies and recombination frequencies in barley’, *Hereditas* **115**, 121–126.
- Nonomura, K. I., Nakano, M., Fukuda, T., Eiguchi, M., Miyao, A., Hirochika, H. and Kurata, N. (2004a), ‘The novel gene Homologous Pairing Aberration In Rice Meiosis1 of rice encodes a putative coiled-coil protein required for homologous chromosome pairing in meiosis’, *Plant Cell* **16**, 1008–1020.
- Nonomura, K. I., Nakano, M., Murata, K., Miyoshi, K., Eiguchi, M., Miyao, A., Hirochika, H. and Kurata, N. (2004b), ‘An insertional mutation in the rice PAIR2 gene, the ortholog of Arabidopsis ASY1, results in a defect in homologous chromosome pairing during meiosis’, *Molecular Genetics and Genomics* **271**, 121–129.
- Nordborg, M. and Weigel, D. (2008), ‘Next-generation genetics in plants’, *Nature* **456**, 720–723.
- Okagaki, R. J. and Weil, C. F. (1997), ‘Analysis of Recombination Sites Within the Maize waxy Locus’, *Genetics* **147**, 815–821.
- Ortiz-Barrientos, D., Engelstädter, J. and Rieseberg, L. H. (2016), ‘Recombination Rate Evolution and the Origin of Species’, *Trends in Ecology and Evolution* **31**, 226–236.

- Otto, S. P. and Lenormand, T. (2002), ‘Resolving the paradox of sex and recombination’, *Nature Reviews Genetics* **3**, 252–261.
- Paape, T., Zhou, P., Branca, A., Briskine, R., Young, N. and Tiffin, P. (2012), ‘Fine-scale population recombination rates, hotspots, and correlates of recombination in the *Medicago truncatula* genome’, *Genome Biology and Evolution* **4**, 726–737.
- Padmore, R., Cao, L. and Kleckner, N. (1991), ‘Temporal comparison of recombination and synaptonemal complex formation during meiosis in *S. cerevisiae*’, *Cell* **66**, 1239–1256.
- Pal, J., Bertheau, R., Buon, L., Qazi, A., Batchu, R. B., Bandyopadhyay, S., Ali-Fehmi, R., Beer, D. G., Weaver, D. W., Reis, R. J. S., Goyal, R. K., Huang, Q., Munshi, N. C. and Shammash, M. A. (2011), ‘Genomic evolution in Barrett’s adenocarcinoma cells: Critical roles of elevated hRAD51, homologous recombination and Alu sequences in the genome’, *Oncogene* **30**, 3585–3598.
- Pan, J., Sasaki, M., Kniewel, R., Murakami, H., Blitzblau, H. G., Tischfield, S. E., Zhu, X., Neale, M. J., Jasin, M., Socci, N. D., Hochwagen, A. and Keeney, S. (2011), ‘A hierarchical combination of factors shapes the genome-wide topography of yeast meiotic recombination initiation’, *Cell* **144**, 719–731.
- Panizza, S., Mendoza, M., Berlinger, M., Huang, L., Nicolas, A., Shirahige, K. and Klein, F. (2011), ‘Spo11-Accessory Proteins Link Double-Strand Break Sites to the Chromosome Axis in Early Meiotic Recombination’, *Cell* **146**, 372–383.
- Parker, J. S. (1975), ‘Aneuploidy and Isolation in Two *Hypochoeris* Species’, *Chromosoma* **52**, 89–101.
- Parniske, M., Hammond-Kosack, K. E., Golstein, C., Thomas, C. M., Jones, D. A., Harrison, K., Wulff, B. B. and Jones, J. D. (1997), ‘Novel Disease Resistance Specificities Result from Sequence Exchange between Tandemly Repeated Genes at the Cf-4/9 Locus of Tomato’, *Cell* **91**, 821–832.
- Pasam, R. K., Sharma, R., Malosetti, M., van Eeuwijk, F. A., Haseneyer, G., Kilian, B. and Graner, A. (2012), ‘Genome-wide association studies for agronomical traits in a world wide spring barley collection’, *BMC Plant Biology* **12**.
- Pawlowski, W. P., Golubovskaya, I. N. and Cande, W. Z. (2003), ‘Altered nuclear distribution of recombination protein RAD51 in maize mutants suggests the involvement of RAD51 in meiotic homology recognition’, *Plant Cell* **15**, 1807–1816.

- Pecinka, A., Fang, W., Rehmsmeier, M., Levy, A. A. and Scheid, O. M. (2011), ‘Polyploidization increases meiotic recombination frequency in Arabidopsis’, *BMC Biology* **9**.
- Peck, J. R. (1994), ‘A ruby in the rubbish: beneficial mutations, deleterious mutations and the evolution of sex.’, *Genetics* **137**, 597–606.
- Pedersen, C., Giese, H. and Linde-Laursen, I. (1995), ‘Towards an Integration of the Physical and the Genetic Chromosome Maps of Barley by in Situ Hybridization’, *Hereditas* **123**, 77–88.
- Perrella, G., Consiglio, M. F., Aiese-Cigliano, R., Cremona, G., Sanchez-Moran, E., Barra, L., Errico, A., Bressan, R. A., Franklin, F. C. H. and Conicella, C. (2010), ‘Histone hyperacetylation affects meiotic recombination and chromosome segregation in Arabidopsis’, *Plant Journal* **62**, 796–806.
- Pessia, E., Popa, A., Mousset, S., Rezvoy, C., Duret, L. and Marais, G. A. (2012), ‘Evidence for widespread GC-biased gene conversion in eukaryotes’, *Genome Biology and Evolution* **4**, 675–682.
- Petroli, C. D., Sansaloni, C. P., Carling, J., Steane, D. A., Vaillancourt, R. E., Myburg, A. A., da Silva, O. B., Pappas, G. J., Kilian, A. and Grattapaglia, D. (2012), ‘Genomic Characterization of DArT Markers Based on High-Density Linkage Analysis and Physical Mapping to the Eucalyptus Genome’, *PLoS ONE* **7**.
- Peñuela, M., Gallo-Franco, J. J., Finke, J., Rocha, C., Gkanogiannis, A., Ghneim-Herrera, T. and Lorieux, M. (2022), ‘Methylation in the CHH Context Allows to Predict Recombination in Rice’, *International Journal of Molecular Sciences* **23**.
- Phillips, D., Jenkins, G., Macaulay, M., Nibau, C., Wnetrzak, J., Fallding, D., Colas, I., Oakey, H., Waugh, R. and Ramsay, L. (2015), ‘The effect of temperature on the male and female recombination landscape of barley’, *New Phytologist* **208**, 421–429.
- Plough, H. H. (1917), ‘The effect of temperature on crossingover in Drosophila’, *Journal of Experimental Zoology* **24**, 147–209.
- Pook, T., Schlather, M. and Simianer, H. (2020), ‘MoBPS - Modular breeding program simulator’, *G3: Genes, Genomes, Genetics* **10**, 1915–1918.

- Powell, W., Caligari, P., Jinks, J. and Hayter, A. (1985), 'The use of doubled haploids in barley breeding. I. Comparison of Hi and H2 generations', *Heredity* **54**, 261–266.
- Preuss, D., Rhee, S. Y. and Davis, R. W. (1994), 'Tetrad Analysis Possible in Arabidopsis with Mutation of the QUARTET (QRT) Genes', *Science* **264**, 1458–1460.
- Prigge, V., Schipprack, W., Mahuku, G., Atlin, G. N. and Melchinger, A. E. (2012), 'Development of in vivo haploid inducers for tropical maize breeding programs', *Euphytica* **185**, 481–490.
- Puchta, H., Swoboda, P. and Hohn, B. (1995), 'Induction of intrachromosomal homologous recombination in whole plants', *The Plant Journal* **7**, 203–210.
- Qi, J., Chen, Y., Copenhaver, G. P. and Ma, H. (2014), 'Detection of genomic variations and DNA polymorphisms and impact on analysis of meiotic recombination and genetic mapping', *Proceedings of the National Academy of Sciences of the United States of America* **111**, 10007–10012.
- Rasmusson, J. (1927), 'Genetically changed linkage values in Pisum', *Hereditas* **10**, 1–152.
- Reddy, K. C. and Villeneuve, A. M. (2004), 'C. elegans HIM-17 Links Chromatin Modification and Competence for Initiation of Meiotic Recombination variety of systems has defined conserved core meiotic recombination machinery responsible for executing crossover recombination. A key player in this core ma', *Cell* **118**, 439–452.
- Reinders, J. and Paszkowski, J. (2009), 'Unlocking the Arabidopsis epigenome', *Epigenetics* **4**, 557–563.
- Rey, M. D., Martín, A. C., Smedley, M., Hayta, S., Harwood, W., Shaw, P. and Moore, G. (2018), 'Magnesium increases homoeologous crossover frequency during meiosis in ZIP4 (Ph1 gene) mutant wheat-wild relative hybrids', *Frontiers in Plant Science* **9**.
- Rhoades, M. M. (1941), 'Different Rates of Crossing over in Male and Female Gametes of Maize', *Agronomy Journal* **33**, 603–615.
- Rice, W. R. (1987), 'The Accumulation of Sexually Antagonistic Genes as a Selective Agent Promoting the Evolution of Reduced Recombination between Primitive Sex Chromosomes', *Source: Evolution* **41**, 911–914.

- Ritz, K. R., Noor, M. A. and Singh, N. D. (2017), ‘Variation in Recombination Rate: Adaptive or Not?’, *Trends in Genetics* **33**, 364–374.
- Robbins, T. P., Walker, E. L., Kermicle, J. L., Alleman, M. and Dellaporta, S. L. (1991), ‘Meiotic instability of the R-r complex arising from displaced intragenic exchange and intrachromosomal rearrangement.’, *Genetics* **129**, 271–283.
- Robertson, D. S. (1984), ‘Different frequency in the recovery of crossover products from male and female translocations in maize gametes of plants hypoploid from B-A’, *Genetics* **107**, 117–130.
- Rockman, M. V. and Kruglyak, L. (2009), ‘Recombinational landscape and population genomics of *Caenorhabditis elegans*’, *PLoS Genetics* **5**.
- Rodgers-Melnick, E., Bradbury, P. J., Elshire, R. J., Glaubitz, J. C., Acharya, C. B., Mitchell, S. E., Li, C., Li, Y. and Buckler, E. S. (2015), ‘Recombination in diverse maize is stable, predictable, and associated with genetic load’, *Proceedings of the National Academy of Sciences of the United States of America* **112**, 3823–3828.
- Ronin, Y. I., Mester, D. I., Minkov, D. G., Akhunov, E. and Korol, A. B. (2017), ‘Building ultra-high-density linkage maps based on efficient filtering of trustable markers’, *Genetics* **206**, 1285–1295.
- Rothenberg, M., Kohli, J. and Ludin, K. (2009), ‘Ctp1 and the MRN-complex are required for endonucleolytic Rec12 removal with release of a single class of oligonucleotides in fission yeast’, *PLoS Genetics* **5**.
- Rowan, B. A., Heavens, D., Feuerborn, T. R., Tock, A. J., Henderson, I. R. and Weigel, D. (2019), ‘An Ultra High-Density *Arabidopsis thaliana* Crossover’, *Genetics* **213**, 771–787.
- Rowan, B. A., Patel, V., Weigel, D. and Schneeberger, K. (2015), ‘Rapid and inexpensive whole-genome genotyping-by-sequencing for crossover localization and fine-scale genetic mapping’, *G3: Genes, Genomes, Genetics* **5**, 385–398.
- Sadhu, M. J., Bloom, J. S., Day, L. and Kruglyak, L. (2016), ‘CRISPR-directed mitotic recombination enables genetic mapping without crosses’, *Science* **352**, 1113–1116.
- Saintenac, C., Falque, M., Martin, O. C., Paux, E., Feuillet, C. and Sourdille, P. (2009), ‘Detailed recombination studies along chromosome 3B provide new insights on crossover distribution in wheat (*Triticum aestivum* L.)’, *Genetics* **181**, 393–403.

- Saintenac, C., Faure, S., Remay, A., Choulet, F., Ravel, C., Paux, E., Balfourier, F., Feuillet, C. and Sourdille, P. (2011), ‘Variation in crossover rates across a 3-Mb contig of bread wheat (*Triticum aestivum*) reveals the presence of a meiotic recombination hotspot’, *Chromosoma* **120**, 185–198.
- Salomé, P. A., Bomblies, K., Fitz, J., Laitinen, R. A., Warthmann, N., Yant, L. and Weigel, D. (2012), ‘The recombination landscape in *Arabidopsis thaliana* F2 populations’, *Heredity* **108**, 447–455.
- Sanchez-Moran, E., Armstrong, S. J., Santos, J. L., Franklin, F. C. H. and Jones, G. H. (2002), ‘Variation in Chiasma Frequency Among Eight Accessions of *Arabidopsis thaliana*’, *Genetics* **162**, 1415–1422.
- Sanchez-Moran, E., Santos, J. L., Jones, G. H. and Franklin, F. C. H. (2007), ‘ASY1 mediates AtDMC1-dependent interhomolog recombination during meiosis in *Arabidopsis*’, *Genes and Development* **21**, 2220–2233.
- Sandor, C., Li, W., Coppieters, W., Druet, T., Charlier, C. and Georges, M. (2012), ‘Genetic variants in REC8, RNF212, and PRDM9 influence male recombination in cattle’, *PLoS Genetics* **8**.
- Sannemann, W., Huang, B. E., Mathew, B. and Léon, J. (2015), ‘Multi-parent advanced generation inter-cross in barley: high-resolution quantitative trait locus mapping for flowering time as a proof of concept’, *Molecular Breeding* **35**.
- Sarno, R., Vicq, Y., Uematsu, N., Luka, M., Lapierre, C., Carroll, D., Bastianelli, G., Serero, A. and Nicolas, A. (2017), ‘Programming sites of meiotic crossovers using Spo11 fusion proteins’, *Nucleic Acids Research* **45**, 1–2.
- Sasaki, A. and Iwasa, Y. (1987), ‘Optimal Recombination Rate in Fluctuating Environments’, *Genetics* **115**, 377–388.
- Sasaki, M., Lange, J. and Keeney, S. (2010), ‘Genome destabilization by homologous recombination in the germ line’, *Nature Reviews Molecular Cell Biology* **11**, 182–195.
- Sato, K., Close, T. J., Bhat, P., Muñoz-Amatriaín, M. and Muehlbauer, G. J. (2011), ‘Single nucleotide polymorphism mapping and alignment of recombinant chromosome substitution lines in barley’, *Plant and Cell Physiology* **52**, 728–737.
- Sato, K., Nankaku, N. and Takeda, K. (2009), ‘A high-density transcript linkage map of barley derived from a single population’, *Heredity* **103**, 110–117.

- Schiex, T. and Gaspin, C. (1997), ‘CARTHAGENE: constructing and joining maximum likelihood genetic maps.’, *Proceedings. International Conference on Intelligent Systems for Molecular Biology* **5**, 258–67.
- Schreiber, M., Gao, Y., Koch, N., Fuchs, J., Heckmann, S., Himmelbach, A., Börner, A., Özkan, H., Maurer, A., Stein, N., Mascher, M. and Dreissig, S. (2022), ‘Recombination Landscape Divergence between Populations is Marked by Larger Low-Recombining Regions in Domesticated Rye’, *Molecular Biology and Evolution* **39**.
- Sebastian, J., Ravi, M., Andreuzza, S., Panoli, A. P., Marimuthu, M. P. and Siddiqi, I. (2009), ‘The plant adherin AtSCC2 is required for embryogenesis and sister-chromatid cohesion during meiosis in Arabidopsis’, *Plant Journal* **59**, 1–13.
- Seguí-Simarro, J. M. (2021), *Doubled Haploid Technology*, Vol. 2288, Springer US.
- Serra, H., Choi, K., Zhao, X., Blackwell, A. R., Kim, J. and Henderson, I. R. (2018), ‘Interhomolog polymorphism shapes meiotic crossover within the Arabidopsis RAC1 and RPP13 disease resistance genes’, *PLoS Genetics* **14**.
- Serra, H., Lambing, C., Griffin, C. H., Topp, S. D., Nageswaran, D. C., Underwood, C. J., Ziolkowski, P. A., Séguéla-Arnaud, M., Fernandes, J. B., Mercier, R. and Henderson, I. R. (2018b), ‘Massive crossover elevation via combination of HEI10 and recq4a recq4b during Arabidopsis meiosis’, *Proceedings of the National Academy of Sciences of the United States of America* **115**, 2437–2442.
- Shalev, G. and Levy, A. A. (1997), ‘The Maize Transposable Element Ac Induces Recombination Between the Donor Site and an Homologous Ectopic Sequence’, *Genetics* **146**, 1143–1151.
- Shao, T., Tang, D., Wang, K., Wang, M., Che, L., Qin, B., Yu, H., Li, M., Gu, M. and Cheng, Z. (2011), ‘Osrec8 is essential for chromatid cohesion and metaphase i monopolar orientation in rice meiosis’, *Plant Physiology* **156**, 1386–1396.
- Shen, C., Wang, N., Huang, C., Wang, M., Zhang, X. and Lin, Z. (2019), ‘Population genomics reveals a fine-scale recombination landscape for genetic improvement of cotton’, *Plant Journal* **99**, 494–505.
- Shen, Y., Tang, D., Wang, K., Wang, M., Huang, J., Luo, W., Luo, Q., Hong, L., Li, M. and Cheng, Z. (2012), ‘ZIP4 in homologous chromosome synapsis and crossover formation in rice meiosis’, *Journal of Cell Science* **125**, 2581–2591.

- Shi, J., Wolf, S. E., Burke, J. M., Presting, G. G., Ross-Ibarra, J. and Dawe, R. K. (2010), ‘Widespread gene conversion in centromere cores’, *PLoS Biology* **8**.
- Shilo, S., Melamed-Bessudo, C., Dorone, Y., Barkai, N. and Levy, A. A. (2015), ‘DNA crossover motifs associated with epigenetic modifications delineate open chromatin regions in Arabidopsis’, *Plant Cell* **27**, 2427–2436.
- Si, W., Yuan, Y., Huang, J., Zhang, X., Zhang, Y., Zhang, Y., Tian, D., Wang, C., Yang, Y. and Yang, S. (2015), ‘Widely distributed hot and cold spots in meiotic recombination as shown by the sequencing of rice F2 plants’, *New Phytologist* **206**, 1491–1502.
- Silva-Junior, O. B. and Grattapaglia, D. (2015), ‘Genome-wide patterns of recombination, linkage disequilibrium and nucleotide diversity from pooled resequencing and single nucleotide polymorphism genotyping unlock the evolutionary history of Eucalyptus grandis’, *New Phytologist* **208**, 830–845.
- Sim, S. C., Durstewitz, G., Plieske, J., Wieseke, R., Ganai, M. W., van Deynze, A., Hamilton, J. P., Buell, C. R., Causse, M., Wijeratne, S. and Francis, D. M. (2012), ‘Development of a large snp genotyping array and generation of high-density genetic maps in tomato’, *PLoS ONE* **7**.
- Simchen, G. and Stamberg, J. (1969), ‘Fine and Coarse Controls of Genetic Recombination’, *Nature* **222**, 329–332.
- Singh, D. K., Andreuzza, S., Panoli, A. P. and Siddiqi, I. (2013), ‘AtCTF7 is required for establishment of sister chromatid cohesion and association of cohesin with chromatin during meiosis in Arabidopsis’, *BMC Plant Biology* **13**.
- Slavov, G. T., Difazio, S. P., Martin, J., Schackwitz, W., Muchero, W., Rodgers-Melnick, E., Lipphardt, M. F., Pennacchio, C. P., Hellsten, U., Pennacchio, L. A., Gunter, L. E., Ranjan, P., Vining, K., Pomraning, K. R., Wilhelm, L. J., Pellegrini, M., Mockler, T. C., Freitag, M., Gerald, A., El-Kassaby, Y. A., Mansfield, S. D., Cronk, Q. C., Douglas, C. J., Strauss, S. H., Rokhsar, D. and Tuskan, G. A. (2012), ‘Genome resequencing reveals multiscale geographic structure and extensive linkage disequilibrium in the forest tree Populus trichocarpa’, *New Phytologist* **196**, 713–725.
- Smagulova, F., Gregoret, I. V., Brick, K., Khil, P., Camerini-Otero, R. D. and Petukhova, G. V. (2011), ‘Genome-wide analysis reveals novel molecular features of mouse recombination hotspots’, *Nature* **472**, 375–378.
- Smith, J. M. (1978), *The Evolution of Sex*, Cambridge University Press.

- Smith, J. M. and Haigh, J. (1974), ‘The hitch-hiking effect of a favourable gene’, *Genetical Research* **23**, 23–35.
- Sommermeier, V., Béneut, C., Chaplais, E., Serrentino, M. E. and Borde, V. (2013), ‘Spp1, a Member of the Set1 Complex, Promotes Meiotic DSB Formation in Promoters by Tethering Histone H3K4 Methylation Sites to Chromosome Axes’, *Molecular Cell* **49**, 43–54.
- Sood, S., Dwivedi, S., Reddy, T. V., Prasanna, P. S. and Sharma, N. (2013), ‘Improving androgenesis-mediated doubled haploid production efficiency of FCV tobacco (*Nicotiana tabacum* L.) through in vitro colchicine application’, *Plant Breeding* **132**, 764–771.
- Sorrells, M. E., Gustafson, J. P., Somers, D., Chao, S., Benscher, D., Guedira-Brown, G., Huttner, E., Kilian, A., McGuire, P. E., Ross, K., Tanaka, J., Wenzl, P., Williams, K. and Qualset, C. O. (2011), ‘Reconstruction of the Synthetic W7984 × Opata M85 wheat reference population’, *Genome* **54**, 875–882.
- Spies, M. and Fishel, R. (2015), ‘Mismatch repair during homologous and homeologous recombination’, *Cold Spring Harbor Perspectives in Biology* **7**.
- Stacey, N. J., Kuromori, T., Azumi, Y., Roberts, G., Breuer, C., Wada, T., Maxwell, A., Roberts, K. and Sugimoto-Shirasu, K. (2006), ‘Arabidopsis SPO11-2 functions with SPO11-1 in meiotic recombination’, *Plant Journal* **48**, 206–216.
- Stack, S. M. and Anderson, L. K. (2002), ‘Crossing over as assessed by late recombination nodules is related to the pattern of synapsis and the distribution of early recombination nodules in maize’, *Chromosome Research* **10**, 329–345.
- Stack, S. M. and Soulliere, D. L. (1984), ‘The relation between synapsis and chiasma formation in *Rhoeo spathacea*’, *Chromosoma* **90**, 72–83.
- Stam, P. (1979), ‘Interference in genetic crossing over and chromosome mapping’, *Genetics* **92**, 573–594.
- Stam, P. (1993), ‘Construction of integrated genetic linkage maps by means of a new computer package: JOINMAP’, *Plant Journal* **3**, 739–744.
- Stein, N., Feuillet, C., Wicker, T., Schlagenhauf, E. and Keller, B. (2000), ‘Subgenome chromosome walking in wheat: A 450-kb physical contig in *Triticum monococcum* L. spans the Lr10 resistance locus in hexaploid wheat (*Triticum aestivum* L.)’, *Proceedings of the National Academy of Sciences of the United States of America* **97**, 13436–13441.

- Stein, N., Prasad, M., Scholz, U., Thiel, T., Zhang, H., Wolf, M., Kota, R., Varshney, R. K., Perovic, D., Grosse, I. and Graner, A. (2007), ‘A 1,000-loci transcript map of the barley genome: New anchoring points for integrative grass genomics’, *Theoretical and Applied Genetics* **114**, 823–839.
- Stephens, J. L., Brown, S. E., Lapitan, N. L. and Knudson, D. L. (2004), ‘Physical mapping of barley genes using an ultrasensitive fluorescence in situ hybridization technique’, *Genome* **47**, 179–189.
- Stevenson, M., Armstrong, S. J., Ford-Lloyd, B. V. and Jones, G. H. (1998), ‘Comparative analysis of crossover exchanges and chiasmata in *Allium cepa* x *fistulosum* after genomic in situ hybridization (GISH)’, *Chromosome Research* **6**, 567–574.
- Stevison, L. S., Hoehn, K. B. and Noor, M. A. (2011), ‘Effects of inversions on within- and between-species recombination and divergence’, *Genome Biology and Evolution* **3**, 830–841.
- Sturtevant, A. H. (1913), ‘The linear arrangement of six sex-linked factors in *Drosophila*, as shown by their mode of association’, *Journal of Experimental Zoology* **14**, 43–59.
- Sturtevant, A. H. (1915), ‘The behavior of the chromosomes as studied through linkage’, *Zeitschrift für Induktive Abstammungs- und Vererbungslehre* **13**, 234–287.
- Sturtevant, A. H. (1921), ‘A Case of Rearrangement of Genes in *Drosophila*’, *Proceedings of the National Academy of Sciences* **7**, 235–237.
- Sturtevant, A. H., Bridges, C. B. and Morgan, T. H. (1919), ‘The spatial relations of genes’, *Proceedings of the National Academy of Sciences* **5**, 168–173.
- Sun, H., Rowan, B. A., Flood, P. J., Brandt, R., Fuss, J., Hancock, A. M., Michelmore, R. W., Huettel, B. and Schneeberger, K. (2019), ‘Linked-read sequencing of gametes allows efficient genome-wide analysis of meiotic recombination’, *Nature Communications* **10**, 1–9.
- Sun, Y., Ambrose, J. H., Haughey, B. S., Webster, T. D., Pierrie, S. N., Muñoz, D. F., Wellman, E. C., Cherian, S., Lewis, S. M., Berchowitz, L. E. and Copenhaver, G. P. (2012), ‘Deep Genome-Wide Measurement of Meiotic Gene Conversion Using Tetrad Analysis in *Arabidopsis thaliana*’, *PLoS Genetics* **8**, e1002968.
- Sybenga, J. (1970), ‘Simultaneous negative and positive chiasma interference across break point in interchange heterozygotes’, *Genetica* **41**, 209–230.

- Szostak, J. W., Orr-Weaver, T. L., Rothstein, R. J. and Stahl, F. W. (1983), ‘The double-strand-break repair model for recombination’, *Cell* **33**, 25–35.
- Säll, T., Flink, J. and Bengtsson, B. O. (1990), ‘Genetic control of recombination in barley: Variation in recombination frequency measured with inversion heterozygotes’, *Hereditas* **112**, 157–170.
- Séguéla-Arnaud, M., Choinard, S., Larchevêque, C., Girard, C., Froger, N., Crismani, W. and Mercier, R. (2017), ‘RMI1 and TOP3 α limit meiotic CO formation through their C-terminal domains’, *Nucleic acids research* **45**, 1860–1871.
- Séguéla-Arnaud, M., Crismani, W., Larchevêque, C., Mazel, J., Froger, N., Choinard, S., Lemhemdi, A., Macaisne, N., Leene, J. V., Gevaert, K., Jaeger, G. D., Chelysheva, L. and Mercier, R. (2015), ‘Multiple mechanisms limit meiotic crossovers: TOP3 α and two BLM homologs antagonize crossovers in parallel to FANCM’, *Proceedings of the National Academy of Sciences of the United States of America* **112**, 4713–4718.
- Söderberg, R. J. and Berg, O. G. (2011), ‘Kick-starting the ratchet: The fate of mutators in an asexual population’, *Genetics* **187**, 1129–1137.
- Taagen, E., Bogdanove, A. J. and Sorrells, M. E. (2020), ‘Counting on Crossovers: Controlled Recombination for Plant Breeding’, *Trends in Plant Science* **25**, 455–465.
- Tan, Y. D. and Fornage, M. (2008), ‘Mapping functions’, *Genetica* **133**, 235–246.
- Tanaka, Y. (1913), ‘A study of mendelian factors in the silkworm, *Bombyx mori*’, *The journal of the College of Agriculture, Tohoku Imperial University*.
- Toyota, M., Matsuda, K., Kakutani, T., Morita, M. T. and Tasaka, M. (2011), ‘Developmental changes in crossover frequency in *Arabidopsis*’, *Plant Journal* **65**, 589–599.
- Underwood, C. J., Choi, K., Lambing, C., Zhao, X., Serra, H., Borges, F., Simorowski, J., Ernst, E., Jacob, Y., Henderson, I. R. and Martienssen, R. A. (2018), ‘Epigenetic activation of meiotic recombination near *Arabidopsis thaliana* centromeres via loss of H3K9me2 and non-CG DNA methylation’, *Genome Research* **28**, 519–531.
- Urry, L. A., Cain, M. L., Wasserman, S. A., Minorsky, P. V., Reece, J. B. and Campbell, N. A. (2017), *Campbell biology*, 11th edn, Pearson Education, Inc.

- Varas, J., Sánchez-Morán, E., Copenhaver, G. P., Santos, J. L. and Pradillo, M. (2015), ‘Analysis of the Relationships between DNA Double-Strand Breaks, Synaptonemal Complex and Crossovers Using the *Atfas1-4* Mutant’, *PLoS Genetics* **11**.
- Villiers, K., Dinglasan, E., Hayes, B. J. and Voss-Fels, K. P. (2022), ‘genomicSimulation: fast R functions for stochastic simulation of breeding programs’, *G3: Genes, Genomes, Genetics* **12**.
- Vojta, A., Dobrinic, P., Tadic, V., Bockor, L., Korac, P., Julg, B., Klasic, M. and Zoldos, V. (2016), ‘Repurposing the CRISPR-Cas9 system for targeted DNA methylation’, *Nucleic Acids Research* **44**, 5615–5628.
- Wang, J., Street, N. R., Scofield, D. G. and Ingvarsson, P. K. (2016), ‘Natural selection and recombination rate variation shape nucleotide polymorphism across the genomes of three related *Populus* species’, *Genetics* **202**, 1185–1200.
- Wang, K., Wang, C., Liu, Q., Liu, W. and Fu, Y. (2015), ‘Increasing the Genetic Recombination Frequency by Partial Loss of Function of the Synaptonemal Complex in Rice’, *Molecular Plant* **8**, 1295–1298.
- Wang, K., Wang, M., Tang, D., Shen, Y., Miao, C., Hu, Q., Lu, T. and Cheng, Z. (2012), ‘The role of rice HEI10 in the formation of meiotic crossovers’, *PLoS Genetics* **8**.
- Wang, M., Wang, K., Tang, D., Wei, C., Li, M., Shen, Y., Chi, Z., Gu, M. and Cheng, Z. (2010), ‘The central element protein ZEP1 of the synaptonemal complex regulates the number of crossovers during meiosis in rice’, *Plant Cell* **22**, 417–430.
- Wang, X., Xu, P., Ren, Y., Yin, L., Li, S., Wang, Y., Shi, Y., Li, H., Cao, X., Chi, X., Yu, T., Pandey, M. K., Varshney, R. K. and Yuan, M. (2020), ‘Genome-wide identification of meiotic recombination hot spots detected by SLAF in peanut (*Arachis hypogaea* L.)’, *Scientific Reports* **10**, 1–11.
- Wang, Z., Chen, M., Yang, H., Hu, Z., Yu, Y., Xu, H., Yan, S., Yi, K. and Li, J. (2023), ‘A simple and highly efficient strategy to induce both paternal and maternal haploids through temperature manipulation’, *Nature Plants* **9**, 699–705.
- Wei, F., Gobelman-Werner, K., Morroll, S. M., Kurth, J., Mao, L., Wing, R., Leister, D., Schulze-Lefert, P. and Wise, R. P. (1999), ‘The *Mla* (Powdery Mildew)

- Resistance Cluster Is Associated With Three NBS-LRR Gene Families and Suppressed Recombination Within a 240-kb DNA Interval on Chromosome 5S (1HS) of Barley', *Genetics* **153**, 1929–1948.
- Wei, F., Wing, R. A. and Wise, R. P. (2002), 'Genome dynamics and evolution of the Mla (powdery mildew) resistance locus in barley', *Plant Cell* **14**, 1903–1917.
- Wei, K. H., Mantha, A. and Bachtrog, D. (2020), 'The theory and applications of measuring broad-range and chromosome-wide recombination rate from allele frequency decay around a selected locus', *Molecular Biology and Evolution* **37**, 3654–3671.
- Weismann, A. F. L. (1886), *Über der Rückschritt in der Natur*, Freiburg i. B. : J.C.B. Mohr.
- Wicker, T., Yahiaoui, N. and Keller, B. (2007), 'Illegitimate recombination is a major evolutionary mechanism for initiating size variation in plant resistance genes', *Plant Journal* **51**, 631–641.
- Wijeratne, A. J., Chen, C., Zhang, W., Timofejeva, L. and Ma, H. (2006), 'The Arabidopsis thaliana PARTING DANCERS Gene Encoding a Novel Protein Is Required for Normal Meiotic Homologous Recombination', *Molecular Biology of the Cell* **17**, 1331–1343.
- Wijnker, E., Dun, K. V., Snoo, C. B. D., Lelivelt, C. L., Keurentjes, J. J., Naharudin, N. S., Ravi, M., Chan, S. W., Jong, H. D. and Dirks, R. (2012), 'Reverse breeding in Arabidopsis thaliana generates homozygous parental lines from a heterozygous plant', *Nature Genetics* **44**, 467–470.
- Wijnker, E., James, G. V., Ding, J., Becker, F., Klasen, J. R., Rawat, V., Rowan, B. A., de Jong, D. F., de Snoo, C. B., Zapata, L., Huettel, B., de Jong, H., Ossowski, S., Weigel, D., Koornneef, M., Keurentjes, J. J. and Schneeberger, K. (2013), 'The genomic landscape of meiotic crossovers and gene conversions in Arabidopsis thaliana', *eLife* **2**, 1–22.
- Williams, C. G., Goodman, M. M. and Stuber, C. W. (1995), 'Comparative recombination distances among Zea mays L. inbreds, wide crosses and interspecific hybrids', *Genetics* **141**, 1573–1581.
- Winckler, W., Myers, S. R., Richter, D. J., Onofrio, R. C., McDonald, G. J., Bontrop, R. E., McVean, G. A., Gabriel, S. B., Reich, D., Donnelly, P. and Altshuler, D. (2005), 'Comparison of fine-scale recombination rates in humans and chimpanzees', *Science* **308**, 107–111.

- Wu, J., Mizuno, H., Hayashi-Tsugane, M., Ito, Y., Chiden, Y., Fujisawa, M., Katagiri, S., Saji, S., Yoshiki, S., Karasawa, W., Yoshihara, R., Hayashi, A., Kobayashi, H., Ito, K., Hamada, M., Okamoto, M., Ikeno, M., Ichikawa, Y., Katayose, Y., Yano, M., Matsumoto, T. and Sasaki, T. (2003), ‘Physical maps and recombination frequency of six rice chromosomes’, *Plant Journal* **36**, 720–730.
- Wu, Y., Bhat, P. R., Close, T. J. and Lonardi, S. (2008a), ‘Efficient and accurate construction of genetic linkage maps from the minimum spanning tree of a graph’, *PLoS Genetics* **4**.
- Wu, Y., Close, T. J. and Lonardi, S. (2008b), ‘On the accurate construction of consensus genetic maps.’, *Computational systems bioinformatics. Computational Systems Bioinformatics Conference* **7**, 285–96.
- Wyatt, H. D. and West, S. C. (2014), ‘Holliday junction resolvases’, *Cold Spring Harbor Perspectives in Biology* **6**:a023192.
- Xiong, Z., Gaeta, R. T. and Pires, J. C. (2011), ‘Homoeologous shuffling and chromosome compensation maintain genome balance in resynthesized allopolyploid *Brassica napus*’, *Proceedings of the National Academy of Sciences of the United States of America* **108**, 7908–7913.
- Yandeau-Nelson, M. D., Nikolau, B. J. and Schnable, P. S. (2006), ‘Effects of trans-acting genetic modifiers on meiotic recombination across the a1-sh2 interval of maize’, *Genetics* **174**, 101–112.
- Yandeau-Nelson, M. D., Zhou, Q., Yao, H., Xu, X., Nikolau, B. J. and Schnable, P. S. (2005), ‘MuDR transposase increases the frequency of meiotic crossovers in the vicinity of a Mu insertion in the maize a1 gene’, *Genetics* **169**, 917–929.
- Yang, S., Yuan, Y., Wang, L., Li, J., Wang, W., Liu, H., Chen, J. Q., Hurst, L. D. and Tian, D. (2012), ‘Great majority of recombination events in *Arabidopsis* are gene conversion events’, *Proceedings of the National Academy of Sciences of the United States of America* **109**, 20992–20997.
- Yao, H. and Schnable, P. S. (2005), ‘Cis-effects on meiotic recombination across distinct a1-sh2 intervals in a common *zea* genetic background’, *Genetics* **170**, 1929–1944.
- Yao, H., Zhou, Q., Li, J., Smith, H., Yandeau, M., Nikolau, B. J. and Schnable, P. S. (2002), ‘Molecular characterization of meiotic recombination across the 140-kb multigenic a1- sh2 interval of maize’, *Proceedings of the National Academy of Sciences of the United States of America* **99**, 6157–6162.

- Yelina, N. E., Choi, K., Chelysheva, L., Macaulay, M., de Snoo, B., Wijnker, E., Miller, N., Drouaud, J., Grelon, M., Copenhaver, G. P., Mezard, C., Kelly, K. A. and Henderson, I. R. (2012), 'Epigenetic Remodeling of Meiotic Crossover Frequency in *Arabidopsis thaliana* DNA Methyltransferase Mutants', *PLoS Genetics* **8**, e1002844.
- Yelina, N. E., Lambing, C., Hardcastle, T. J., Zhao, X., Santos, B. and Henderson, I. R. (2015), 'DNA methylation epigenetically silences crossover hot spots and controls chromosomal domains of meiotic recombination in *Arabidopsis*', *Genes and Development* **29**, 2183–2202.
- Yelina, N. E., Ziolkowski, P. A., Miller, N., Zhao, X., Kelly, K. A., Muñoz, D. F., Mann, D. J., Copenhaver, G. P. and Henderson, I. R. (2013), 'High-throughput analysis of meiotic crossover frequency and interference via flow cytometry of fluorescent pollen in *Arabidopsis thaliana*', *Nature Protocols* **8**, 2119–2134.
- Youds, J. L. and Boulton, S. J. (2011), 'The choice in meiosis - Defining the factors that influence crossover or non-crossover formation', *Journal of Cell Science* **124**, 501–513.
- Yu, J., Holland, J. B., McMullen, M. D. and Buckler, E. S. (2008), 'Genetic design and statistical power of nested association mapping in maize', *Genetics* **178**, 539–551.
- Yuan, W., Li, X., Chang, Y., Wen, R., Chen, G., Zhang, Q. and Wu, C. (2009), 'Mutation of the rice gene PAIR3 results in lack of bivalent formation in meiosis', *Plant Journal* **59**, 303–315.
- Zhang, C., Song, Y., Cheng, Z. H., Wang, Y. X., Zhu, J., Ma, H., Xu, L. and Yang, Z. N. (2012a), 'The *Arabidopsis thaliana* DSB formation (AtDFO) gene is required for meiotic double-strand break formation', *Plant Journal* **72**, 271–281.
- Zhang, P., Zhang, Y., Sun, L., Sinumporn, S., Yang, Z., Sun, B., Xuan, D., Li, Z., Yu, P., Wu, W., Wang, K., Cao, L. and Cheng, S. (2017), 'The rice AAA-ATPase *osfigl1* is essential for male meiosis', *Frontiers in Plant Science* **8**.
- Zhang, W. W., Pan, J. S., He, H. L., Zhang, C., Li, Z., Zhao, J. L., Yuan, X. J., Zhu, L. H., Huang, S. W. and Cai, R. (2012), 'Construction of a high density integrated genetic map for cucumber (*Cucumis sativus* L.)', *Theoretical and Applied Genetics* **124**, 249–259.

- Zhong, W. and Priest, N. K. (2011), ‘Stress-induced recombination and the mechanism of evolvability’, *Behavioral Ecology and Sociobiology* **65**, 493–502.
- Zickler, D. and Kleckner, N. (1999), ‘Meiotic Chromosomes: Integrating Structure and Function’, *Annual Review of Genetics* **33**, 603–754.
- Zickler, H. (1934), ‘Genetische untersuchungen an einem hetero-thallischen askomyzeten (*Bombardia lunata* nov. spec.)’, *Planta* pp. 573–613.
- Zilberman, D., Gehring, M., Tran, R. K., Ballinger, T. and Henikoff, S. (2007), ‘Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription’, *Nature Genetics* **39**, 61–69.
- Ziolkowski, P. A., Berchowitz, L. E., Lambing, C., Yelina, N. E., Zhao, X., Kelly, K. A., Choi, K., Ziolkowska, L., June, V., Sanchez-Moran, E., Franklin, C., Copenhaver, G. P. and Henderson, I. R. (2015), ‘Juxtaposition of heterozygous and homozygous regions causes reciprocal crossover remodelling via interference during *Arabidopsis* meiosis’, *eLife* **4**.
- Ziolkowski, P. A., Underwood, C. J., Lambing, C., Martinez-Garcia, M., Lawrence, E. J., Ziolkowska, L., Griffin, C., Choi, K., Franklin, F. C. H., Martienssen, R. A. and Henderson, I. R. (2017), ‘Natural variation and dosage of the HEI10 meiotic E3 ligase control *Arabidopsis* crossover recombination’, *Genes and Development* **31**, 306–317.


3 Genomic prediction of the recombination rate variation in barley – A route to highly recombinogenic genotypes

Authors:

Federico Casale, Delphine van Inghelandt, Marius Weisweiler, Jinquan Li, and Benjamin Stich.

Own contribution: First author. I performed the data analyses and wrote the manuscript.

Genomic prediction of the recombination rate variation in barley – A route to highly recombinogenic genotypes

Federico Casale¹ , Delphine Van Inghelandt¹ , Marius Weisweiler¹ , Jinquan Li^{2,4} and Benjamin Stich^{1,2,3,*} 

¹Institute of Quantitative Genetics and Genomics of Plants, Heinrich Heine University, Düsseldorf, Germany

²Max Planck Institute for Plant Breeding Research, Köln, Germany

³Cluster of Excellence on Plant Sciences, From Complex Traits Towards Synthetic Modules, Düsseldorf, Germany

⁴Strube D&S GmbH, Söllingen, Germany

Received 28 June 2021;

revised 6 October 2021;

accepted 7 November 2021.

*Correspondence (Tel +49 211 81 13395;

fax +49 211 81 13395; email

benjamin.stich@hhu.de)

Summary

Meiotic recombination is not only fundamental to the adaptation of sexually reproducing eukaryotes in nature but increased recombination rates facilitate the combination of favourable alleles into a single haplotype in breeding programmes. The main objectives of this study were to (i) assess the extent and distribution of the recombination rate variation in cultivated barley (*Hordeum vulgare* L.), (ii) quantify the importance of the general and specific recombination effects, and (iii) evaluate a genomic selection approach's ability to predict the recombination rate variation. Genetic maps were created for the 45 segregating populations that were derived from crosses among 23 spring barley inbreds with origins across the world. The genome-wide recombination rate among populations ranged from 0.31 to 0.73 cM/Mbp. The crossing design used in this study allowed to separate the general recombination effects (GRE) of individual parental inbreds from the specific recombination effects (SRE) caused by the combinations of parental inbreds. The variance of the genome-wide GRE was found to be about eight times the variance of the SRE. This finding indicated that parental inbreds differ in the efficiency of their recombination machinery. The ability to predict the chromosome or genome-wide recombination rate of an inbred ranged from 0.80 to 0.85. These results suggest that a reliable screening of large genetic materials for their potential to cause a high extent of genetic recombination in their progeny is possible, allowing to systematically manipulate the recombination rate using natural variation.

Keywords: recombination rate, genomic prediction, GBLUP, plant breeding.

Introduction

The reciprocal genetic exchange between homologous chromosomes is termed crossover (CO), and it is required for the proper chromosomal segregation during the first meiotic division (Morgan, 1916). This genetic reshuffling in addition to the independent segregation of chromosomes enables meiosis to produce new allelic combinations in the resulting gametes. This process is called meiotic recombination. Consequently, the rate and distribution pattern of recombination events along the genome determine the effectiveness of selection in removing deleterious mutations and increasing the frequency of beneficial allele combinations (Henderson, 2012). This makes meiotic recombination a fundamental element not only for the adaptation of sexually reproducing eukaryotes in nature but also for stacking many favourable alleles into a single haplotype in breeding schemes (Nachman, 2002; Tiley and Burleigh, 2015). The manipulation of the factors influencing the rate and distribution of recombination events along the genome, therefore, has the potential to accelerate plant and animal breeding (Choi and Henderson, 2015).

Studies on model plants have increased our knowledge about the mechanism and regulation of recombination considerably (for review, see Mercier *et al.*, 2015; Wang and Copenhaver, 2018). While this has opened up possibilities for the manipulation of genetic recombination by environmental factors such as temperature (Arrieta *et al.*, 2020; Higgins *et al.*, 2012), an even higher

impact is expected from approaches that rely on altering the genetics of recombination (Taagen *et al.*, 2020). The use of genome-editing approaches that induce double-stranded breaks (DSBs) or modify epigenomes at the desired sites of recombination (Hayut *et al.*, 2017; Underwood *et al.*, 2018), and the manipulation of CO factors (Mieulet *et al.*, 2018; Sarno *et al.*, 2017; Tam *et al.*, 2011) are increasingly applicable for achieving this goal. However, such approaches still face technical challenges such as to the genotype-specific efficiency of genetic transformation to be effectively applied (Altpeter *et al.*, 2016; Hayta *et al.*, 2019). In addition to technical challenges, a constraint for the adoption of gene-edited crops is government regulation (Taagen *et al.*, 2020). As an alternative to controlled recombination via genome editing, the utilization of natural variation remains a possible way to manipulate recombination in plants.

The meiotic recombination rate is known to vary within and among species (Nachman, 2002). Over the last years, several studies have examined the intraspecific variation of recombination rate in animals (Booker *et al.*, 2017; Chan *et al.*, 2012; Coop *et al.*, 2008; Dumont *et al.*, 2009; Fledel-Alon *et al.*, 2011; Hunter *et al.*, 2016; Kong *et al.*, 2010; Mezner *et al.*, 2010; Petit *et al.*, 2017; Sandor *et al.*, 2012) and plants such as *Arabidopsis* (Kim *et al.*, 2007; Salomé *et al.*, 2012; Ziolkowski *et al.*, 2017), maize (Bauer *et al.*, 2013; McMullen *et al.*, 2009; Rodgers-Melnick *et al.*, 2015), wheat (Darrier *et al.*, 2017; Gardiner *et al.*, 2019; Jordan *et al.*, 2018), rice (Marand *et al.*, 2019), cotton (Shen

et al., 2019), and *Eucalyptus* (Gion et al., 2016), which showed high extent of variation in the frequency and distribution of recombination events across the genomes among genotypes of the same species. Until now, the recombination rate variation in barley (*Hordeum vulgare* L.) was examined in crosses between wild and cultivated barley (Dreissig et al., 2020), which may be due to the structural variants (SVs) between both genomes, thus not fully representative of intraspecific variation. In this sense, information about the intraspecific recombination rates for cultivated barley, which is an important crop species and a model for the Triticeae tribe, is lacking.

The most robust approach to assess the recombination rate variation is to examine the co-segregation pattern of alleles at linked loci in populations with known pedigree relatedness (Petit et al., 2017; Salomé et al., 2012). The pedigrees that were examined in an animal genetic context were designed such that one female was mainly recombined with one male (e.g. Smeds et al., 2016; Weng et al., 2014). Similarly, the nested association mapping designs that were used for evaluating the recombination rate variation among plants consist of progenies derived from the crosses between a diverse set of genotypes and a common parent (Dreissig et al., 2020; Jordan et al., 2018; McMullen et al., 2009). Such pedigrees, however, do not allow to assess whether a high recombination rate in the progenies is due to the general parental effect or from the specific combination of both parental genotypes. This information is crucial for designing experiments to alter the recombination rate systematically.

The idea to exploit the natural recombination rate variation to construct highly recombinogenic genotypes was proposed more than 30 years ago (Cederberg, 1985). Despite the dramatic advances in genotyping technology and the availability of new resources for genetic mapping (Beyer et al., 2008; Lee et al., 2002; Yu et al., 2008), such highly recombinogenic genotypes have not yet been developed in any crop species. This might be explained by the fact that the assessment of recombination properties requires considerable experimental efforts to generate and genotype one to several segregating populations for each accession of interest. Genomic selection approaches employ all available markers across the genomes to predict genotypic values and are nowadays used in most animal and crop breeding programmes because of their high prediction accuracy (Meuwissen et al., 2001). To our knowledge, no previous study has evaluated the potential of genomic selection (GS) to predict the genome-wide and local recombination rate variation. These approaches may permit the development of highly recombinogenic lines in recurrent genomic selection programmes.

The main objectives of this study were to (i) assess the extent and distribution of recombination rate variation in cultivated barley, (ii) quantify the importance of the general and specific recombination effects, and (iii) evaluate a genomic selection approach's ability to predict the recombination rate variation.

Results

Genetic variation and parental segregation among the DRR populations

A principal coordinate analysis (PCoA) was performed for the diversity panel, three *ssp. spontaneum* and one *ssp. agriocrithon* accessions, and Morex (Figure 1b). The first axis separated the two rows from the six-row genotypes, where the four wild

barley accessions clustered with the latter. The result of the PCoA suggested that the parents of the double round-robin (DRR) populations well represent the genotypic space of the diversity panel. In the PCoA of the DRR populations and their parental inbreds, the inbreds of each DRR population clustered together in between the position of their parental inbreds (Figure 1c), thereby illustrating the absence of pedigree errors. The assessment of segregation distortion (SD) demonstrated that for 38 of the 45 DRR populations, one or several genome regions were observed with an allele frequency that significantly ($P < 0.05$) deviated from 0.5. Several of the observed SD regions were found to be large with up to 300 Mbp (Figure S1). Interestingly, some of the DRR populations exhibited shared segregation trends depending on the parental inbred. Allelic segregation favoured the allele of IG128216 in chromosome 5H (50–350 Mbp) of the populations HvDRR30 and HvDRR31. Contrastingly, the allelic segregation on chromosome 6H (390–530 Mbp) of the populations HvDRR43, HvDRR44, HvDRR45, and HvDRR46 disfavoured the allele of Kombyne. Because the reported regions were long, many genes might have been responsible for the segregation bias.

The intraspecific recombination variation in cultivated barley

The high-density linkage maps that were constructed for the 45 DRR populations comprised 6,569–12,962 single-nucleotide polymorphisms (SNPs) (Figure S2). This resulted in genetic maps with average distances between adjacent bins varying from 0.88 to 3.17 cM and the median of the longest gap across all populations being 23.99 cM. The median of Pearson's correlation coefficient between the genetic and physical map position was 0.9 across all populations. This, together with the median fraction of 0.008 of the SNPs that were at a threshold of 5 cM non-collinear to the physical map in each DRR population, indicated high collinearity between the obtained genetic maps and the reference genome.

The previously described genetic maps were the basis for the assessment of the recombination rates (Figures S3 and S4). The recombination rate per chromosome observed on average across all populations ranged from 0.37 (4H) to 0.58 cM/Mbp (5H) (Figure 2a). The same pattern of recombination rate along the chromosomes was also noted for the different populations (Figure 3a,b). Similar to what has been widely observed in species with large genomes such as grasses (Melamed-Bessudo et al., 2016), the recombination rate was consistently found to be almost negligible in the pericentromeric region, while an increase was noticed in the distal regions. As a result, the recombination rate was found to be positively correlated with gene density ($P < 0.001$). The same trend was detected in the analysis of historical recombination in the diversity panel (Figure 3c). This was supported by the observation of a significant ($P < 0.05$) correlation coefficient between the historical and meiotic recombination rate assessed on a consensus map basis that ranged from 0.81 to 0.93. The variability of the genome-wide recombination rate among populations was higher compared with that observed among the chromosomes, and it ranged from 0.31 to 0.73 cM/Mbp (median = 0.45 cM/Mbp). Local differences in the recombination rate were detected among populations with a median of 4.5-fold variation in 10 Mbp windows, although some windows showed up to even a 198-fold variation among populations (Figure 2c). The differences in the recombination rate among

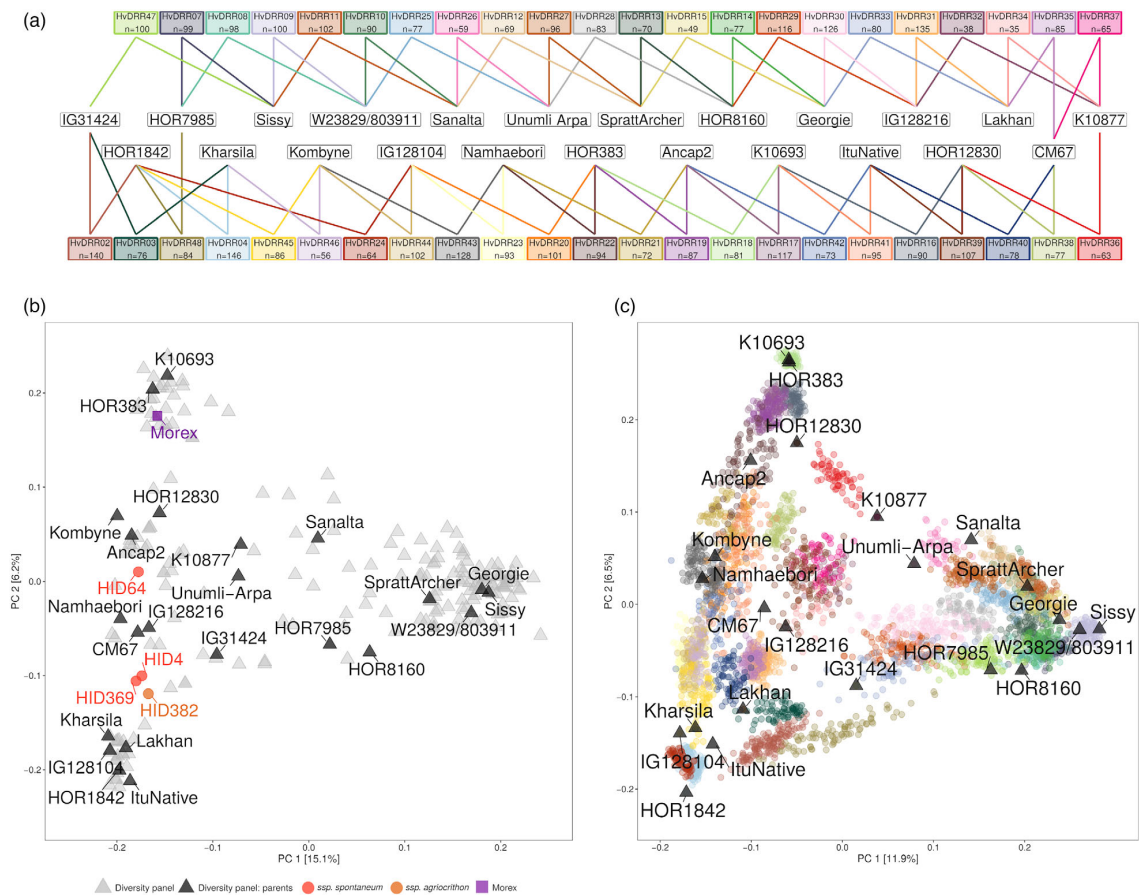


Figure 1 (a) The crossing scheme underlying the double round-robin populations of barley. The number of recombinant inbred lines available per population is indicated below each population's name. (b) Principal coordinate analysis of the diversity panel, Morex and three *ssp. spontaneum* and one *ssp. agriocrithon* accessions based on 36,077 SNP markers. PC 1 and PC 2 are the first and second principal coordinate, respectively, and the number in parentheses refers to the proportion of variance explained by the principal coordinates. (c) Principal coordinate analysis of the double round-robin (DRR) populations and their parental inbreds based on 36,077 SNP markers. PC 1 and PC 2 are the first and second principal coordinate, respectively, and the number in parentheses refers to the proportion of variance explained by the principal coordinates. Parental inbreds are indicated by black triangles. Individuals from the same DRR population are indicated with dots of the same colour.

populations were particularly large in the distal regions, where the recombination rates varied between 1.21 and 6.45 cM/Mbp. However, when correcting for the mean differences in recombination rate across the genome by calculating the coefficient of variation, a higher recombination rate variation among populations was observed in the pericentromeric region compared with the distal regions. In addition, distal and pericentromeric regions were found distinctively correlated with the genome-wide recombination rate ($P < 0.001$) (Figure 4). This trend was also detected when evaluating the general recombination effect (GRE) values (Figure S5).

The extent to which the above-explained differences in recombination rates among DRR populations were due to the GRE of each of the two parental inbreds compared to the specific recombination effect (SRE) of the combination of the two involved inbreds was quantified by comparing their variances. The variance of the phenotypic estimated GRE (GRE_P) values calculated genome-wide ($\sigma_{GRE_P}^2$) was about eight times the variance of the SREs ($\sigma_{SRE_P}^2$) (Table 1). The ratio between $\sigma_{GRE_P}^2$

and $\sigma_{SRE_P}^2$ was observed to be even higher for some of the chromosomes and examined windows. The proportion of recombination rates' variation that was due to genetic differences was measured using broad-sense heritability (H^2) and was 0.37 on a genome-wide level. For individual chromosomes, H^2 had values between 0.30 and 0.37, which is slightly lower compared with that for the genome-wide recombination rate variation. Regardless of the analysed scale level, the inbreds with the largest GRE_P were ItuNative, CM67, Ancap2, Lakhan, and HOR12830 (Figure 2b).

The potential reasons for the considerable differences among the GRE_P values were examined. The GRE_P values of the parental inbreds were found to be significantly ($P < 0.001$) and positively (0.68) correlated with the average temperature of the geographical area where they originated from. The inbreds with the highest recombination rates originated from regions with high mean temperatures, and they were mostly six-row types ($P < 0.05$). In contrast, annual precipitation and germplasm type were not found to be significantly correlated with GRE_P . Additionally, the

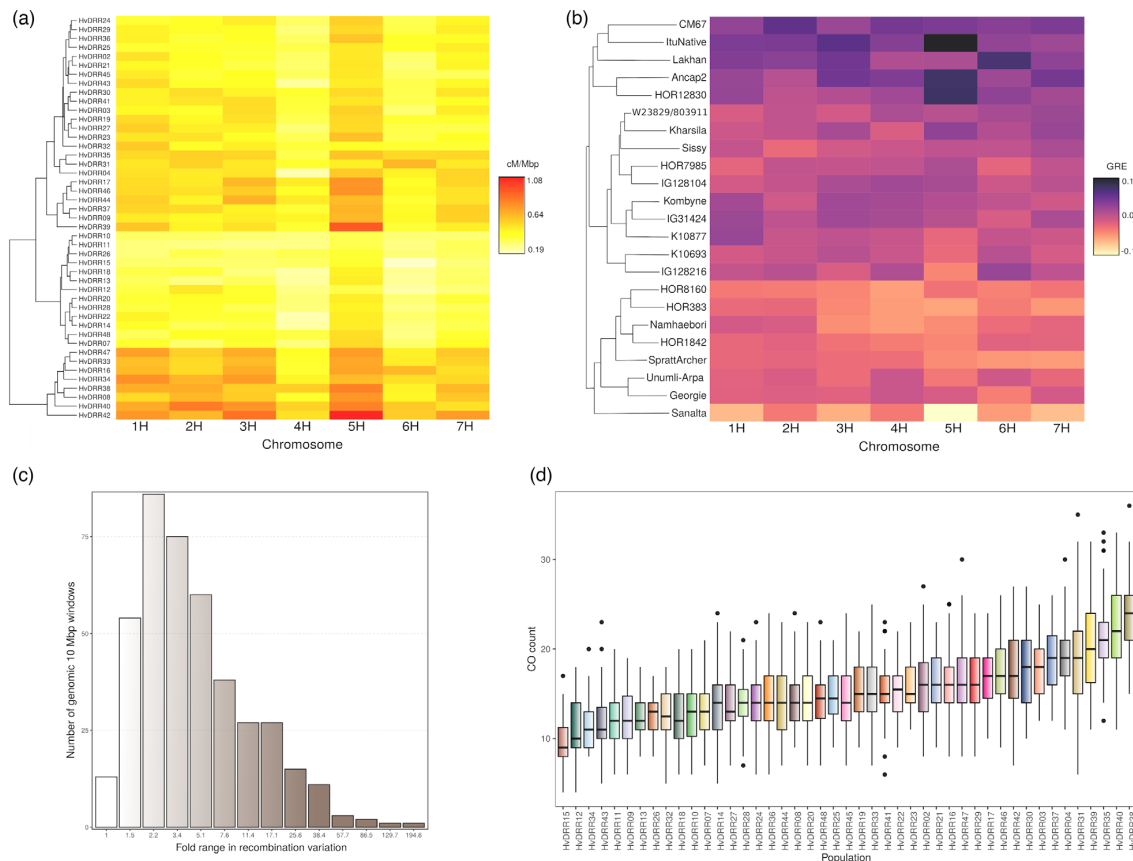


Figure 2 (a) Heat map of the chromosome-wide recombination rates for the 45 double round-robin populations. (b) Heat map of the chromosome-wide phenotypic-estimated general recombination effect (GRE_p) for the 23 parental inbreds. Darker colours indicate higher recombination rates or GRE_p values. On the y-axis, populations and inbreds are ordered according to their hierarchical complete clustering based on Euclidean distances of recombination rates and GRE_p values per chromosome respectively. (c) Histogram of the number of 10 Mbp windows by the fold range of recombination rate variation. (d) Boxplot of the number of counted genome-wide crossovers (CO) for all DRR populations.

sequence divergence between the parental inbreds was evaluated as a factor contributing to the recombination rate variation. No significant correlation was observed between the recombination rate and allelic parental similarity neither on a population basis averaged across the genome ($P = 0.423$) nor on a genome basis averaged across the populations ($P = 0.510$; Figure S6). Across the pericentromeric region, such correlation was not significant ($P = 0.06$) on a population basis, while on a genome basis, a high negative correlation ($P < 0.001$) was observed.

QTL analysis of CO counts

The genome-wide CO counts ranged from 7 to 59 per DRR population with a median of 20 COs (Figure 2d). Pearson's correlation coefficient between the average CO count per population and their recombination rates was with 0.52 highly significant ($P < 0.001$). The 1.44-fold variation found between

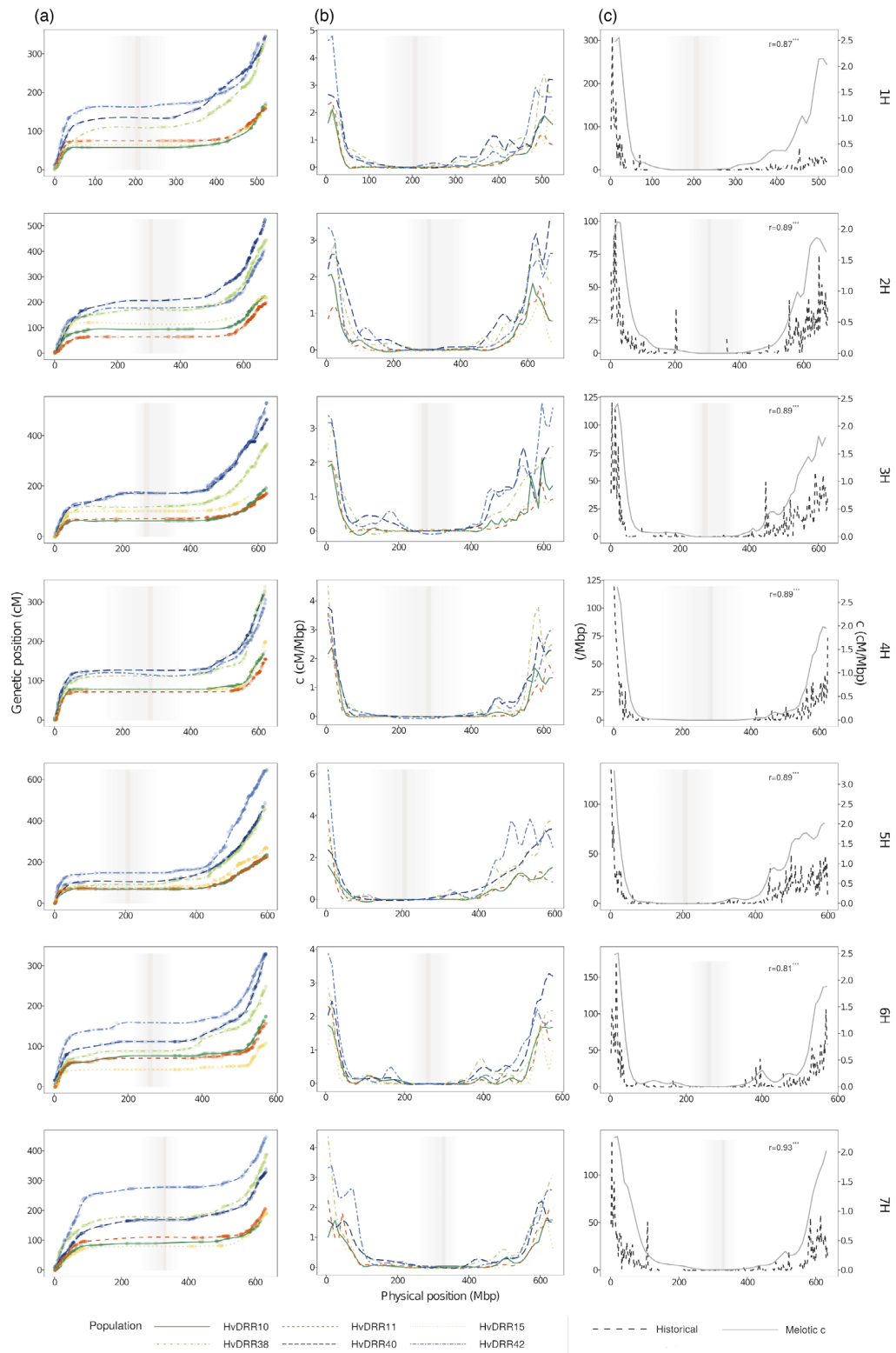
the populations with the lowest and highest CO count is consistent with the respective 1.35-fold variation found for the recombination rate. On a chromosomal level, the CO counts ranged from 0 to 14 with a median between 2 and 3 COs depending on the chromosome. COs per chromosome were noted to be significantly ($P < 0.001$) correlated with the chromosomes' physical length. Across the 8 examined CO counts, 16 quantitative trait loci (QTLs) were detected using a multi-population analysis (Table S1). Although each detected QTL was significant ($P < 0.05$) in at least five populations (Figures S7–S10), it explained <3% of the total phenotypic variance.

The genomic prediction of recombination-related estimates

The ability to predict recombination-related estimates for individual populations using the genome-wide SNP profiles (Figure S11)

Figure 3 (a) Marey map and (b) meiotic recombination rate (c) landscape across the seven chromosomes of the three double round-robin populations with the highest and lowest genomic recombination rates. (c) Historical recombination estimates ρ_w (Mbp, black-dashed line) for the diversity panel and consensus meiotic recombination rate c (cM/Mbp, grey-solid line) across all DRR populations along the seven barley chromosomes. The vertical line indicates the position of the centromere in the reference map and the shadow denotes the pericentromeric region. r is Spearman's correlation coefficient between the historical recombination estimate $\hat{\rho}_w$ and the consensus (c) across 10 Mbp windows.

680 Federico Casale et al.



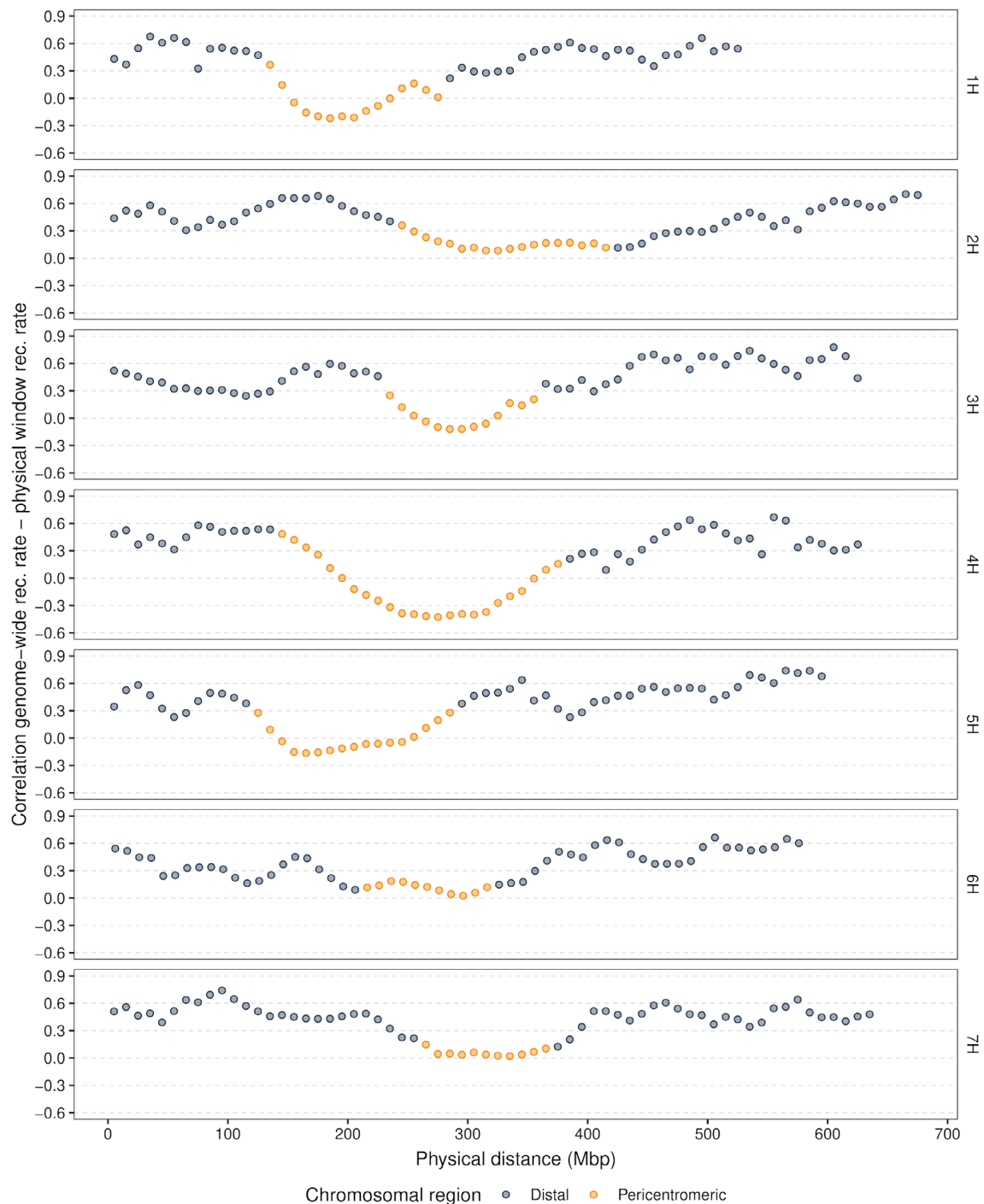


Figure 4 Pearson's correlation coefficient between the 45 double round-robin populations' recombination rate values for 10 Mbp physical windows and their respective genome-wide recombination rate values across the seven barley chromosomes.

of the parental inbreds was assessed. When using all DRR populations as the training set, the genomic best linear unbiased prediction (GBLUP) model resulted in high prediction abilities concerning the recombination rate with values between 0.74 and

0.94 for the entire genome and the individual chromosomes (Figure 5a). Cross-validation (CV) approaches were utilized to obtain unbiased prediction abilities. The fivefold CV approach led to prediction abilities with values between 0.40 and 0.53, about

Table 1 The variances for the phenotypic estimated general ($\sigma_{GRE_P}^2$) and specific ($\sigma_{SRE_P}^2$) recombination effects as well as the phenotypic estimated broad-sense heritability (H^2) for recombination rate variation.

Molecular level	$\sigma_{GRE_P}^2$	$\sigma_{SRE_P}^2$	H^2
Genome-wide	0.00229	0.00028	0.37
Chromosomes			
1H	0.00222	0.00104	0.32
2H	0.00174	0.00107	0.30
3H	0.00302	0.00116	0.32
4H	0.00193	0.00014	0.38
5H	0.00544	0.00048	0.37
6H	0.00238	0.00040	0.36
7H	0.00223	0.00031	0.36
Windows			
Min	0.0000069	0	0.21
Max	0.17	0.14	0.63

40% lower than that observed without CV. A further reduction in the training set (TS) size from 36 to 27 or even 18 lessened the ability to predict genome-wide recombination rate by only 0.03 and 0.11, respectively, but it was slightly more pronounced for the recombination rate of the individual chromosomes. A second CV approach was implemented to test the model's ability to predict the recombination rate for populations that were unrelated by pedigree with the populations of the TS. The prediction ability observed in these two scenarios was about 0.1 lower than those in the first CV approach with a comparable TS size.

In the same way, the ability to predict the GRE of parental inbreds was evaluated. The same trends outlined previously were also valid for the predicted GRE values (Figure 5b). However, the general level of the prediction ability was 13% higher compared with that concerning the recombination rate of the populations. The aforementioned analyses were also performed for the recombination rate in 10 Mbp physical windows. The model's ability to predict recombination across individual windows was more variable and consistently lower than for entire chromosomes or genome-wide predictions (Figure S12). In addition, the model was utilized to predict the genomic estimated GRE (GRE_G) of the 3,959 DRR recombinant inbred lines (RILs). The predicted GRE_G values of the RILs and the resulting genomic estimated breeding value (GEBV) of all 7,838,820 possible RILs' hybrid combinations were observed to represent the recombination variation in parental inbreds and the DRR populations respectively (Figures 6a,b).

The ability of the model to predict the recombination of individual populations decreased by 0.02, 0.03, and 0.05 on average across the scenarios when randomly sampling a uniform distribution of SNPs across the physical map with 1 SNP per 1, 5, and 10 Mbp respectively (Figure S13). The reduction to predict the GRE of parental inbreds with the three reduced marker densities was of similar size with 0.03, 0.04, and 0.05.

Discussion

Recent advances in understanding the mechanisms and regulation of recombination opened up biotechnological possibilities to manipulate genetic recombination (Mercier *et al.*, 2015). The use of genome-editing approaches that induce DSBs or modify

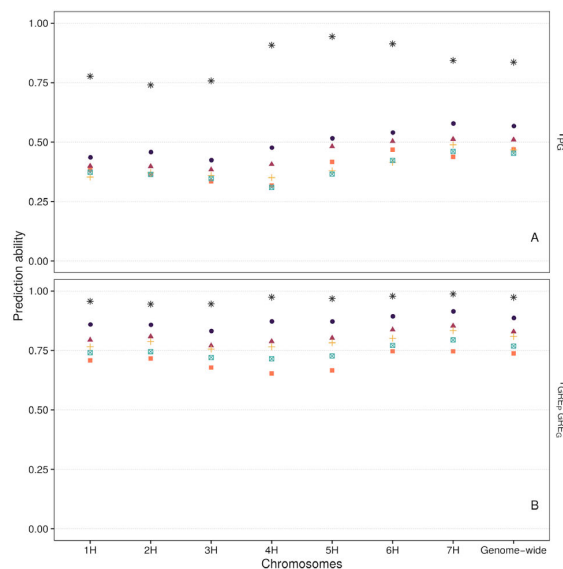


Figure 5 Genomic prediction ability concerning the recombination rate variation of individual chromosomes and the genome-wide level, using different approaches. (a) Pearson's correlation coefficient between the observed recombination rate and genomic estimated breeding values of the double round-robin (DRR) populations, r_{PG} . (b) Pearson's correlation coefficient between the phenotypic and genomic estimated general recombination effects of the parental inbreds, $r_{GRE_P GRE_G}$. Cross-validation (CV) scenarios for genomic prediction are detailed in the table, where the number of populations per validation set (N_{pop} VS), the number of parents involved in the populations in each validation set (N_{par} VS), the number of populations per training set (N_{pop} TS), and the number of parents involved in the populations in each training set (N_{par} TS) are given.

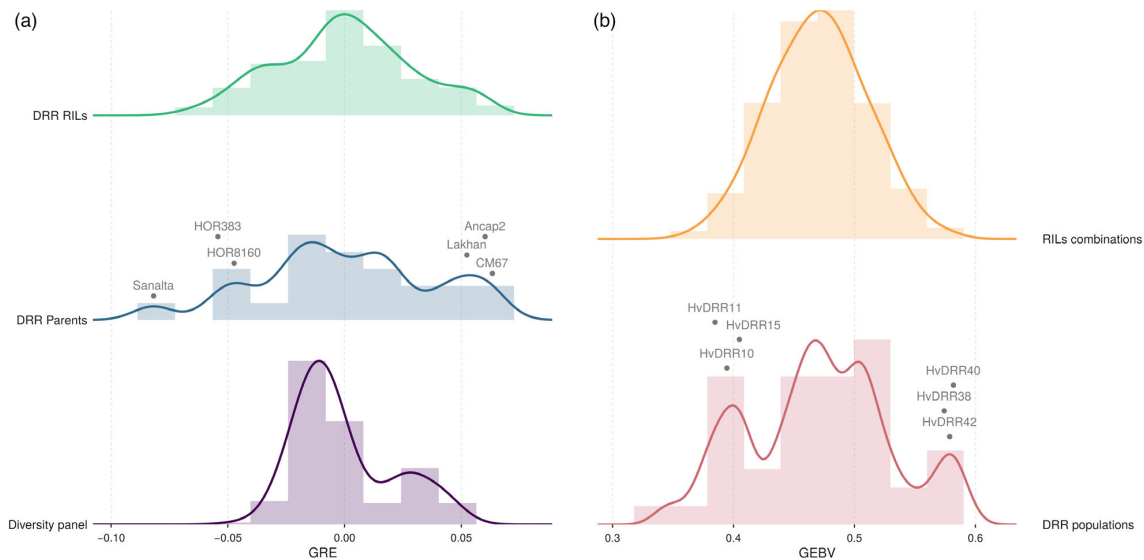


Figure 6 (a) Frequency distribution of the genome-wide general recombination effects (GRE_G) of the diversity panel, double round-robin (DRR) populations' parental lines, and DRR recombinant inbred lines (RILs) predicted by GBLUP. The GRE_G for the parental inbreds of the three DRR populations with the respective lowest and highest recombination rate is displayed. (b) Frequency distribution of the genomic estimated breeding values (GEBVs) concerning the genome-wide recombination rate for the DRR populations and for all possible populations derived from the DRR RILs. The GEBVs for the three DRR populations with the lowest and highest genome-wide recombination rate are displayed.

epigenomes at desired sites (Hayut *et al.*, 2017; Underwood *et al.*, 2018) and the manipulation of CO factors (Mieulet *et al.*, 2018; Sarno *et al.*, 2017; Tam *et al.*, 2011) are increasingly applicable for achieving this goal. However, the utilization of natural variation explored in the present study remains an alternative way to the mentioned approaches to manipulate recombination.

The intraspecific variation of recombination rate in cultivated barley

At the genome-wide level, the observed meiotic recombination rate ranged from 0.31 to 0.73 cM/Mbp (median = 0.45 cM/Mbp) among DRR populations (Figure 2a). Populations with the highest and lowest genome-wide recombination rates were also among the most frequent extreme recombining populations when ranking them according to the recombination rate per individual 10 Mbp window (data not shown). This indicates that the observed genome-wide recombination rates are not due to the effect of a few windows with a particularly high or low recombination rate in a population but because of the genome-wide tendency.

The observed level of recombination rate variation in the DRR populations was higher than that reported in populations derived from crosses between domesticated and wild barley accessions (Dreissig *et al.*, 2020). Moreover, it was also higher than most observations for other plant species such as *Arabidopsis* (Salomé *et al.*, 2012; Ziolkowski *et al.*, 2017), maize (Bauer *et al.*, 2013), and wheat (Gardiner *et al.*, 2019; Jordan *et al.*, 2018), but it was lower than the values observed in the animal kingdom (Booker *et al.*, 2017; Chan *et al.*, 2012; Fledel-Alon *et al.*, 2011; Mezner *et al.*, 2010). The high recombination rate variation observed in this study is presumably due to the high extent of genetic variation among the parental inbreds, which represent most of

the genetic variation of cultivated barley (Figure 1c). A previously reported nested association mapping (NAM) population that comprised a vast genetic diversity of cultivated maize was also found to show a high variation in recombination rates among populations (McMullen *et al.*, 2009).

The considerable differences noticed in the genome- and chromosome-wide recombination rates among the DRR populations led to the question of whether they are caused by the effect of parental inbreds (GRE) or the specific combination of two parental inbreds (SRE); being this work the first to report them. The general effect of parental inbreds on recombination was found to be about eight times higher than the specific effect of parental combinations across the different analysed scale levels. This finding suggested that the examined parental inbreds differ in the efficiency of their recombination machinery. The segregation of structural variants in the individual DRR populations, which has been described to influence the extent and distribution of recombination events (Muñoz-Amatrián *et al.*, 2013; Rowan *et al.*, 2019; Saxena *et al.*, 2014), was determined in this study as part of the SRE. This is because the same SVs will not segregate in all populations in which a common parent is involved.

Because of the high importance of the GRE in relation to the SRE in determining the recombination rate variation, it was interesting to understand the causes of the variation in GRE observed among the 23 parental inbreds. The environmental conditions at the stage of meiosis have the potential to influence the recombination rate (Wang and Copenhaver, 2018). In particular, the effect of temperature on meiosis has been studied for a long time in *Drosophila* (Plough, 1917) and plants (Dowrick, 1957), revealing that the rate of formation and distribution of COs varies depending on the temperature, as recently demonstrated in *Arabidopsis* (Choi *et al.*, 2013; Lloyd *et al.*, 2018; Modliszewski *et al.*, 2018), barley (Higgins *et al.*, 2012; Phillips

et al., 2015), and other plants (Wang and Copenhaver, 2018), with no common trends across species. In this sense, the observed GRE and SRE might be a result of the recombination plasticity of the studied inbreds interacting with the particular environmental conditions where our crossing experiment took place. These conditions might have uncovered the different temperature responses of the examined inbreds, making those detectable as recombination rate variation. However, that is unlikely because the controlled environmental conditions under which the experiment took place are standard for barley cultivation.

Nevertheless, the GRE_p values of the parental inbreds were found to be significant ($P < 0.001$) and positively correlated with the average temperature of the geographic areas where they originated from. This observation was in agreement with a previous report (Dreissig et al., 2019). Other environmental factors were not significantly ($P < 0.05$) correlated with the observed GRE_p values.

In addition to environmental factors, the importance of genetic factors in determining the recombination rate variation was explored. As proposed in previous studies, a QTL analysis using the number of CO of each RIL as the dependent variable was performed (Esch et al., 2007; McMullen et al., 2009). To identify shared controllers of recombination across the genetic diversity of cultivated barley, a multi-population analysis was performed, and 16 QTLs were detected. Eight of the detected QTLs were located on different chromosomes than where the CO count used as phenotype was assessed. In addition, three loci were found to be responsible for genome-wide effects. Both observations are in accordance with reports on other plant (Dreissig et al., 2020; Esch et al., 2007; Jordan et al., 2018; Yandeu-Nelson et al., 2006; Ziolkowski et al., 2017) and animal species (Fleidel-Alon et al., 2011; Kong et al., 2010), suggesting the existence of trans-acting control of recombination. This result supports the previous explanation that the high importance of the GRE relative to the SRE in determining the recombination rate variation suggests that the barley inbreds used in this study differ in the efficiency of their recombination machinery. However, each of the detected QTLs explained <3% of the phenotypic variance (Table S1). This result is in agreement with previous studies on the genetics of intraspecific recombination in maize, wherein no shared controllers of recombination have been detected (McMullen et al., 2009). Furthermore, it suggests that meiotic recombination has a polygenic inheritance. The same conclusion can be drawn from the observation of a rather uniform distribution of marker effects across the genome when fitting genomic prediction models to the recombination rate estimates (Figure S11).

Finally, it must be considered that, in addition to allelic variation at the single nucleotide level, epigenetics factors are also known to play a role in the recombination of the plant genome (Henderson, 2012). In particular, recombination rate is negatively associated with the level of DNA methylation and nucleosome density (Apuli et al., 2020; Choi et al., 2013; Rowan et al., 2019), which partly explains the suppression for recombination in the centromeric region of chromosomes in plants where the content of heterochromatin is high (Choi et al., 2018). However, further research is needed to understand the effect of methylation on the recombination rate variation among different genotypes.

Breeding for recombination rate

The amount of genetic variation generated per meiosis is determined by the extent of the recombination rate (Henderson, 2012). Therefore, recombination influences the population size as

well as the number of generations required to stack multiple favourable alleles in any breeding programme (Choi and Henderson, 2015). Developing genotypes that carry alleles providing high recombination will, thus, increase the gain of selection (Jordan et al., 2018). The present study provides an evaluation of the procedure required to perform selection for increased recombination rates.

The considerable differences observed among the recombination rate of individual DRR populations, and especially among the GRE of parental inbreds (Figures 2a,b) as well as the moderate heritabilities (Table 1), indicate the high potential for using natural variation to manipulate the rate and distribution of recombination in the genome by systematic breeding. This approach requires the evaluation of the genetic material for its recombination properties. The high relative importance of GRE compared with SRE in determining the recombination rate variation observed in this study suggests that the recombination properties of genetic materials do not need to be evaluated in crosses with several other parental genotypes but can instead be evaluated in a resource-efficient manner by crossing them with only one other parent. However, such approaches still require the generation of segregating material from each of the genotypes of interest, which, even with today's genotyping and sequencing techniques, is a task that demands considerable resources. As an alternative, the prediction of recombination-related estimates based on genome-wide SNP profiles was evaluated for the first time in the present study. The GBLUP model using genome-wide SNP profile data has been shown to provide a high ability to predict recombination-related estimates (e.g. GRE) as well as the observed meiotic recombination (Figure 5). The observed prediction abilities were on a similar level to those concerning agronomic and quality traits (Haile et al., 2018; Heffner et al., 2011), where genomic prediction is routinely used in many commercial breeding programmes.

To evaluate if the observed predictive abilities were determined by the fact that the SNPs used to calculate the genetic relationship matrix were mainly located in genome regions with a high recombination rate (Figure S11), resampling simulations have been performed. Only a mild reduction in the model's predictive ability was observed when using uniformly distributed SNPs across the genome compared to the original set of SNPs (Figure S13). That illustrates that the observed high predictive abilities are not due to the overrepresentation of those regions of the genome with high recombination rates when estimating the genetic relationship of the parental inbreds.

The impact of TS size on the prediction ability of the model was assessed through cross-validation. As expected, the prediction ability of the model significantly decreased with smaller TS sizes in both the CV approaches performed in this study (Figure 5), which is consistent with what was previously observed in genomic selection studies in animals (VanRaden et al., 2009) and plants (e.g. Heffner et al., 2011; Lorenzana and Bernardo, 2009; Technow et al., 2013). However, the high prediction ability obtained with TS sizes of 27 or 18 segregating populations suggests that it is still possible to accurately select genotypes for their recombination properties using datasets that are considerably smaller than the one used in our study.

An aspect that is very important for the design of breeding programmes was examined: the ability of the model to predict the recombination rate of segregating populations derived from new inbreds, i.e., inbreds for which no other segregating populations are available in the TS. For these scenarios, only

about 10% lower prediction abilities were observed compared with the CV scenarios wherein related segregating populations were in the TS of similar size (Figure 5). This finding indicates that even for new inbreds, recombination properties can be predicted reasonably well. To illustrate it further, the genome-wide GRE_G of different sets of new inbreds was predicted (Figure 6a), in addition to the GEBVs of populations derived from such new inbreds (Figure 6b). The high variation among predicted GEBVs demonstrates that the proposed method makes it possible to screen and select highly recombinogenic genotypes based on their SNPs' profiles. These are then evaluated and recombined in the next step of recurrent selection schemes for altered recombination properties. In such a breeding scheme, the genotyped individuals are used not only to quantify the recombination rate of the parental genotypes but also to start the next cycle of a breeding programme. This, however, would not be possible when using high-throughput pollen sequencing (Dreissig et al., 2015; Drouaud et al., 2013) instead of genotyping individual plants.

When considering the standard deviation in recombination rates among populations across the genome, it might be concluded that no variation among populations exists in the pericentromeric region (Figure 3b). However, when adjusting the variation of recombination rate among populations using the mean recombination rate in a window, to consider the coefficient of variation, a higher recombination rate variation among populations in the pericentromeric region than in the distal regions was detected. This indicated that also the recombination rate variation near the centromere should be considered when selecting genotypes for recombination.

The potential impact of increasing genetic variation in the pericentromeric region on barley breeding is particularly high, as a big portion of barley's functional genes is present in this region (Mascher et al., 2017; Phillips et al., 2015). Therefore, the ability to predict recombination properties was evaluated not only on a genome- or chromosome-wide scale but also in smaller windows across the genome. The prediction abilities observed for the recombination rate in 10 Mbp windows were considerably lower compared with that on a chromosome- or genome-wide level (Figure S12). However, the ability of the model to predict the GRE for individual windows was on average >0.5 after cross-validation, which suggests that an alteration of recombination properties in individual windows is possible.

Nevertheless, very low prediction abilities were noticed in some windows, presumably because of the low extent of variation among the GRE_P of the parental inbreds for these windows. The pairwise correlations among the GRE_P of individual windows and the genome-wide GRE_P were calculated to investigate if the inbreds with high genome-wide GRE_P were also those that cause a high recombination rate in the pericentromeric regions of their progenies. In this analysis, strong positive correlations were detected in the distal regions of all chromosomes, whereas most correlations in the pericentromeric region were considerably lower (Figure S5). This suggests that the mechanisms influencing recombination rate variation in distal regions differ from those in the pericentromeric region of barley chromosomes, as demonstrated in *Arabidopsis* (Choi et al., 2018; Rowan et al., 2019) and rice (Marand et al., 2019).

One aspect that could prevent breeders from employing our proposed procedure is if the increased recombination rate variation is negatively correlated with other important agronomic characters. However, such correlations were not observed for the barley inbreds used in this study (Table S2).

Conclusion

The present study revealed a considerable recombination rate variation among segregating populations of the model species barley. In addition, this variation was observed to be mainly due to the general effects of individual parental inbreds, and only to about 12% of the variance was caused by combinations of both parents. Furthermore, we suggested a route and characterized the required methods to systematically manipulate recombination rates by using natural variation that might serve as an alternative or complement to controlled recombination via transgenesis.

Experimental procedures

Plant materials and genotypic characterization

In this work, two different genetic materials were analysed. On the one hand, 23 inbreds were selected from a diversity panel as parental inbreds of the DRR population. The inbreds were selected based on the maximal combined genotypic and phenotypic richness index (Weisweiler et al., 2019). The parental inbreds were then crossed following the DRR design (Stich, 2009) to initiate biparental populations (Figure 1a). Within each of the 45 populations, randomly chosen genotypes in the F2 generation underwent subsequent selfing generations to produce 35–146 RILs per population. This resulted in a total set of 3,959 RILs across 45 biparental populations, hereafter referred to as DRR populations. The cultivation of the parental inbreds to make the crosses for the F1 generation and the subsequent selfing generations, until S4 of each RIL, were carried out under identical environmental conditions in a greenhouse.

On the other hand, the diversity panel of 224 unrelated inbreds from which the parental inbreds were selected was analysed. It mostly consisted of landraces and improved varieties, representing the worldwide genetic diversity of cultivated spring barley (Haseneyer et al., 2009). In addition, the inbred Morex, three *ssp. spontaneum* (HID 4, HID 64, and HID 369), and one *ssp. agriocrithon* (HID 382) accessions were added to the analysis.

The RILs of the DRR populations were genotyped at the S4 generation as the occurrence of recombination events after this generation can hardly be detected because of the high degree of homozygosity. Both the RILs of the DRR populations and the inbreds from the diversity panel were genotyped by TraitGenetics GmbH (Gatersleben, Germany) using the 50K Illumina Infinium iSelect SNP genotyping array that includes 40,040 SNP markers (Bayer et al., 2017).

Statistical analysis

Characterization of the recombination rate in segregating populations

Estimation of the meiotic recombination rate. The details of the data cleaning process and the procedure employed to construct the DRR populations' linkage maps are provided in the supporting information section (Methods S1 and S2). The positions of the SNPs on the genetic maps were used together with their positions on the reference physical map (Monat et al., 2019) to establish a Marey map (Chakravarti, 1991) for each chromosome–DRR population combination. SNPs that did not generate a monotonously increasing trend were removed from the map but those with 2 cM diversions were tolerated (Bauer et al., 2013). Afterwards, a cubic spline model was fit to the coordinates of the Marey map for each chromosome–DRR

population combination. Model parameters were subsequently adjusted to smooth the curve when needed (Berloff, 2002; Perperoglou et al., 2019). The meiotic recombination rate (c , cM/Mbp) (Falconer and Mackay, 1996; Petit et al., 2017) was calculated as the slope of the fitted curve on a 10 Mbp window basis. In case a Marey map had no SNPs in any of the extreme windows of the respective chromosome, the recombination rate for that window was estimated by deriving the predicted curve value of the average position of the five nearest SNPs.

The average recombination rate for each chromosome was calculated as the average of the recombination rates in the 10 Mbp windows of that particular chromosome. The genome-wide recombination rate for a given population was calculated as the average recombination rate across the seven chromosomes.

The pairwise genetic similarity between both parental inbreds of each population was calculated as the fraction of shared SNP alleles on a 10 Mbp window basis. The parental similarity was correlated with the recombination rate on a genome basis (i.e. the population-based recombination rates were averaged per physical window) as well as a population basis (i.e., the window-based recombination rates were averaged per population). The recombination rate per physical window was correlated with the gene density in each respective window using the gene annotation provided by Monat et al. (2019).

Consensus map. A consensus map based on the 45 linkage maps of the DRR populations was developed based on the following approach: First, the average recombination rate per window across all 45 DRR populations was calculated. Then, the physical distances in Mbp between the adjacent markers (Monat et al., 2019) in each window were converted into cM according to the average recombination rate for that particular window. In our study, the pericentromeric region of each chromosome was defined as the continuous region surrounding the centromere for which the average recombination rate across the 45 DRR populations was 20-fold lower than the average across the barley genome (cf. Baker et al., 2014). Since the pericentromeric region mostly represent the proximal region in this species, the regions of the chromosome which do not belong to the pericentromeric region were designated distal regions.

Calculation of historical recombination and comparison with the meiotic recombination

Historical recombination rates (ρ) were estimated using PHASE version 2.1 (Li and Stephens, 2003) for the diversity panel of 224 inbreds. To allow that the effective population size N_e can vary along the genome, the estimation of the historical recombination rate ($\rho = 4N_e c$) was performed in 2.5 Mbp windows, with an overlap of 200 Kb with neighbouring windows to avoid border effect. PHASE was evaluated with different priors of the mean historical recombination parameters ($\mu = 0.000002$, 0.00001 , 0.00002 , and 0.001). Because of the observed high correlation coefficient among ρ , PHASE was finally used with the default parameter settings. The number of main iterations was increased to obtain 1,000 posterior samples ($\times 10$), as recommended by the authors for more accurate recombination estimates. In addition, all individuals were used in the estimations to obtain the posterior distribution for the historical recombination rate for each window ($\hat{\rho}_w$). ρ_w was set to NA for windows with < 2 variants for the diversity panel. The median of the 1,000 posterior samples of ρ_w was used as the point estimate ($\hat{\rho}_w$).

To compare the patterns of meiotic (c) and historical recombination (ρ) across the barley genome, Spearman's correlation coefficient between the average meiotic recombination rate (c) across the 45 DRR populations and $\hat{\rho}_w$ of the diversity panel was assessed across the 10 Mbp physical windows.

QTL analysis of crossover counts

The number of COs for each chromosome of each RIL as well as the sum of genome-wide COs was the basis for this analysis. To ensure genotypic data's quality, SNPs with a GenTrain score lower than 0.7 were excluded. In each DRR population, SNPs with missing data $> 10\%$ were also discarded. In addition, RILs with $> 10\%$ residual heterozygosity or missing data were discarded from each population. The CO count of each RIL was estimated using the function 'countXO' of the 'R/qtl' package (Broman et al., 2003). Any RIL with a CO count that exceeded by 2 COs, the last consecutive bin of its population's frequency distribution was considered as an outlier and excluded from the analysis. A multi-population QTL analysis was conducted using the R package 'mppR' (Garin et al., 2015), and a cross-specific QTL effect model was considered. The significance threshold above which a position can be selected as QTL was determined as the 0.95 quantile of the null distribution created by performing 1,000 permutations.

Genomic prediction of the recombination rate

Genetic estimates of the recombination rate. The average recombination effect of a parental inbred in a series of different populations was defined as the general recombination effect (GRE), and the recombination effect of a particular population adjusted for the GRE of both involved parental inbreds was defined as the specific recombination effect (SRE). In this sense, the recombination rate c_{ij} in a population created by crossing the parental inbreds i and j was modelled as:

$$c_{ij} = \mu + GRE_i + GRE_j + SRE_{ij}, \quad (1)$$

where μ is the intercept, GRE_i and GRE_j are the GRE effects of the i^{th} and j^{th} parental inbred, respectively, and SRE_{ij} is the SRE effect of the population between parental inbred i and j .

In the present study, two ways to estimate the GRE were evaluated: the genomically estimated GRE (GRE_G) and phenotypic estimated GRE (GRE_P) whose estimation procedure is described below. When the text refers to GRE without specifying whether it is genomic or phenotypic estimated, it refers to both. The estimation of SRE as well as the nomenclature of SRE_G and SRE_P were in analogy to that of GRE.

BLUP model. The meiotic recombination rate (c) was modelled using best linear unbiased prediction (BLUP)

$$\mathbf{c} = \mathbf{1}_n \mu + \mathbf{Z}_{GRE} \mathbf{u}_{GRE} + \mathbf{Z}_{SRE} \mathbf{u}_{SRE} + \mathbf{e}, \quad (2)$$

where \mathbf{c} is the vector of the recombination rates for the 45 DRR populations; $\mathbf{1}_n$ is the unit vector of length n , where n is the number of DRR populations; μ is the general mean; \mathbf{u}_{GRE} and \mathbf{u}_{SRE} are the vectors of GRE and SRE effects; and \mathbf{e} is the vector of random residuals. \mathbf{Z}_{GRE} and \mathbf{Z}_{SRE} are the incidence matrices of the GRE and SRE effects, relating \mathbf{c} to the additive (**A**) and dominance (**D**) genomic relationship matrices respectively. It is assumed that $\mathbf{u}_{GRE} \sim N(0, \mathbf{A}\sigma_a^2)$, $\mathbf{u}_{SRE} \sim N(0, \mathbf{D}\sigma_d^2)$, and $\mathbf{e} \sim N(0, \mathbf{I}\sigma_e^2)$, where σ_a^2 is the additive genetic variance, σ_d^2 is the dominance variance, and σ_e^2 is the residual variance. For the calculation of

GRE_P and SRE_P , **A** and **D** were identity matrices, as was **I**. The model fit and variance component estimation based on REML were performed using the 'sommer' package (Covarrubias-Pazaran, 2016).

GBLUP model. For the calculation of GRE_G and SRE_G , **A** and **D** were matrices from genome-wide SNP markers, thus turning the BLUP into a GBLUP model. The SNP effect's profiles were calculated using ridge regression best linear unbiased prediction (RR-BLUP) (Meuwissen et al., 2001).

Prediction ability of the GBLUP model. The aforementioned model was tested to predict the GRE_P of parental inbreds and the recombination rate of a population at three different scales: genome-wide, individual chromosomes, and 10 Mbp physical windows. The ability of the model to predict the GRE_P of parental inbreds was calculated using Pearson's correlation coefficient between the phenotypic and the genomic estimated GRE of the parental inbreds (r_{GRE_P, GRE_G}).

Moreover, the ability of the model to predict the recombination rate of a population was calculated using Pearson's correlation coefficient between the DRR population's recombination rate (c) and the GEBV of the DRR population's recombination rate (r_{PG}). The latter was calculated using the model [1]. Differences between the correlations were tested for their significance using the approach proposed by Zou (2007). The broad-sense heritability (H^2) for recombination rate was calculated as: $\frac{\sigma_{GRE_P}^2 + \sigma_{SRE_P}^2}{\sigma_{GRE_P}^2 + \sigma_{SRE_P}^2 + \sigma_e^2}$.

Additionally, the GRE_G of each RIL of each DRR population was predicted using the model [2]. The GEBV of all possible combinations among DRRs' RILs was calculated using the model [1].

Prediction ability evaluated by cross-validation. Two different CV procedures were employed, with one comprising three scenarios and the other having two. The first approach was intended to evaluate the ability to predict the recombination rate of new segregating populations from parental inbreds for which already segregating populations are available. In this sense, a fivefold cross-validation was performed by randomly dividing the full set of DRR populations into five disjoint subsets of equal size. For each of the five possible runs, one subset was left out to be used as the validation set (VS), whereas the other four subsets were used as the training set (TS). This procedure was repeated 100 times resulting in 500 cross-validation runs in total. In addition, scenarios with different TS sizes (N_{pop}) were evaluated by reducing the number of subsets in the TS from four ($N_{pop} = 36$) to three (27) and two (18).

The second CV approach focussed on the ability of the model to predict the GRE of new inbreds for which no segregating populations are available yet. In this approach, all populations derived from three randomly selected inbreds ($N_{par} = 3$) were used as VS and all other populations as TS. This procedure was performed 1,000 times. This analysis was also performed for $N_{par} = 5$.

The median of Pearson's correlation coefficients across all runs of each scenario was reported and compared using a t-test. The aforementioned CV approaches were performed at three different scale levels: genome-wide, individual chromosomes, and 10 Mbp windows across the genome.

The impact of the number and distribution of SNPs on the prediction ability of the model was evaluated by repeating the above-described procedure but using different subsets of

uniformly spaced SNPs in every CV run. Three different distributions were tested: 1 SNP per 1, 5, and 10 Mbp.

Correlation between the recombination estimates and the characteristics of inbreds

Pearson's correlation coefficient between the GRE_P values and the variables characterizing the inbreds (such as row type, germplasm type, and geographic origin) was calculated in addition to the environment of their locations of origin (such as annual precipitation and temperature). Information about annual precipitation and the average temperature was estimated for the region of origin of each parental inbred based on historical data (1901–2016) from the World Bank's database (The World Bank, 2020). Furthermore, Pearson's correlation coefficient between the GRE_P values and phenotypic trait scores of the parental inbreds was calculated. Procedures employed to assess phenotypic traits in field experiments are described in the supporting information section (Method S3).

Acknowledgements

Computational infrastructure and support were provided by the Centre for Information and Media Technology at Heinrich Heine University Düsseldorf. The authors give thanks to the IPK for providing the seeds of the diversity panel. We thank our former colleagues Andrea Lossow, Nicole Kliche-Kamphaus, Nele Kaul, Isabelle Scheibert, Marianne Haperscheid and George Alskief, as well as our present colleague, Florian Esser, for their technical assistance in creating and maintaining the DRR populations. In addition, we thank Franziska Wessel from Saatzucht Breun and her team for performing field experiments in Quedlinburg. Finally, we thank Raphael Mercier for his valuable suggestions to improve this manuscript. This research is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy (EXC 2048/1, Project ID: 390686111).

Conflict of interest

The authors declare no conflict of interest.

Author contributions

FC performed all analyses related to meiotic recombination, DVI performed the analyses related to historical recombination and the consensus map, and contributed phenotypic information, MW contributed to SNP data analysis, JL created the segregating populations, BS designed and coordinated the project. FC, DVI, and BS wrote the manuscript. All authors read and approved the final manuscript.

Data availability statement

The genotypic data utilized in this study, as well as the genetic maps and crossover counts, are available at <https://doi.org/10.5281/zenodo.5495951>.

References

Altpeter, F., Springer, N.M., Bartley, L.E., Blechl, A.E., Brutnell, T.P., Citovsky, V., Conrad, L.J. et al. (2016) Advancing crop transformation in the era of genome editing. *Plant Cell*, **28**, 1510–1520.

- Apuli, R.P., Bernhardtsson, C., Schiffthaler, B., Robinson, K.M., Jansson, S., Street, N.R. and Ingvarsson, P.K. (2020) 'Inferring the genomic landscape of recombination rate variation in European Aspen (*Populus tremula*). *G3: Genes - Genomes - Genetics*, **10**, 299–309.
- Arrieta, M., Willems, G., DePessemer, J., Colas, I., Burkholz, A., Darracq, A., Vanstraelen, S. et al. (2020) The effect of heat stress on sugar beet recombination. *Theor. Appl. Genet.* **13**, 81–93.
- Baker, K., Bayer, M., Cook, N., Dreißig, S., Dhillon, T., Russell, J., Hedley, P.E. et al. (2014) The low-recombining pericentromeric region of barley restricts gene diversity and evolution but not gene expression. *Plant J.* **79**, 981–992.
- Bauer, E., Falque, M., Walter, H., Bauland, C., Camisan, C., Campo, L., Meyer, N. et al. (2013) Intraspecific variation of recombination rate in maize. *Genome Biol.* **14**, R103.
- Bayer, M.M., Rapazote-Flores, P., Ganai, M., Hedley, P.E., Macaulay, M., Plieske, J., Ramsay, L. et al. (2017) Development and evaluation of a barley 50k iSelect SNP array. *Front. Plant Sci.* **8**, 1792.
- Berloff (2002) Physical and genetic maps. *J. Comput. Biol.* **9**, 465–475.
- Beyer, P., Morell, M., Mackay, I. and Powell, W. (2008) From mutations to MAGIC: resources for gene discovery, validation and delivery in crop plants. *Curr. Opin. Plant Biol.* **11**, 215–221.
- 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**, 297–309.
- Broman, K.W., Wu, H., Sen, S. and Churchill, G.A. (2003) R/qtl: QTL mapping in experimental crosses. *Bioinformatics*, **19**, 889–890.
- Cederberg, H. (1985) Recombination in other chromosomal regions than the interval subjected to selection, in lines of *Neurospora crassa* selected for high and for low recombination frequency. *Heredity*, **103**, 89–97.
- Chakravarti, A. (1991) A graphical representation of genetic and physical maps: the Marey map. *Genomics*, **11**, 219–222.
- Chan, A.H., Jenkins, P.A. and Song, Y.S. (2012) Genome-wide fine-scale recombination rate variation in *Drosophila melanogaster*. *PLoS Genet.* **8**, e1003090.
- Choi, K. and Henderson, I.R. (2015) Meiotic recombination hotspots – a comparative view. *Plant J.* **83**, 52–61.
- Choi, K., Zhao, X., Kelly, K.A., Venn, O., Higgins, J.D., Yelina, N.E., Hardcastle, T.J. et al. (2013) Arabidopsis meiotic crossover hot spots overlap with H2A.Z nucleosomes at gene promoters. *Nat. Genet.* **45**, 1327–1338.
- Choi, K., Zhao, X., Tock, A.J., Lambing, C., Underwood, C.J., Hardcastle, T.J., Serra, H. et al. (2018) Nucleosomes and DNA methylation shape meiotic DSB frequency in Arabidopsis thaliana transposons and gene regulatory regions. *Genome Res.* **28**, 532–546.
- Coop, G., Wen, X., Ober, C., Pritchard, J.K. and Przeworski, M. (2008) High-resolution mapping of crossovers reveals extensive variation in fine-scale recombination patterns among humans. *Science*, **319**, 1395–1398.
- Covarrubias-Pazarán, G. (2016) Genome-assisted prediction of quantitative traits using the R package sommer. *PLoS One*, **11**, e0156744.
- Darrier, B., Rimbaut, H., Balfourier, F., Pingault, L., Josselin, A.A., Servin, B., Navarro, J. et al. (2017) High-resolution mapping of crossover events in the hexaploid wheat genome suggests a universal recombination mechanism. *Genetics*, **206**, 1373–1388.
- Dowrick, G.J. (1957) The influence of temperature on meiosis. *Heredity*, **11**, 37–49.
- Dreißig, S., Fuchs, J., Cápai, P., Kettles, N., Byrne, E. and Houben, A. (2015) Measuring meiotic crossovers via multi-locus genotyping of single pollen grains in barley. *PLoS One*, **10**, 1–10.
- Dreißig, S., Mascher, M., Heckmann, S. and Purugganan, M. (2019) Variation in recombination rate is shaped by domestication and environmental conditions in barley. *Mol. Biol. Evol.* **36**, 2029–2039.
- Dreißig, S., Maurer, A., Sharma, R., Milne, L., Flavell, A.J., Schmutzer, T. and Pillen, K. (2020) Natural variation in meiotic recombination rate shapes introgression patterns in intraspecific hybrids between wild and domesticated barley. *New Phytol.* **228**, 1852–1863.
- Drouaud, J., Khademan, H., Giraut, L., Zanni, V., Bellalou, S., Henderson, I.R., Falque, M. et al. (2013) Contrasted patterns of crossover and non-crossover at *Arabidopsis thaliana* meiotic recombination hotspots. *PLoS Genet.* **9**, e1003922.
- Dumont, B.L., Broman, K.W. and Payseur, B.A. (2009) Variation in genomic recombination rates among heterogeneous stock mice. *Genetics*, **182**, 1345–1349.
- Esch, E., Szymaniak, J.M., Yates, H., Pawlowski, W.P. and Buckler, E.S. (2007) Using crossover breakpoints in recombinant inbred lines to identify quantitative trait loci controlling the global recombination frequency. *Genetics*, **177**, 1851–1858.
- Falconer, D.S. and Mackay, T.F.C. (1996) *Introduction to Quantitative Genetics*, 4th ed. Essex, UK: Longman.
- Fledel-Alon, A., Leffler, E.M., Guan, Y., Stephens, M., Coop, G. and Przeworski, M. (2011) Variation in human recombination rates and its genetic determinants. *PLoS One*, **6**, e20321.
- Gardiner, L.J., Wingen, L.U., Bailey, P., Joynson, R., Brabbs, T., Wright, J., Higgins, J.D. et al. (2019) Analysis of the recombination landscape of hexaploid bread wheat reveals genes controlling recombination and gene conversion frequency. *Genome Biol.* **20**, 1.
- Garin, V., Wimmer, V. and Malosetti, M. (2015) *mppR: an R package for QTL analysis in multi-parent populations using linear mixed models*. R Vignette. https://cran.r-project.org/package=mppR/vignettes/mppR_vignette.pdf
- Gion, J.M., Hudson, C.J., Lesur, I., Vaillancourt, R.E., Potts, B.M. and Freeman, J.S. (2016) Genome-wide variation in recombination rate in Eucalyptus. *BMC Genom.* **17**, 1–12.
- Haile, J.K., N'Diaye, A., Clarke, F., Clarke, J., Knox, R., Rutkoski, J., Bassi, F.M. et al. (2018) Genomic selection for grain yield and quality traits in durum wheat. *Mol. Breed.* **38**, 75.
- Haseneyer, G., Stracke, S., Paul, C., Einfeldt, C., Broda, A., Piepho, H.-P., Graner, A. et al. (2009) Population structure and phenotypic variation of a spring barley world collection set up for association studies. *Plant Breed.* **129**, 271–279.
- Hayta, S., Smedley, M.A., Demir, S.U., Blundell, R., Hinchliffe, A., Atkinson, N. and Harwood, W.A. (2019) An efficient and reproducible *Agrobacterium*-mediated transformation method for hexaploid wheat (*Triticum aestivum* L.). *Plant Methods*, **15**, 1–15.
- Hayut, S.F., Bessudo, C.M. and Levy, A.A. (2017) Targeted recombination between homologous chromosomes for precise breeding in tomato. *Nat. Commun.* **8**, 15605.
- Heffner, E.L., Jannink, J.L., Iwata, H., Souza, E. and Sorrells, M.E. (2011) Genomic selection accuracy for grain quality traits in biparental wheat populations. *Crop Sci.* **51**, 2597–2606.
- Henderson, I.R. (2012) Control of meiotic recombination frequency in plant genomes. *Curr. Opin. Plant Biol.* **15**, 556–561.
- Higgins, J.D., Perry, R.M., Barakate, A., Ramsay, L., Waugh, R., Halpin, C., Armstrong, S.J. et al. (2012) Spatiotemporal asymmetry of the meiotic program underlies the predominantly distal distribution of meiotic crossovers in barley. *Plant Cell*, **24**, 4096–4109.
- Hunter, C.M., Huang, W., Mackay, T.F. and Singh, N.D. (2016) The genetic architecture of natural variation in recombination rate in *Drosophila melanogaster*. *PLoS Genet.* **12**, 1–31.
- Jordan, K.W., Wang, S., He, F., Chao, S., Lun, Y., Paux, E., Sourdille, P. et al. (2018) The genetic architecture of genome-wide recombination rate variation in allopolyploid wheat revealed by nested association mapping. *Plant J.* **95**, 1039–1054.
- Kim, S., Plagnol, V., Hu, T.T., Toomajian, C., Clark, R.M., Ossowski, S., Ecker, J.R. et al. (2007) Recombination and linkage disequilibrium in *Arabidopsis thaliana*. *Nat. Genet.* **39**, 1151–1155.
- Kong, A., Thorgeirsson, G., Gudbjartsson, D.F., Masson, G., Sigurdsson, A., Jonasdottir, A.A., Walters, G.B. et al. (2010) Fine-scale recombination rate differences between sexes, populations and individuals. *Nature*, **467**, 1099–1103.
- Lee, M., Sharopova, N., Beavis, W.D., Grant, D., Katt, M., Blair, D. and Hallauer, A. (2002) Expanding the genetic map of maize with the intermated B73 x Mo17 (IBM) population. *Plant Mol. Biol.* **48**, 453–461.
- Li, N. and Stephens, M. (2003) Modeling linkage disequilibrium and identifying recombination hotspots using single-nucleotide polymorphism data. *Genetics*, **165**, 2213–2233.
- Lloyd, A., Morgan, C., Franklin, F.C.H. and Bomblies, K. (2018) Plasticity of meiotic recombination rates in response to temperature in arabidopsis. *Genetics*, **208**, 1409–1420.

- Lorenzana, R.E. and Bernardo, R. (2009) Accuracy of genotypic value predictions for marker-based selection in biparental plant populations. *Theor. Appl. Genet.* **120**, 151–161.
- Marand, A.P., Zhao, H., Zhang, W., Zeng, Z., Fang, C. and Jianga, J. (2019) Historical meiotic crossover hotspots fueled patterns of evolutionary divergence in rice. *Plant Cell*, **31**, 645–662.
- Mascher, M., Gundlach, H., Himmelbach, A., Beier, S., Twardziok, S.O., Wicker, T., Radchuk, V. et al. (2017) A chromosome conformation capture ordered sequence of the barley genome. *Nature*, **544**, 427–433.
- McMullen, M.D., Kresovich, S., Villeda, H.S., Bradbury, P., Li, H., Sun, Q., Flint-Garcia, S. et al. (2009) Genetic properties of the maize nested association mapping population. *Science*, **325**, 737–740.
- Melamed-Bessudo, C., Shilo, S. and Levy, A.A. (2016) Meiotic recombination and genome evolution in plants. *Curr. Opin. Plant Biol.* **30**, 82–87.
- Mercier, R., Mézard, C., Jenczewski, E., Macaisne, N. and Grelon, M. (2015) The molecular biology of meiosis in plants. *Annu. Rev. Plant Biol.* **66**, 297–327.
- Meuwissen, T.H., Hayes, B.J. and Goddard, M.E. (2001) Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, **157**, 1819–1829.
- Meznar, E.R., Gadau, J., Koeniger, N. and Rueppell, O. (2010) Comparative linkage mapping suggests a high recombination rate in all honeybees. *J. Hered.* **101**, S118–S126.
- Mieulet, D., Aubert, G., Bres, C., Klein, A., Droc, G., Vieille, E., Rond-Coissieux, C. et al. (2018) Unleashing meiotic crossovers in crops. *Nat. Plants*, **4**, 1010–1016.
- Modliszewski, J.L., Wang, H., Albright, A.R., Lewis, S.M., Bennett, A.R., Huang, J., Ma, H. et al. (2018) Elevated temperature increases meiotic crossover frequency via the interfering (Type I) pathway in *Arabidopsis thaliana*. *PLoS Genet.* **14**, 1–15.
- Monat, C., Padmarasu, S., Lux, T., Wicker, T., Gundlach, H., Himmelbach, A., Ens, J. et al. (2019) TRITEX: chromosome-scale sequence assembly of Triticeae genomes with open-source tools. *Genome Biol.* **20**, 284.
- Morgan, H. (1916) *A Critique of the Theory of Evolution*. Princeton, NJ: Princeton University Press.
- Muñoz-Amatrián, M., Eichten, S.R., Wicker, T., Richmond, T.A., Mascher, M., Steuernagel, B., Scholz, U. et al. (2013) Distribution, functional impact, and origin mechanisms of copy number variation in the barley genome. *Genome Biol.* **14**, R58.
- Nachman, M.W. (2002) Variation in recombination rate across the genome: evidence and implications. *Curr. Opin. Genet. Dev.* **12**, 657–663.
- Perperoglou, A., Sauerbrei, W., Abrahamowicz, M. and Schmid, M. (2019) A review of spline function procedures in R. *BMC Med. Res. Methodol.* **19**, 1–16.
- Petit, M., Astruc, J.-M., Sarry, J., Drouilhet, L., Fabre, S., Moreno, C.R. and Servin, B. (2017) Variation in recombination rate and its genetic determinism in sheep populations. *Genetics*, **207**, 767–784.
- Phillips, D., Jenkins, G., Macaulay, M., Nibau, C., Wnetrzak, J., Fallding, D., Colas, I. et al. (2015) The effect of temperature on the male and female recombination landscape of barley. *New Phytol.* **208**, 421–429.
- Plough, H.H. (1917) The effect of temperature on crossingover in *Drosophila*. *J. Exp. Zool.* **24**, 147–209.
- Rodgers-Melnick, E., Bradbury, P.J., Elshire, R.J., Glaubitz, J.C., Acharya, C.B., Mitchell, S.E., Li, C. et al. (2015) Recombination in diverse maize is stable, predictable, and associated with genetic load. *Proc. Natl Acad. Sci. USA*, **112**, 3823–3828.
- Rowan, B.A., Heavens, D., Feuerborn, T.R., Tock, A.J., Henderson, I.R. and Weigel, D. (2019) An ultra high-density *Arabidopsis thaliana* crossover. *Genetics*, **213**, 771–787.
- Salomé, P.A., Bombliès, K., Fitz, J., Laitinen, R.A., Warthmann, N., Yant, L. and Weigel, D. (2012) The recombination landscape in *Arabidopsis thaliana* F2 populations. *Heredity (Edinb)*, **108**, 447–455.
- Sandor, C., Li, W., Coppieters, W., Druet, T., Charlier, C. and Georges, M. (2012) Genetic variants in REC8, RNF212, and PRDM9 influence male recombination in cattle. *PLoS Genet.* **8**, e1002854.
- Sarno, R., Vicq, Y., Uematsu, N., Luka, M., Lapierre, C., Carroll, D., Bastianelli, G. et al. (2017) Programming sites of meiotic crossovers using Spo11 fusion proteins. *Nucleic Acids Res.* **45**, e164.
- Saxena, R.K., Edwards, D. and Varshney, R.K. (2014) Structural variations in plant genomes. *Brief. Funct. Genomics*, **13**, 296–307.
- Shen, C., Wang, N., Huang, C., Wang, M., Zhang, X. and Lin, Z. (2019) Population genomics reveals a fine-scale recombination landscape for genetic improvement of cotton. *Plant J.* **99**, 494–505.
- Smeds, L., Mugal, C.F., Qvarnström, A. and Ellegren, H. (2016) High-resolution mapping of crossover and non-crossover recombination events by whole-genome re-sequencing of an avian pedigree. *PLoS Genet.* **12**, 1–23.
- Stich, B. (2009) Comparison of mating designs for establishing nested association mapping populations in maize and *Arabidopsis thaliana*. *Genetics*, **183**, 1525–1534.
- Taagen, E., Bogdanove, A.J. and Sorrells, M.E. (2020) Counting on crossovers: controlled recombination for plant breeding. *Trends Plant Sci.* **25**, 455–465.
- Tam, S.M., Hays, J.B. and Chetelat, R.T. (2011) Effects of suppressing the DNA mismatch repair system on homeologous recombination in tomato. *Theor. Appl. Genet.* **123**, 1445–1458.
- Technow, F., Bürger, A. and Melchinger, A.E. (2013) Genomic prediction of northern corn leaf blight resistance in maize with combined or separated training sets for heterotic groups. *G3: Genes - Genomes - Genetics*, **3**, 197–203.
- The World Bank (2020) *World Development Indicators*. Accessed June 2020. <https://data.worldbank.org/>.
- Tiley, G.P. and Burleigh, G. (2015) The relationship of recombination rate, genome structure, and patterns of molecular evolution across angiosperms. *BMC Evol. Biol.* **15**, 1–14.
- Underwood, C.J., Choi, K., Lambing, C., Zhao, X., Serra, H., Borges, F., Simorowski, J. et al. (2018) Epigenetic activation of meiotic recombination near *Arabidopsis thaliana* centromeres via loss of H3K9me2 and non-CG DNA methylation. *Genome Res.* **28**, 519–531.
- VanRaden, P.M., Van Tassell, C.P., Wiggans, G.R., Sonstegard, T.S., Schnabel, R.D., Taylor, J.F. and Schenkel, F.S. (2009) Invited review: reliability of genomic predictions for north american Holstein bulls. *J. Dairy Sci.* **92**, 16–24.
- Wang, Y. and Copenhaver, G.P. (2018) Meiotic recombination: mixing it up in plants. *Annu. Rev. Plant Biol.* **69**, 577–609.
- Weisweiler, M., Montaigu, A.D., Ries, D., Pfeifer, M. and Stich, B. (2019) Transcriptomic and presence/absence variation in the barley genome assessed from multi-tissue mRNA sequencing and their power to predict phenotypic traits. *BMC Genom.*, **20**, 1–15.
- Weng, Z.Q., Saatchi, M., Schnabel, R.D., Taylor, J.F. and Garrick, D.J. (2014) Recombination locations and rates in beef cattle assessed from parent-offspring pairs. *Genet. Sel. Evol.* **46**, 1–14.
- Yandeu-Nelson, M.D., Nikolau, B.J. and Schnabel, P.S. (2006) Effects of trans-acting genetic modifiers on meiotic recombination across the a1-sh2 interval of maize. *Genetics*, **174**, 101–112.
- Yu, J., Holland, J.B., McMullen, M.D. and Buckler, E.S. (2008) Genetic design and statistical power of nested association mapping in maize. *Genetics*, **178**, 539–551.
- Ziolkowski, P.A., Underwood, C.J., Lambing, C., Martinez-Garcia, M., Lawrence, E.J., Ziolkowska, L., Griffin, C. et al. (2017) Natural variation and dosage of the HEI10 meiotic E3 ligase control *Arabidopsis* crossover recombination. *Genes Dev.* **31**, 306–317.
- Zou, G.Y. (2007) Toward using confidence intervals to compare correlations. *Psychol. Methods*, **12**, 399–413.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1 The segregation pattern of the parental alleles across the chromosomes of the 45 double round-robin populations.

Figure S2 The bins' genetic map of the seven barley chromosomes across the double round-robin populations.

Figure S3 The Marey maps of the seven barley chromosomes across the double round-robin populations.

690 Federico Casale et al.

Figure S4 The recombination landscape of the seven barley chromosomes across the double round-robin populations.

Figure S5 Pearson's correlation coefficient between the GRE_p values of the parental inbreds for 10 Mbp physical windows and their respective genome-wide GRE_p values across the seven barley chromosomes.

Figure S6 The sequence similarity between the parental inbreds across the seven barley chromosomes of the three double round-robin populations with the highest and lowest genomic recombination rate.

Figure S7 The effect of the QTLs associated with the number of crossovers on chromosome 2H across the 45 double round-robin populations.

Figure S8 The effect of the QTLs associated with the number of crossovers on chromosome 5H across the 45 double round-robin populations.

Figure S9 The effect of the QTLs associated with the number of crossovers on chromosome 7H across the 45 double round-robin populations.

Figure S10 The effect of the QTLs associated with the number of crossovers on the genome across the 45 double round-robin populations.

Figure S11 The distribution of SNPs' effects predicted by RR-BLUP across the genome.

Figure S12 The genomic prediction ability of recombination rate in 10 Mbp window level across the genome, using different cross-validation scenarios.

Figure S13 Genomic prediction ability concerning the recombination rate variation of individual chromosomes and the genome-wide level, using different approaches and subsets of equally spaced SNPs.

Table S1 Summary table of the QTLs detected for crossovers count of chromosomes 2H, 5H, and 7H, and genome-wide using a multi-population analysis.

Table S2 Pearson's correlation between the phenotypic trait mean and the GRE of the parental inbreds (r_{TGRE}) for different agronomic traits of barley.

Method S1 Data cleaning.

Method S2 Linkage map construction.

Method S3 Assessment of phenotypic traits.

SUPPORTING INFORMATION

Methods S1: Data cleaning

SNPs with a GenTrain score lower than 0.4 were excluded. In each DRR population, monomorphic SNPs and those with a minor allele frequency (MAF) < 10% were discarded. Further, SNPs with > 20% of missing data or exhibiting heterozygosity > 20% were eliminated from the data set. SNPs showing segregation distortion were identified with a Chi-square (X^2) Goodness-of-fit test ($P < 0.05$) after applying the Bonferroni correction for multiple tests and were excluded from map construction (Ott and Longnecker, 2006). A segregation distortion region (SDR) was considered large if it occupied a segment longer than 50 Mbp in which more than 80% of the markers were found significantly distorted. In the same way as SNPs, RILs with > 20% residual heterozygosity or missing data were discarded. RILs carrying nonparental alleles were also removed from the dataset. The resulting dataset comprised 36,077 SNPs across the total set of DRR populations with 34 to 143 RILs per population. A principal coordinate analysis was performed on the DRR populations and their parental inbreds based on the modified Roger's distance calculated from that dataset. Following the same procedure, a principal coordinate analysis of the diversity panel, Morex, three *spp. spontaneum*, and one *spp. agriocrithon* accessions, was performed.

In the context of assessing the historical recombination, monomorphic SNPs and SNPs with a MAF < 5% were removed from the dataset of the diversity panel. Moreover, SNPs with > 10% heterozygotes and > 20% missing data were discarded. For those individuals with > 10% heterozygote SNPs, the SNP genotypes were replaced by NA. Only SNPs with a position on the physical reference map were retained (Monat et al., 2019). This cleaning process resulted in a total of 35,980 SNPs.

Methods S2: Linkage map construction

Only one of the perfectly co-segregating SNPs (referred to as duplicated SNPs) was retained to increase computational efficiency. To start the map construction, SNPs in each DRR population were categorized to seven linkage groups using the R (R Core Team, 2020) package “qtl” (Broman et al., 2003). Then, Carthagene (de Givry et al., 2005) was used to create the genetic map of each linkage group in four steps: sketch, scaffold, framework, and addition of the remaining markers. First, a comprehensive map for each linkage group was sketched using “mfmapd” function. SNPs 10 centimorgans (cM) apart on the sketch map were selected to construct the scaffold map of the respective linkage group, using “buildfw” function. The resulting scaffold map was re-ordered with the “flips” command. In the next step, the remaining SNPs of each linkage group were added to the respective scaffold map using the “buildfw” function to construct high-density framework maps followed by using the “flips” function to improve the order. This addition of

SNPs and the ordering was repeated twice with decreasing LOD thresholds. After the framework map's construction, the only SNPs that remained unmapped were those that did not produce a change in the log-likelihood of the map after being added to the framework map. These SNPs were then added to the respective framework map at their best position. Finally, duplicated SNPs were included in their co-segregating bin to obtain the final map. The genetic distances were estimated using Haldane's mapping function (Haldane, 1919).

Methods S3: Assessment of phenotypic traits

For the assessment of phenotypic traits under field conditions, the 23 parental inbreds were planted as replicated check genotypes in an experiment with other entries which was laid out as an augmented row-column design. This experiment was repeated in seven environments in Germany: Cologne (2017, 2018, and 2019), Mechernich (2018 and 2019), and Quedlinburg (2018 and 2019). The leaf angle was assessed in four-week-old plants, the heading date was scored as the number of days after planting (DAP), and the plant height was assessed after heading. The seed traits were measured with MARVIN seed analyzer (GTA Sensorik, Neubrandenburg, Germany).

REFERENCES

- Broman, K. W., Wu, H., Sen, S. and Churchill, G. A. (2003), ‘R/qtl: QTL mapping in experimental crosses’, *Bioinformatics* **19**, 889–890.
- de Givry, S., Bouchez, M., Chabrier, P., Milan, D. and Schiex, T. (2005), ‘CARTHAGENE: multipopulation integrated genetic and radiation hybrid mapping’, *Bioinformatics* **21**, 1703–1704.
- Haldane, J. B. (1919), ‘The combination of linkage values, and the calculation of distances between the loci of linked factors’, *J. Genet.* **8**, 299–309.
- Monat, C., Padmarasu, S., Lux, T., Wicker, T., Gundlach, H., Himmelbach, A., Ens, J., Li, C., Muehlbauer, G. J., Schulman, A. H., Waugh, R., Braumann, I., Pozniak, C., Scholz, U., Mayer, K. F., Spannagl, M., Stein, N. and Mascher, M. (2019), ‘TRITEX: chromosome-scale sequence assembly of Triticeae genomes with open-source tools’, *Genome Biol.* **20**, 284.
- Ott, R. L. and Longnecker, M. T. (2006), *Introduction to statistical methods and data analysis*, Duxbury Press, USA.
- R Core Team (2020), *R: a language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria.

Table S1: Summary table of the QTLs detected for crossovers count of chromosomes 2H, 5H, and 7H, and genome-wide using a multi-population analysis. R^2 is the percentage of the explained phenotypic variance. N populations is the number of double round-robin populations in which the QTL was found significant.

QTL	Phenotype	Chromosome	Physical interval	R^2	N populations
1	chr: 2H	2H	19,706,917 – 21,564,939	2.13	9
2	chr: 2H	3H	540,528,678 – 549,939,046	2.14	10
3	chr: 2H	4H	476,635,174 – 496,775,512	2.17	12
4	chr: 2H	5H	526,241,270 – 528,726,782	2.95	9
5	chr: 2H	6H	478,413,233 – 501,926,097	2.84	11
6	chr: 2H	7H	5,146,581 – 13,654,310	1.92	10
7	chr: 5H	2H	5,283,335 – 10,970,690	2.19	9
8	chr: 5H	5H	447,633,911 – 523,261,050	2.17	10
9	chr: 5H	5H	484,772,895 – 488,160,271	3.27	9
10	chr: 5H	5H	585,716,037 – 588,620,278	2.49	11
11	chr: 7H	4H	529,604,097 – 557,962,172	2.24	9
12	chr: 7H	5H	533,736,593 – 548,597,684	2.79	11
13	chr: 7H	7H	632,263,733 – 632,995,979	2.19	5
14	genome-wide	1H	477,085,105 – 478,327,181	2.60	7
15	genome-wide	4H	580,387,883 – 586,286,949	2.57	10
16	genome-wide	5H	516,799,912 – 519,129,769	2.69	10

Table S2: Pearson's correlation between the phenotypic trait mean and the GRE of the parental inbreds (r_{TGRE}) for different agronomic traits of barley.

Trait	r_{TGRE}
Leaf angle	-0.15^{ns}
Heading date	-0.31^{ns}
Plant height	-0.11^{ns}
Thousand seed weight	-0.25^{ns}
Seed length	-0.07^{ns}
Seed width	-0.24^{ns}
Seed area	-0.19^{ns}

ns, not significant

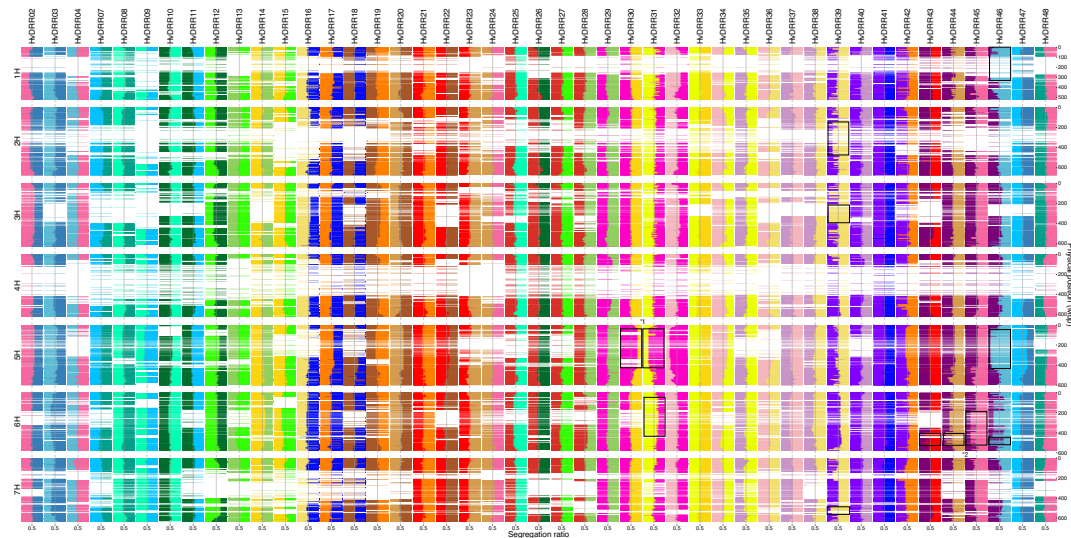


Fig. S1: The segregation pattern of the parental alleles across the chromosomes of the 45 double round-robin populations. For each SNP, the allele frequency of the two parental alleles in each population is displayed as colored horizontal bar. A particular parental inbred is represented by the same color across all populations. Large regions with significant segregation distortion (SDR) are marked as dark rectangles, where an SDR that favored a common parent is identified with *1, and an SDR that disfavored a common parent is identified with *2.

9

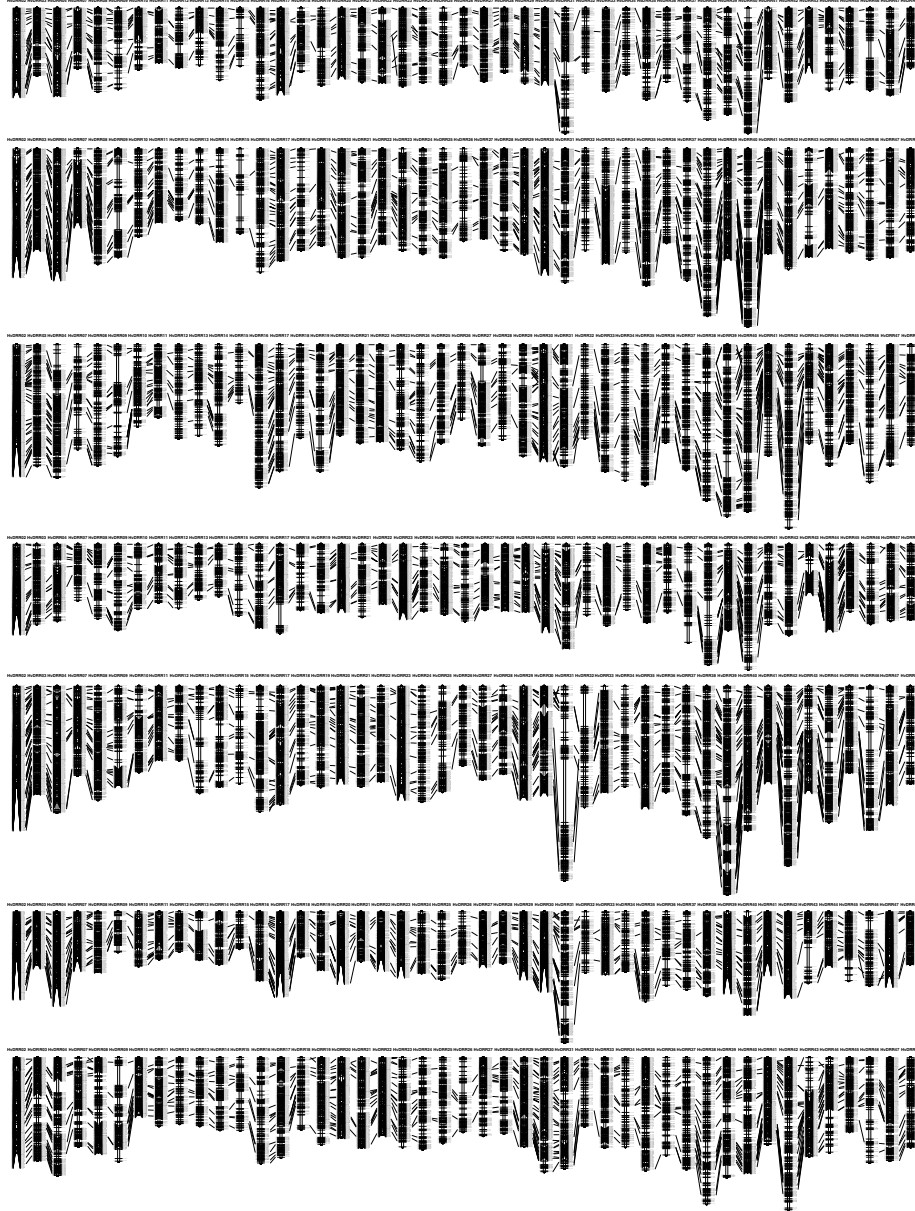


Fig. S2: The bins' genetic map of the seven barley chromosomes across the double round-robin populations.

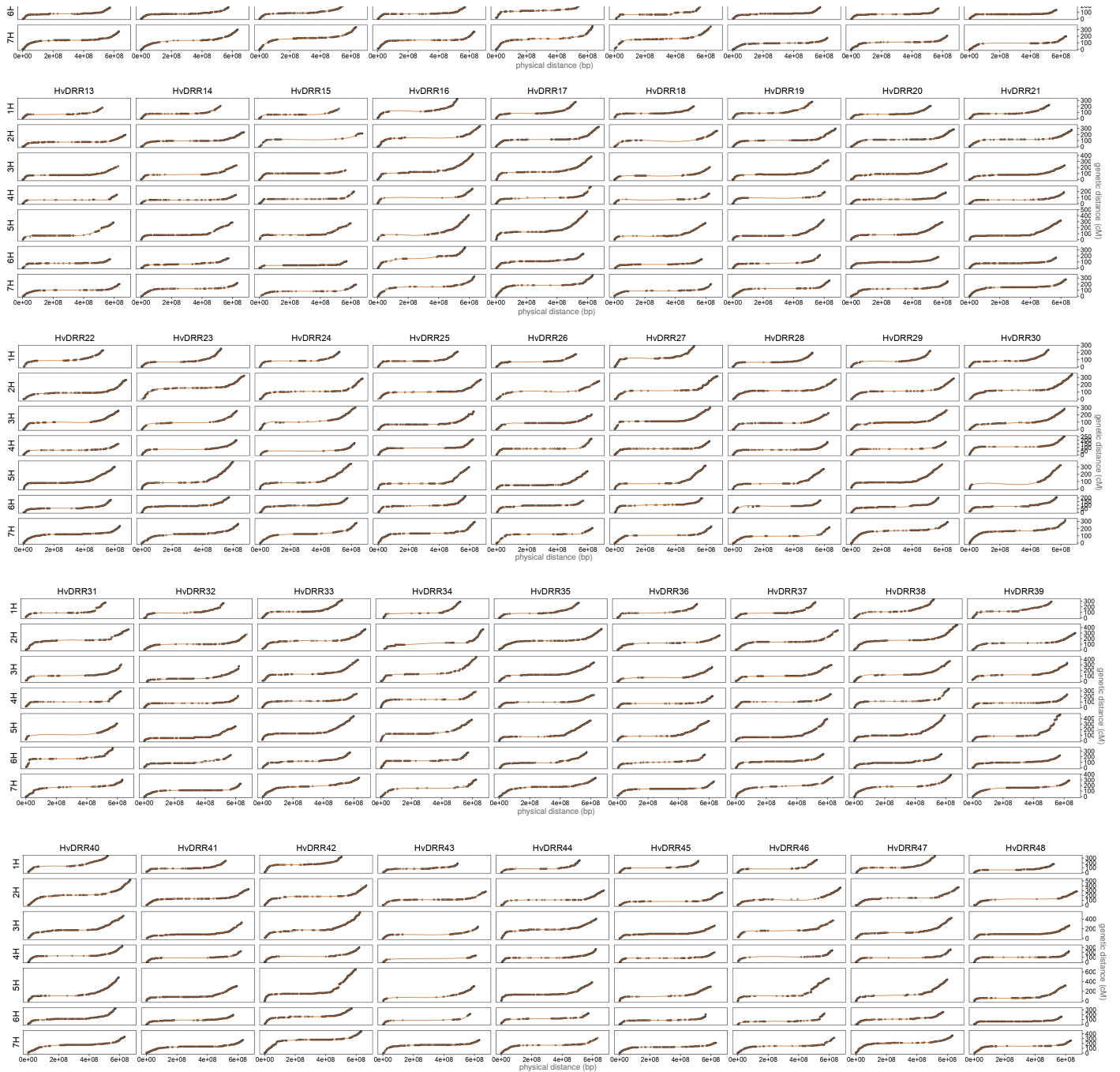


Fig. S3: The Marey maps of the seven barley chromosomes across the double round-robin populations.

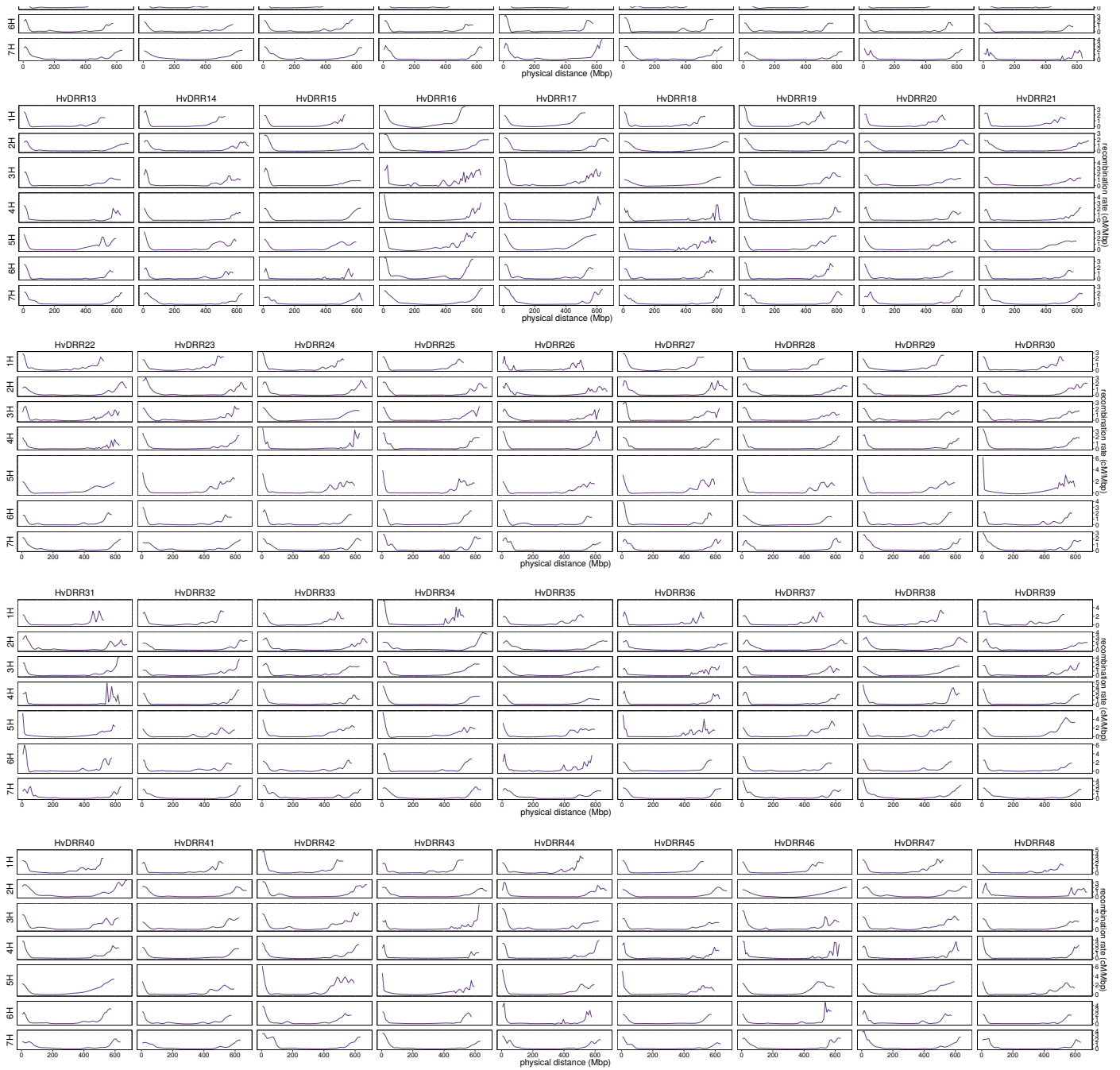


Fig. S4: The recombination landscape of the seven barley chromosomes across the double round-robin populations.

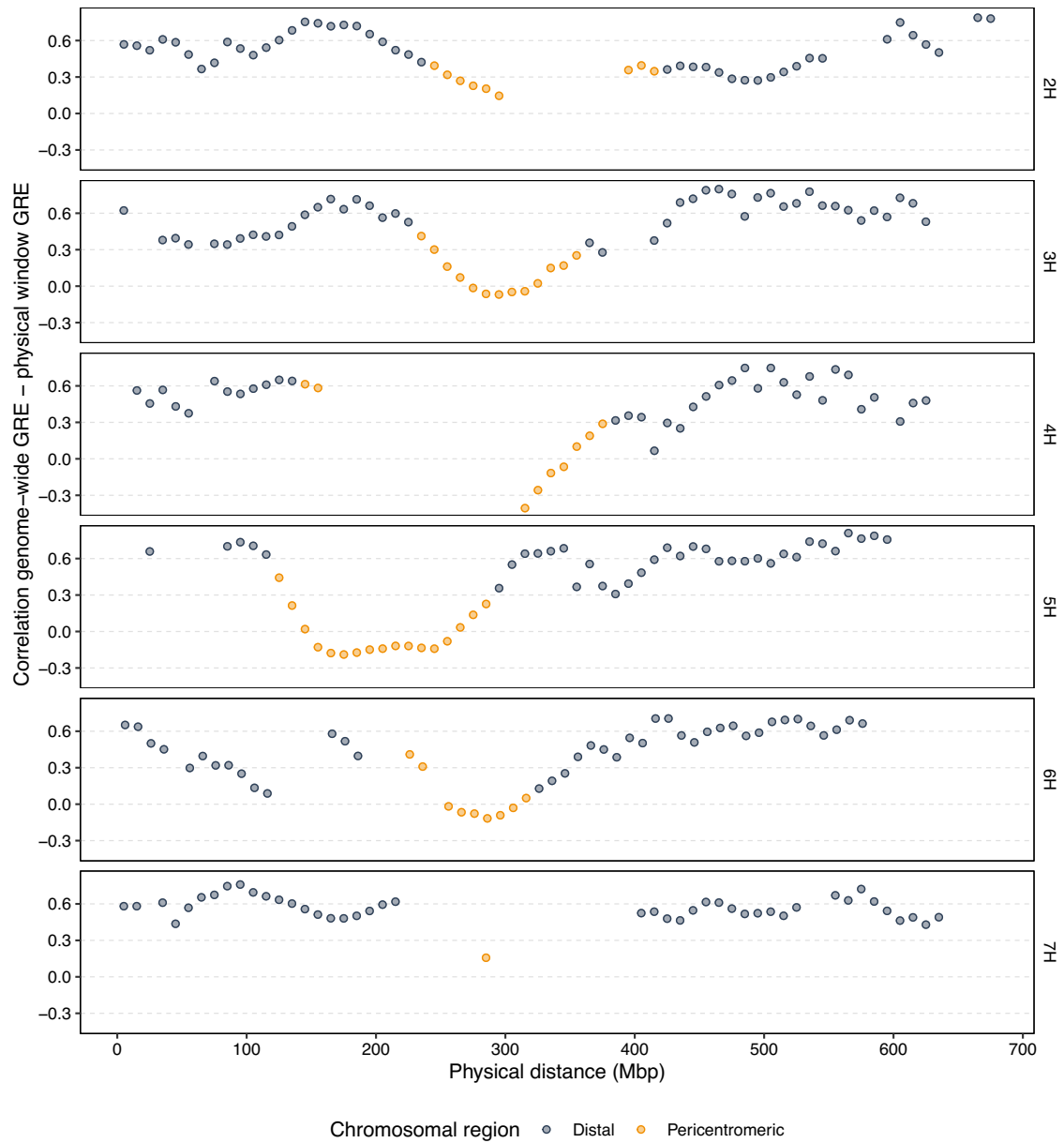


Fig. S5: Pearson's correlation coefficient between the GRE_P values of the parental inbreds for 10 Mbp physical windows and their respective genome-wide GRE_P values across the seven barley chromosomes.

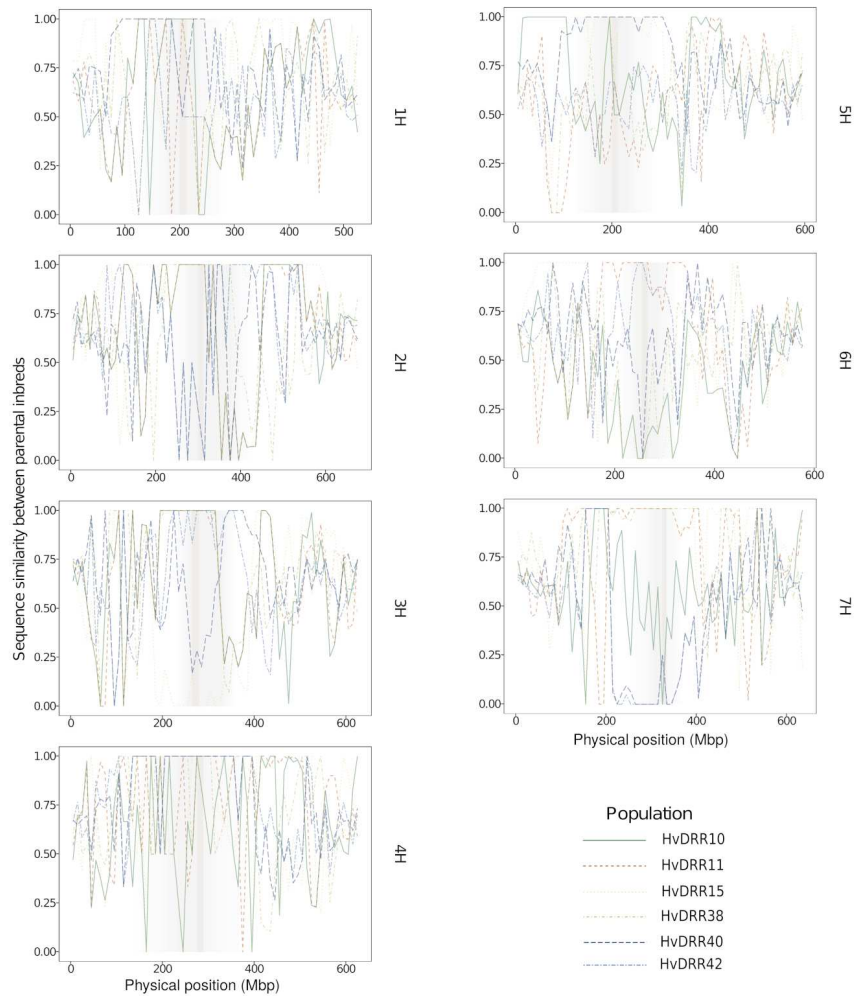


Fig. S6: The sequence similarity between the parental inbreds across the seven barley chromosomes of the three double round-robin populations with the highest and lowest genomic recombination rate. The vertical line in the background indicates the centromere position in the reference map, and the expanded shadow indicates the pericentromeric region in the consensus map.

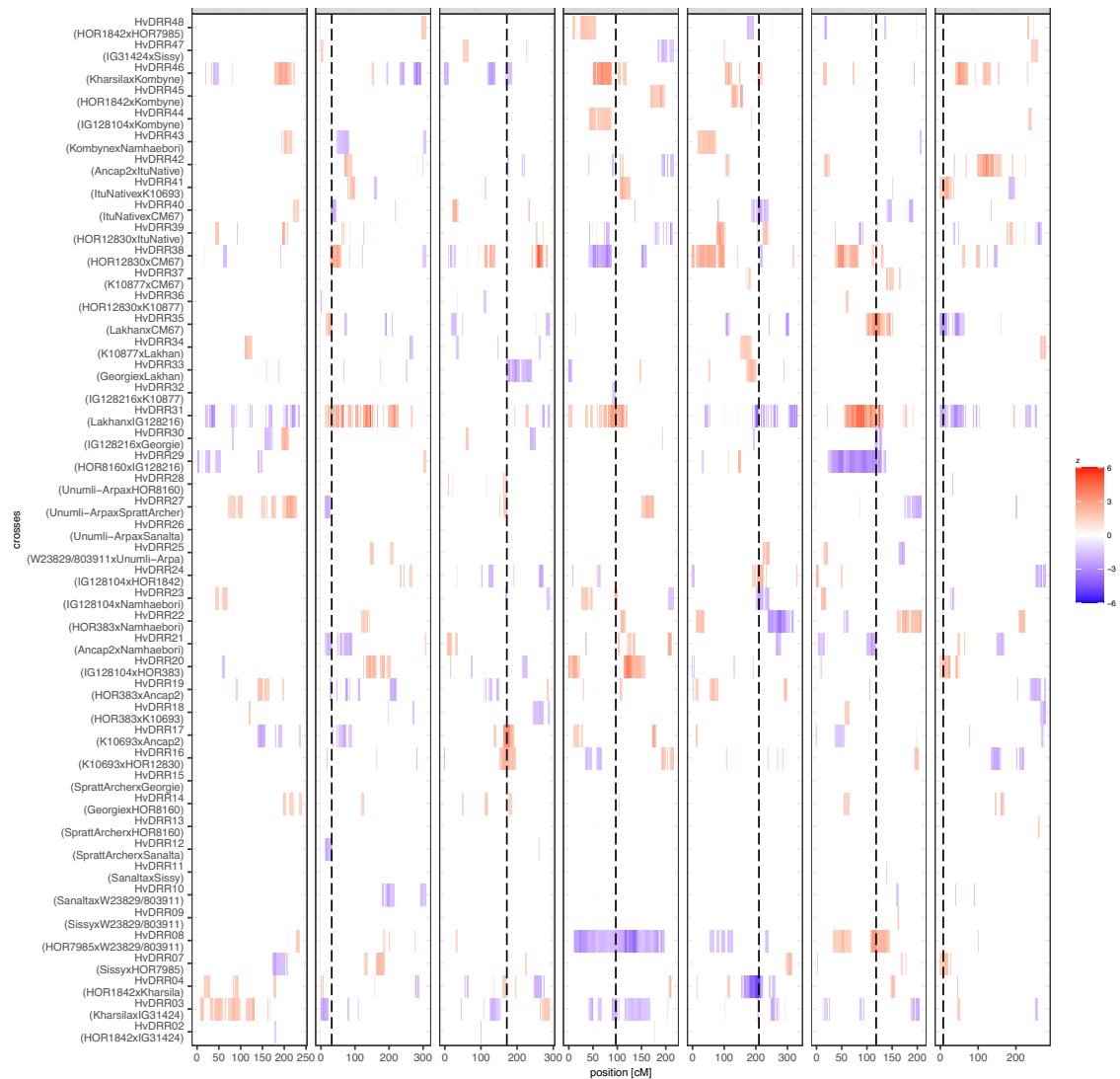


Fig. S7: The effect of the QTLs associated with the number of crossovers on chromosome 2H across the 45 double round-robin populations. The QTLs' locations are identified with a dashed black line.

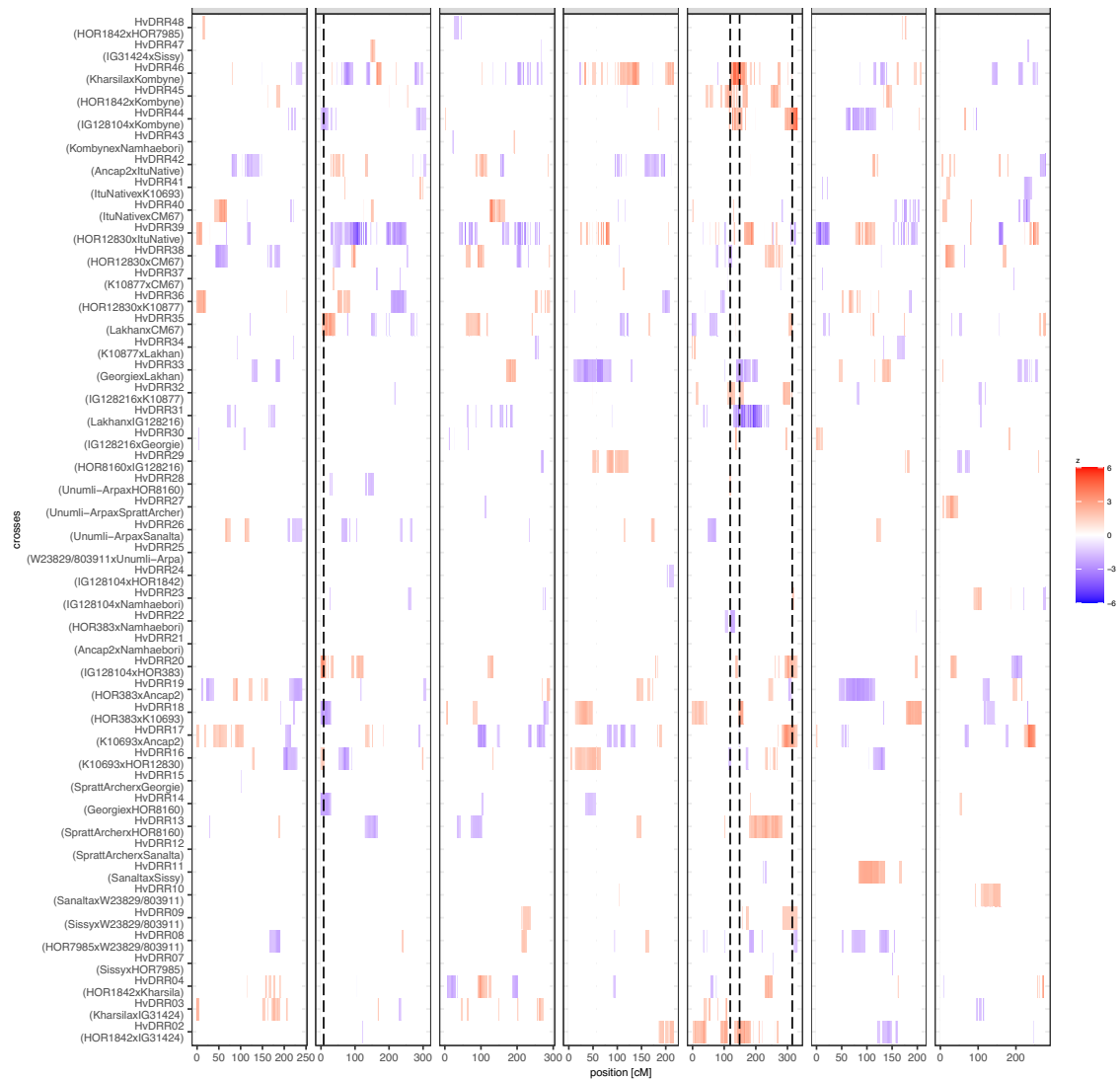


Fig. S8: The effect of the QTLs associated with the number of crossovers on chromosome 5H across the 45 double round-robin populations. The QTLs' locations are identified with a dashed black line.

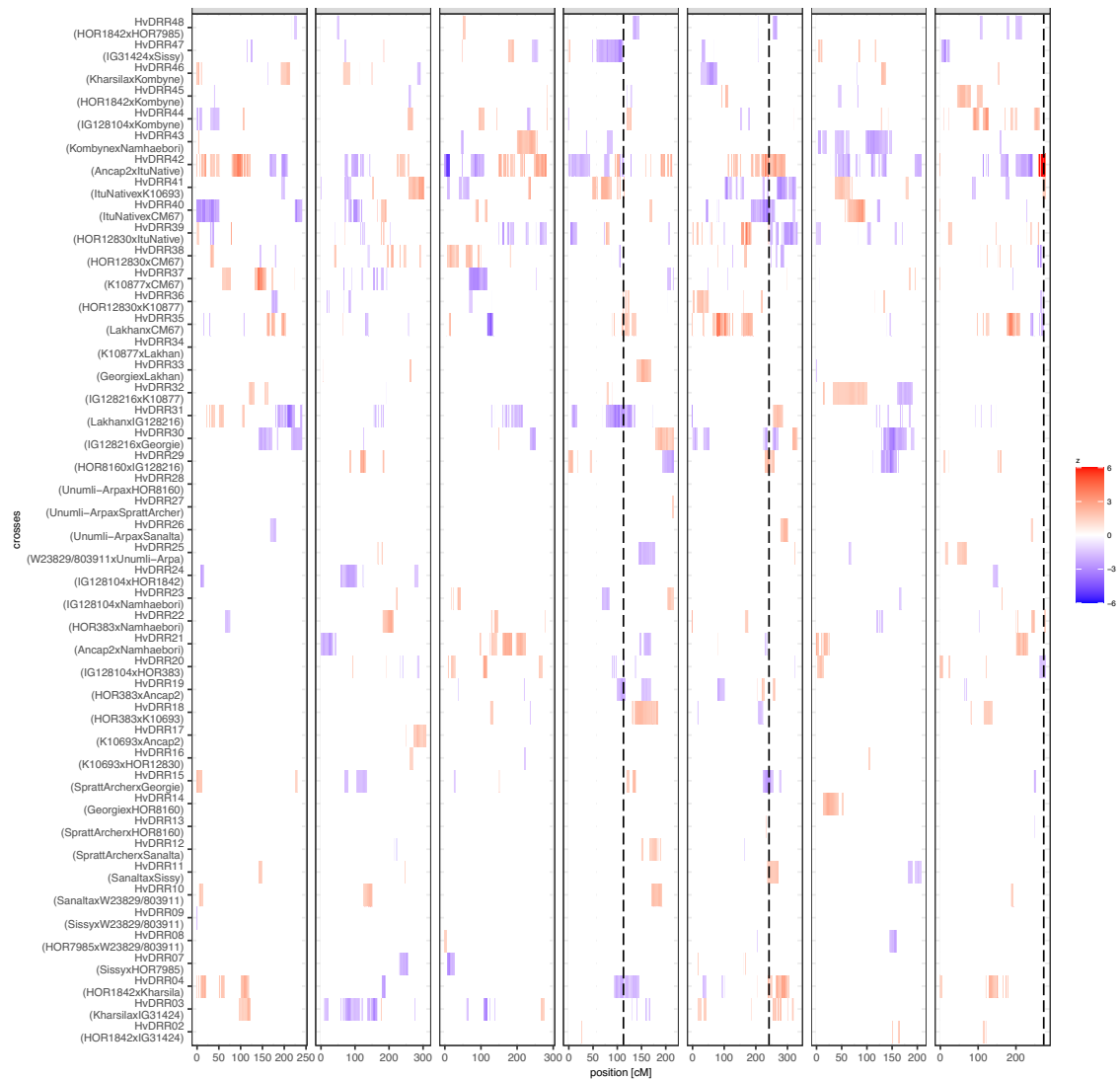


Fig. S9: The effect of the QTLs associated with the number of crossovers on chromosome 7H across the 45 double round-robin populations. The QTLs' locations are identified with a dashed black line.

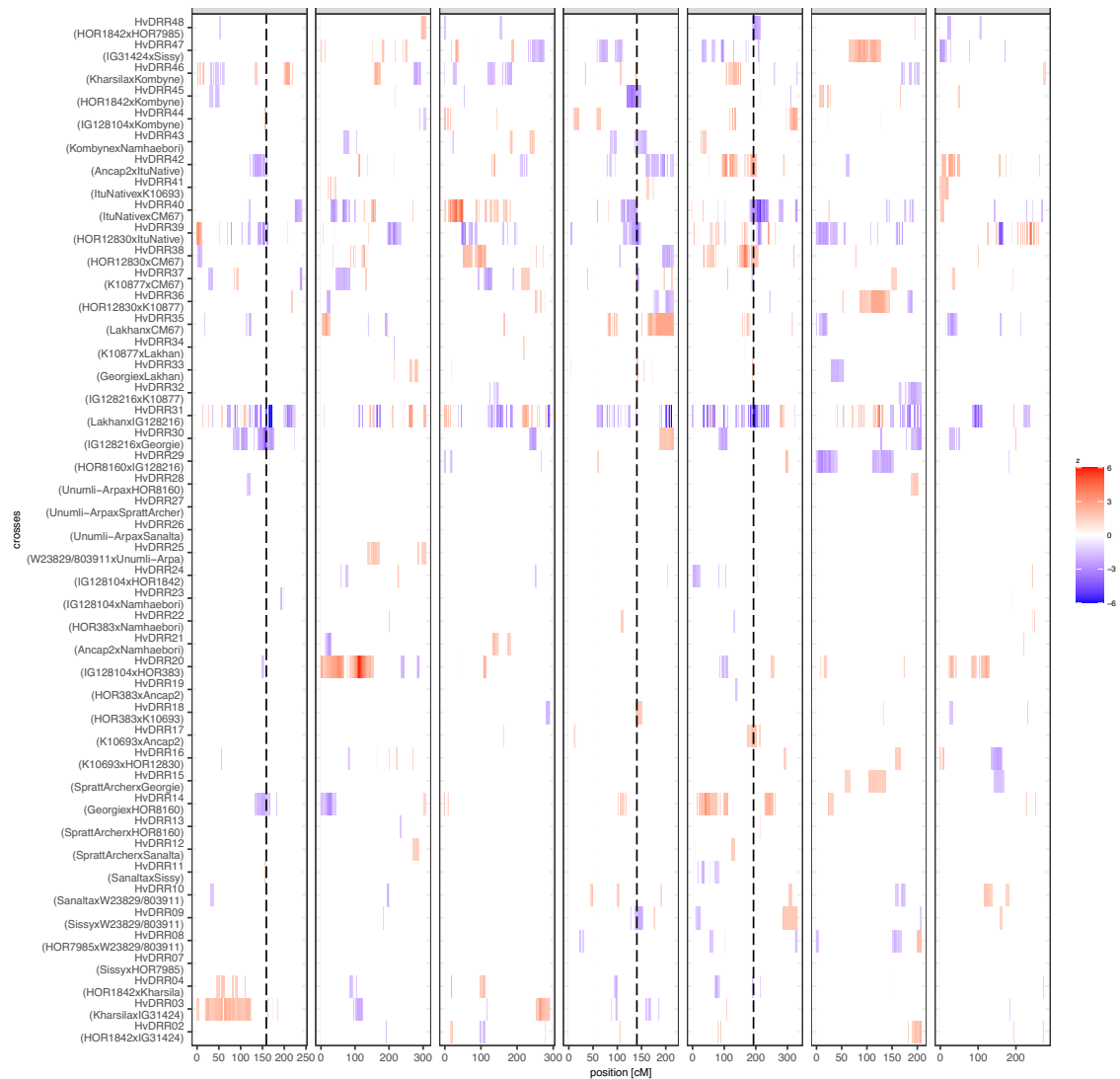


Fig. S10: The effect of the QTLs associated with the number of crossovers on the genome across the 45 double round-robin populations. The QTLs' locations are identified with a dashed black line.

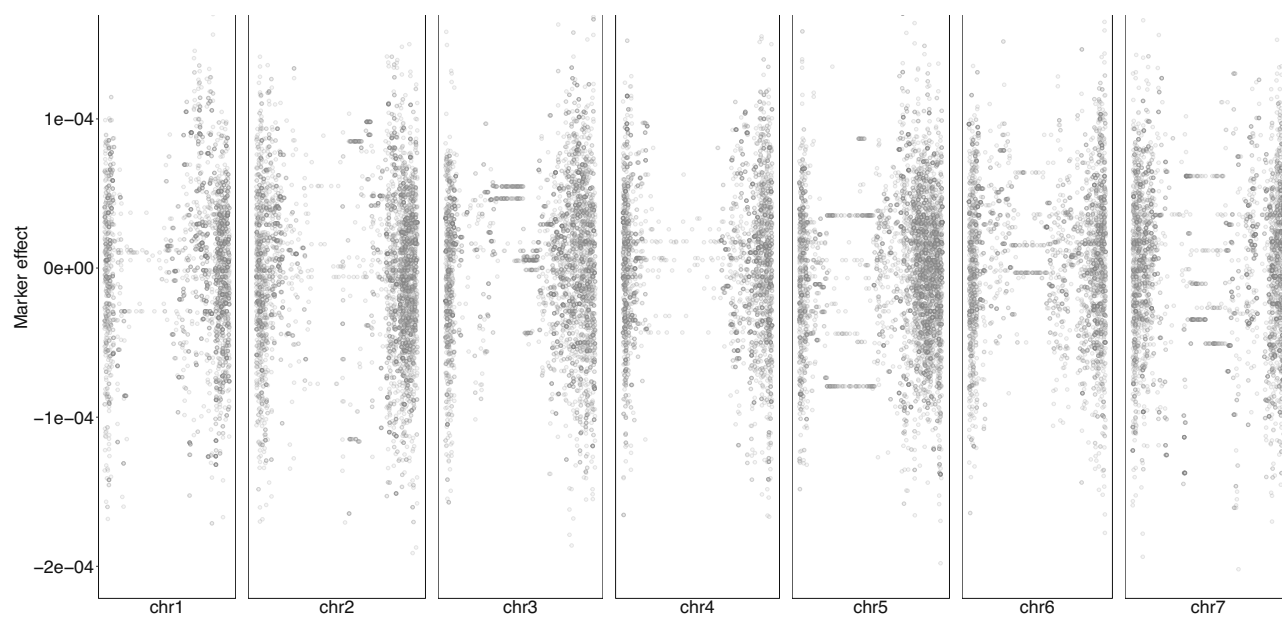


Fig. S11: The distribution of SNPs' effects predicted by RR-BLUP across the genome.

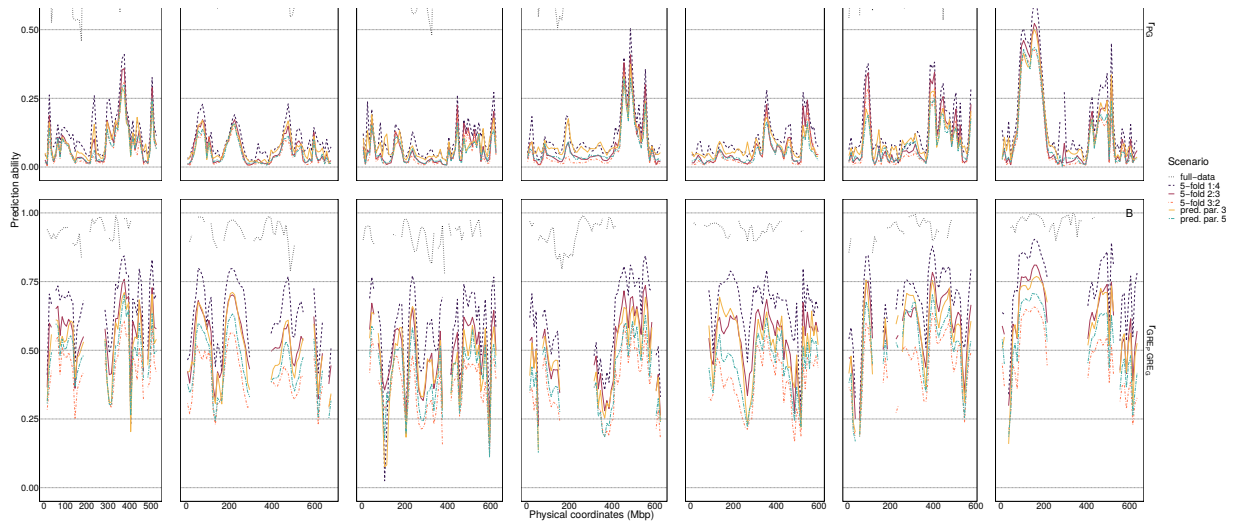


Fig. S12: The genomic prediction ability of recombination rate in 10 Mbp window level across the genome, using different cross-validation scenarios. (A) Pearson's correlation coefficient between the observed recombination rate and genomic estimated breeding values of the DRR populations, r_{PG} . (B) Pearson's correlation coefficient between the phenotypic and genomic estimated general recombination effects of the parental inbreds, $r_{GREpGRE_G}$. The cross-validation scenarios for genomic prediction are detailed in Figure 5.

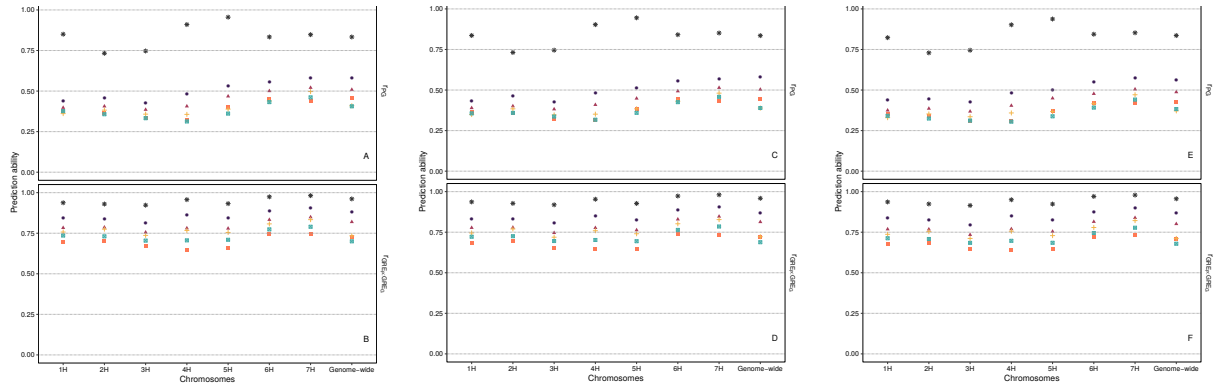


Fig. S13: Genomic prediction ability concerning the recombination rate variation of individual chromosomes and the genome-wide level, using different approaches and subsets of equally spaced SNPs. Genomic predictions in A and B were calculated using 1 SNP per 1 Mbp, in C and D using 1 SNP per 5 Mbp, and in E and F, using 1 SNP per 10 Mbp, respectively. (A, C, and E) Pearson's correlation coefficient between the observed recombination rate and genomic estimated breeding values of the DRR populations, r_{PG} . (B, D, and F) Pearson's correlation coefficient between the phenotypic and genomic estimated general recombination effects of the parental inbreds, $r_{(GRE_P GRE_G)}$. The cross-validation scenarios for genomic prediction are detailed in Figure 5.

4 The role of methylation and structural variants in shaping the recombination landscape of barley

This manuscript has been submitted for publication.

Authors:

Federico Casale, Christopher Arlt, Marius Köhl, Jinquan Li, Julia Engelhorn, Thomas Hartwig, and Benjamin Stich.

Own contribution: First author. I contributed to the design of the project, performed the data analyses, and wrote the manuscript.

The role of methylation and structural variants in shaping the recombination landscape of barley

Federico Casale¹, Christopher Arlt^{1,2}, Marius Kühl^{1,2}, Jinqian Li^{3,§}, Julia Engelhorn^{3,4}, Thomas Hartwig^{3,4}, and Benjamin Stich^{1,2,3,5*}

¹Institute of Quantitative Genetics and Genomics of Plants, Heinrich Heine University, Universitätsstraße 1, 40225 Düsseldorf, Germany

²Julius Kuehn Institute, Federal Research Center for Cultivated Plants, Institute for Breeding Research on Agricultural Crops, 18190 Sanitz, Germany

³Max Planck Institute for Plant Breeding Research, 50829 Köln, Germany

⁴Institute for Molecular Physiology, Heinrich Heine University, Universitätsstraße 1, 40225 Düsseldorf, Germany

⁵Cluster of Excellence on Plant Sciences, From Complex Traits toward Synthetic Modules, Universitätsstraße 1, 40225 Düsseldorf, Germany

*Corresponding author: Benjamin Stich; benjamin.stich@julius-kuehn.de

§Current address: Strube D&S GmbH, Hauptstrasse 1, 38387 Söllingen

July 31, 2024

ABSTRACT

Meiotic recombination is not only a key mechanism for sexual adaptation in eukaryotes but crucial for the accumulation of beneficial alleles in breeding populations. The effective manipulation of recombination requires, however, a better understanding of the mechanisms regulating the rate and distribution of recombination events in genomes. Here, we identified the genomic features that best explain the recombination variation among a diverse set of segregating populations of barley at a resolution of 1 Mbp and investigated how methylation and structural variants (SVs) determine crossover (CO) hotspots and coldspots at a high-resolution of 10 kb. Hotspots were found to be in proximity to genes and the genetic effects not assigned to methylation were found to be the most important factor explaining differences in recombination rates among populations along with the methylation and the parental sequence divergence. Interestingly, the inheritance of a highly-methylated genomic fragment from one parent only was enough to generate a coldspot, but both parents must be equally low methylated at a genomic segment to allow a hotspot. The parental sequence divergence was shown to have a sigmoidal correlation with recombination indicating an upper limit of mismatch among homologous chromosomes for CO formation. SVs were shown to suppress COs, and their type and size were not found to influence that effect. Methylation and SVs act jointly determining the location of coldspots in barley and the weight of their relative effect depends on the genomic region. Our findings suggest that

recombination in barley is highly predictable, occurring mostly in multiple short sections located in the proximity to genes and being modulated by local levels of methylation and SV load.

Keywords: recombination rate, hotspot, methylation, structural variant.

INTRODUCTION

In sexual reproduction, the recombination between the maternal and paternal homologous chromosomes followed by their random assortment during meiosis generates new combinations of alleles that can be transmitted to the next generation (Barton and Charlesworth 1998). This mechanism for generating genetic diversity is widely conserved among eukaryotes since it provides a possibility for the adaptation of species: the reshuffling of alleles breaks the linkage between beneficial and deleterious mutations, allowing the accumulation of beneficial mutations into one haplotype, and, ultimately, generating new phenotypes upon which selection can act (Muller 1932; Peck 1994).

Meiosis consists of a single round of DNA replication in which the bivalent is generated –a pair of physically linked homologous chromosomes each composed of two replicated sister chromatids– followed by two rounds of chromosome segregation (for a review see Mercier et al. 2015). During the first round, meiotic recombination occurs in prophase I when programmed DNA double-strand breaks (DSBs) are repaired as either crossovers (COs), the reciprocal exchange of large regions between homologous nonsister chromatids, or noncrossovers (NCOs), –the unidirectional copies of small fragments from any of the intact homologous chromatids (Szostak et al. 1983). In addition, when both CO and NCO occur, mismatches at the site of the strand invasion are produced if sequence polymorphisms exist among homologous chromosomes. Such mismatches are either restored to the original

allelic state (i.e., intersister repair) or repaired in favor of the homologous allele resulting in gene conversion (GC): the nonreciprocal exchange of alleles between homologous nonsister chromatids (Burt 2000; Wijnker et al. 2013).

The rate of recombination events is strongly regulated and has been proposed to be related to an optimal level of recombination for adaptation in eukaryotes (Mercier et al. 2015). The presence of at least one CO per bivalent, termed the obligate CO, is required for the correct segregation of homologous chromosomes (Hall 1972). In addition, the number of generated DSBs exceeds the number of observed COs which are rarely more than three per chromosome per meiosis in most species (Martini et al. 2006; Baudat and Massy 2007; Bauer et al. 2013). Most of the observed COs, furthermore, prevail in small regions of a few kilobases called recombination hotspots where CO rates are several times greater than the chromosome average (Mézard 2006; Choi and Henderson 2015). Remarkably, both the rate and distribution of COs in the genome have been shown to exhibit extensive inter- and intraspecific variation (Nachman 2002; Ritz et al. 2017; Lawrence et al. 2017). The mechanisms behind such variation, however, are not completely understood. Such knowledge is required for the effective manipulation of recombination, e.g., for the purpose of plant breeding. This recombination determines the frequency of breaking undesirable linkages and stacking favorable alleles in the genetic background of breeding populations and defines marker resolution to map quantitative traits (Bauer et al. 2013; Blary and Jenczewski 2019).

Earlier studies provided valuable information for deciphering the mechanisms

behind recombination variation in plants. Most of the recombination in plants occurs in euchromatic regions where chromatin is accessible while heterochromatin is suppressed for meiotic recombination (Henderson 2012). For example, earlier studies in *Arabidopsis thaliana* and other species have shown that recombination events tend to occur in genomic regions with hypomethylated DNA (Yelina et al. 2012; Rodgers-Melnick et al. 2015; Marand et al. 2019; Apuli et al. 2020; Fernandes et al. 2024) and depleted nucleosome density (Choi et al. 2013; Wijnker et al. 2013). Moreover, COs in plants are typically located in close proximity to genes and in association with chromatin marks that favor transcription (Choi et al. 2013; Mercier et al. 2015). A positive correlation between the recombination rate and gene density has been observed in many plant families (Paaape et al. 2012; Choi et al. 2013; Silva-Junior and Grattapaglia 2015; Gion et al. 2016; Wang et al. 2016; Apuli et al. 2020), including most grasses (Bauer et al. 2013; Darrier et al. 2017; Jordan et al. 2018; Gardiner et al. 2019; Marand et al. 2019; Casale et al. 2022). It is worth noting that some studies have instead reported a negative correlation between recombination rate and gene density (Kim et al. 2007; Giraut et al. 2011; Yang et al. 2012; Rodgers-Melnick et al. 2015).

Polymorphisms among homologous chromosomes are expected to prevent recombination due to defective strand invasion and homology pairing caused by the increase in mismatches among nonsister chromatids (Henderson 2012). Earlier studies in plants, however, reported contrasting correlations between recombination rate and parental sequence divergence (Saintenac et al. 2011; Salomé et al. 2012;

Yang et al. 2012; Bauer et al. 2013; Jordan et al. 2018; Marand et al. 2019). Recent reports in *Arabidopsis* suggested that the recombination rate has a positive correlation with parental allelic divergence until a level of mismatch among homologous chromosomes prevents CO formation (Blackwell et al. 2020; Hsu et al. 2022). Accordingly, large structural variants (SVs) were shown to suppress local recombination in several plant species (Rodgers-Melnick et al. 2015; Shen et al. 2019; Rowan et al. 2019; Fernandes et al. 2024).

The above-mentioned contrasting findings impose the necessity to characterize the associations between recombination and genomic features at the species level to avoid making incorrect assumptions. In addition, due to differences in the employed research methods, disagreements were observed among studies on the same species (Apuli et al. 2020). In this respect, most reported recombination rates in plants were calculated based on a coarse genomic resolution that failed to capture the complete genetic variation generated from meiosis. At present, most reported recombination assessments at high-resolution have been performed in *Arabidopsis* (Sun et al. 2012; Lu et al. 2012; Yang et al. 2012; Wijnker et al. 2013; Rowan et al. 2019; Fernandes et al. 2024), while only a few have been performed in major cereal crops such as maize (*Zea mays*) (Rodgers-Melnick et al. 2015; Li et al. 2015), rice (*Oryza sativa*) (Si et al. 2015; Marand et al. 2019), and wheat (*Triticum aestivum*) (Jordan et al. 2018). In barley (*Hordeum vulgare*), the fourth of this list, earlier low-resolution recombination studies successfully revealed the genetic basis of recombination as well as the association of recombination with some environmental and genomic

features on a broad genomic scale (Higgins et al. 2012; Dreissig et al. 2019, 2020; Casale et al. 2022); however, a high-resolution study depicting the complete meiotic recombination landscape and the respective associations of genetic and epigenetic features in the barley genome is still lacking.

Consequently, in the present study, we aimed to (*i*) identify the genomic features that best explain the recombination variation among the double round-robin (DRR) populations, (*ii*) detect recombination events in the barley genome at high-resolution, and (*iii*) analyze the effect of genomic features in determining the location of recombination hotspots and coldspots in the genome.

RESULTS

The genomic features associated with recombination rate variation in the barley genome

The recombination rate was almost null in the pericentromeric region, but increased toward the distal regions of the chromosomes (Figure 1A). The SV load, the physical fraction spanned by the analyzed SVs (insertions, deletions, inversions, duplications, and translocations) increased toward the distal regions of the chromosomes as did the recombination rate and gene density (Figures 1A and S1). In this way, the presence of SVs and the sequence divergence among parents showed a significant positive Pearson's correlation ($P < 0.05$) of 0.35 with the recombination rate along the distal regions of the chromosomes (Figures 1C and S3A), whereas a lower but significant correlation ($P < 0.05$) of 0.1 was observed in the pericentromeric region (Figures 1D and S3B). The sequence divergence among the parental inbred lines, which showed a significant positive correlation ($P < 0.05$) with the SV load of 0.27, was found to have a significant positive correlation ($P < 0.05$) with recombination rates of only 0.1 and 0.14 in the distal and pericentromeric regions, respectively (Figures 1A, 1C, and 1D). The methylation level in the sequence contexts differed along the barley chromosomes, where the level in the CpG and CHG contexts reached a maximum in the pericentromeric region and decreased toward the telomeres, while the methylation level in the CHH context increased toward the telomeres (Figure 1B). Because the CpG and CHG contexts had higher overall

methylation levels than did the CHH context, the average methylation level along the chromosomes mostly represented the trend observed for the CpG and CHG contexts. The observed significant negative correlation ($P < 0.05$) between the average methylation level and recombination rate along the barley chromosomes was therefore due to the CpG and CHG contexts but not to the CHH context (Figures 1C and 1D). The difference in the methylation level between the parental inbred lines of any of the analyzed populations was greater in the distal regions than in the pericentromeric regions of the chromosomes in all three analyzed sequence contexts (Figure S4). This explained the positive correlation ($P < 0.05$) of 0.17 of such difference and the recombination rate along the distal regions in the barley chromosomes (Figures 1C, 1D, S3A, and S3B).

***The genomic features associated with recombination rate
variation among barley populations***

The best subset of genomic features explaining the recombination rate variation among the 45 DRR populations in 1 Mbp genomic windows along the barley chromosomes was identified using a stepwise regression model (Figure 2). The fraction of 1 Mbp windows of the barley genome in which a given genomic feature was found to be significantly correlated and the direction of such correlation are given in parentheses below. The genetic effects, which were calculated from the GRE, were found to be the most determining factor in explaining differences in recombination rate among populations with a promoting effect on both the distal (positive, 0.76) and

the pericentromeric regions (positive, 0.40) of the chromosomes. The parental sequence divergence was also found to be positively correlated with the recombination rate in both mentioned chromosomal regions (positive, 0.24 and 0.19, respectively). A few windows showed a significant negative correlation between parental sequence divergence and the recombination rate (negative, 0.04 and 0.01 in both respective chromosomal regions). Notably, the windows with a negative correlation had a parental sequence divergence near its relative maximum, which appeared to be associated with a high SV load (Figure S5A). The average methylation level across the different sequence contexts was negatively correlated with the recombination rate in the distal (negative, 0.43) and pericentromeric (negative, 0.12) regions of the barley chromosomes. This means that populations with a higher methylation level than others in a particular window showed a lower recombination rate than others in that window, and vice-versa. Additionally, the multiple regression model used to calculate genetic effects revealed a correlation between the methylation level and the sum of the GREs ranging from -0.29 to 0.03 per 1 Mbp genomic window with an average of -0.06 in the distal region. The difference in the average methylation level between the parental inbred lines of the respective populations was found to have only a low impact on the recombination rate variation among populations (in distal regions, positive, 0.04; negative, 0.03). The physical fraction of 1 Mbp genomic windows spanned by all SVs was found to have a low (and mostly positive) impact on the differences in recombination rate among populations (in distal regions, positive, 0.05; negative, 0.02).

The methylation level in the three sequence contexts –CpG, CHG, and CHH– analyzed independently in an extended model was shown to have the same repressive effect on recombination as the average across the three sequence contexts (Figure S6 and Table S2). In addition, the parental methylation difference for the sequence contexts CHG and CHH was found to be positively associated with the recombination rate, and the inverse was found for the context CpG. Furthermore, the extended model revealed no impact of the different SV types on recombination rate variation among populations.

***The key parameters for detecting recombination events at
high-resolution***

The mRNA sequencing of recombinant inbred lines (RILs) of the evaluated populations yielded 4.1 M sequence variants of which, 858 K SNPs remained after being intersected with reported SNV parental data generated by DNA sequencing. A total of 12 K SNPs were added from the iselect array, resulting in a total of approximately 870 K SNPs genome-wide. After removing SNPs carrying nonparental alleles or with 100% missing data, the final number of genome-wide SNPs ranged from 214 to 259 K per population (Table S3 and Figure S7). The median inter-SNP distance was 132 bp on average across populations, which was 79 times shorter than the 10,475 bp utilized to count COs for the same three populations analyzed in a previous study (Casale et al. 2022). The maximum inter-SNP distance ranged from 5.45 to 11.99 Mbp among populations (Table S3), denoting large regions that

were identical by descent (IBD) among the parental inbred lines involved in a given population and thus among their respective offspring. This resulted in an average density of 57 SNPs per Mbp across populations. The mean length for each parental SNP block category was 14.8, 44.1, and 111.5 Mbp for the short, medium, and long CO-related blocks, respectively (Table S4). The block length was positively correlated with the number of SNPs per block (> 0.75 , $P < 0.001$). Therefore, the false positive rate for detecting CO was inversely proportional to the marker block length, supporting the identification of block length categories with different CO layers.

On average, there were 30, 87, and 269 genome-wide COs accumulated per RIL across populations for the long, medium, and short block lengths, respectively (Table S5). Considering only the layer of COs generated by blocks longer than 3 Mbp, the genome-wide CO counts per RIL ranged from 14 to 65 across populations (Figures S8 and S9). Since a given CO breakpoint was determined as the midpoint of the CO interval, the breakpoint location accuracy depended on the CO interval length (Table S6). The lengths of the detected CO intervals ranged from less than 20 bp to 10.8 Mbp with a median that varied from 41.9 kb to 151.6 kb depending on the considered CO layer. The average number of genome-wide GC events that accumulated per RIL across populations was 58, 251, and 6,521 for long, medium, and short GC-related block lengths, respectively.

An average of 80 recombination hotspot windows per chromosome were found across the three selected populations (Table S7). Among these, 12 were found

in the pericentromeric regions, and the rest were found in the distal regions of the barley chromosomes (Table S8 and Figure S10). The recombination hotspot windows contained 0.26, 0.14, and 0.25 of the total observed COs in the HvDRR13, HvDRR27, and HvDRR28 populations, respectively. Less than 10% of the hotspot windows in a given population were shared with another population and less than 1% of the total counted hotspot windows were shared among the three analyzed populations (Figure S11). Interestingly, both the number and conservation level of coldspot windows far exceeded those of hotspots. On average, across populations, more than half (10,436 out of the 19,496) of the distal windows were recombination coldspots. More than 60% of the coldspot windows in a given population were shared with the other populations, and 16.7% of the coldspot windows were present in all three analyzed populations. The majority of the coldspot windows were located contiguously in regions with lengths that varied from 10 kb to 17 Mbp with an average of 322 kb across the three analyzed populations.

More than 15% of the GC hotspot windows in a given population were shared among the three analyzed populations (Figure S12). The GC hotspot windows were found to overlap with the CO hotspot windows in the distal region of the barley chromosome significantly more than they did under a random distribution across the three analyzed populations (Table S9). The GC hotspot windows detected in a given population overlapped with 12–15% of the CO hotspot windows in the same population and with 7.5–9.5% of the CO hotspot windows detected in the other two analyzed populations (Table S10). There was no significant ($P > 0.05$) difference

between such overlap proportions in the HvDRR27 and HvDRR28 populations.

***The effect of methylation and structural variants on
recombination rate variation at high-resolution***

The coldspot and hotspot windows have different methylation level and SV load than the rest of the genome

The coldspot and hotspot windows in a given chromosome region showed distinct methylation patterns compared to the average remaining windows in the same region in the three analyzed populations (Figure 3). The average methylation level across all three methylated sequence contexts in the coldspot windows of the distal telomeric subregion was significantly greater ($P < 0.001$) than that across the other windows in both distal subregions. In contrast, the coldspot windows in the distal proximal region were not found to be differentially methylated ($P > 0.001$) from other windows in any of the distal subregions. However, when analyzing the CpG and CHG sequence contexts separately, the methylation level in the coldspot windows of the distal proximal subregion was found to be significantly greater ($P < 0.001$) than that in the other windows of both distal subregions. Differently, the methylation level in the coldspot windows of the distal telomeric subregion was significantly greter ($P < 0.001$) than that in the other windows in this subregion but significantly lower ($P < 0.001$) than that in the windows of the distal proximal subregion. The coldspot windows in such comparisons that were below the critical value (methylation levels of 0.89 and 0.59 for the sequence contexts CpG and CHG,

respectively) were found to have a significantly ($P < 0.001$) greater total SV load fraction than the coldspot windows above the critical value (Table S11).

The average methylation level across the three sequence contexts in the hotspot windows was significantly lower ($P < 0.001$) than that across the other windows in any region of the barley chromosomes. The hotspot windows in the pericentromeric region, however, were not found to be differentially methylated from the rest of the windows in such region or from the windows in other regions of the genome, including coldspots. However, by analyzing the methylated sequence contexts separately, the hotspot windows were found to be significantly less methylated ($P < 0.001$) in the CpG and CHG sequence contexts than in the total windows in the pericentromeric region.

The coldspot windows in the distal telomeric regions were found to have a significantly greater total SV load ($P < 0.001$) than the rest of the windows in that subregion. In contrast, the coldspot windows in the distal proximal region did not show such an increase in total SV load. However, the coldspot windows below the critical value (SV loads equal to 0.187, 0.174, and 0.187 for the HvDRR13, HvDRR27, and HvDRR28 populations, respectively) in such comparisons had a significantly increased methylation level ($P < 0.05$) compared to the windows above the critical value for the CpG and CHG sequence contexts (Table S12). The hotspot windows were not observed to have a significantly different ($P > 0.001$) span of total SVs compared to the total windows in their respective chromosome regions. However, the observed overlaps between CO intervals and insertions/deletions and

duplications were found to be significantly less frequent ($P < 0.001$) than such overlaps under a random distribution of the COs and the respective SVs in the distal regions of the barley chromosomes in the three analyzed populations (Table S13). In the case of inversions, such a pattern was observed only for the HvDRR27 population. Moreover, the distance between the CO breakpoints and the closest SV of any type was significantly greater ($P < 0.001$) than the CO-SV distances expected by chance (Table S14).

The genomic environment neighboring hotspot and coldspot windows

The 10 kb genomic windows adjacent to coldspot regions were found to have a significantly lower ($P < 0.001$) average methylation level across the three sequence contexts than coldspots in both distal subregions of the chromosome in the three analyzed populations (Figure 4). This observation reflected the pattern produced at the methylated sequence contexts CpG and CHG, individually (Figure S13). In addition, any 10 kb window in the considered range from -40 kb to +40 kb around coldspot regions was found to have a significantly lower ($P < 0.001$) total SV load than the neighboring coldspot. The 10 kb genomic windows adjacent to hotspot regions were not found to have significantly ($P > 0.001$) different methylation levels or SV loads than any of the analyzed chromosomal regions or populations.

The windows neighboring coldspot regions were found to have a significantly lower ($P < 0.001$) gene density than these regions, except for the windows located 20 kb upstream of coldspots. However, the overlap between the coldspot regions

and genes in the distal regions of the barley chromosomes was not significantly different ($P > 0.001$) from such overlap under a random distribution (Table 1). In contrast, a visual increase in the gene density from the hotspots to 20 kb upstream was observed in all the analyzed genomic regions, although this pattern was not significant ($P > 0.001$). Furthermore, the overlap between the hotspot regions and genes was found to be significantly ($P > 0.001$) greater than expected under a random distribution in the three analyzed populations, while the overlap between hotspots and intergenic regions was not found to be significantly ($P < 0.001$) greater than random, with the exception of the HvDRR27 population. In addition, a high proportion of the windows surrounding hotspot regions in both the proximal (0.37–0.49) and telomeric (0.33–0.45) subregions of the distal region of the genome were coldspot windows (Table S15).

The variation in methylation and SVs in coldspot and hotspot windows among barley populations

Significant differences ($P < 0.016$) were observed in the methylation levels of the three analyzed populations, either by analyzing the methylated sequence contexts separately or by analyzing their average (Figure 3). Such differences among populations observed for the total windows were also detected in the coldspot windows in all of the analyzed chromosome regions. In contrast, the methylation level in the hotspot windows was found to be equal ($P > 0.016$) among populations in any of the analyzed chromosome regions, either by analyzing the methylated sequence

contexts separately or their average. A similar trend was observed for the total SV load fraction: while observing significant differences ($P < 0.016$) among populations in the total windows but also in the coldspots of both the pericentromeric region and the distal subregions, such differences were not observed for the hotspot windows.

The methylation level of the two parental inbred lines of each of the analyzed populations differed significantly ($P < 0.008$) at the CpG and CHG sequence contexts of the genomic windows identified as coldspots in their respective offspring (Table 2). Thus, the increased methylation of only one of the parental genotypes at a genomic region might be enough to generate a coldspot in the offspring. In hotspot windows, no significant differences ($P > 0.008$) between parental inbred lines were found in the methylation level at any of the three analyzed sequence contexts, indicating that parents must have equally low methylation at a genomic segment to allow a recombination hotspot.

DISCUSSION

Detection of recombination events at high-resolution in barley

The substantial decrease of the median inter-SNP distance compared to a previous study with the same three populations (Casale et al., 2022) produced a slight increase of the CO discovery rate of 0.31 times when considering the COs related to > 3 Mbp marker blocks. However, such increase jumps to 2.74 and 10.59 times, if considering the COs related to > 500 kb and > 10 kb blocks, respectively (Tables S3 and S5). The assumed part of the observed increase due to the additional recombination that occurred at heterozygous regions in the selfing generation analyzed in the first study is expected to be small due to the decreasing remaining heterozygosity after every selfing generation that produces fewer new observable COs per generation during inbreeding (Esch et al. 2007). By analyzing the detected CO rate between comparable low and high-resolution analyses reported in previous studies, only small differences were detected in populations of *Arabidopsis thaliana* and maize (*Zea mays*) (McMullen et al. 2009; Rodgers-Melnick et al. 2015), but substantial differences were reported in populations of wheat (Esch et al. 2007; Gutierrez-Gonzalez et al. 2019; Gardiner et al. 2019) and *Populus* (Apuli et al. 2020). In addition to the different utilized resolutions, other reasons behind observing differences in the recombination rate in the same population may include the genotyping error rate, the data filtering criteria, and the size of considered CO events.

The observed CO rate per RIL per chromosome per generation in the present study, when considering the COs related to > 3 Mbp marker blocks, is in line with high-resolution studies in *Arabidopsis* (Sun et al. 2012; Lu et al. 2012; Yang et al. 2012; Wijnker et al. 2013; Qi et al. 2014; Rowan et al. 2019) and rice (*Oryza sativa*) (Si et al. 2015) that reported rates of 1.5–2.2 COs per chromosome per generation, and with another high-resolution study on wheat (Gardiner et al. 2019) when looking at the COs related to > 500 kb marker blocks.

The COs per RIL per chromosome per generation observed in our study when considering all COs related to blocks (10 kb) should be compared with values observed by Yang et al. (2012) in *Arabidopsis* and Gardiner et al. (2019) in wheat. In such comparisons, however, considering every marker block shorter than 10 kb as a GC is an arbitrary threshold. This has the potential to cause misclassification between COs and GCs among the categories of COs related to blocks > 10 kb and those related to GCs between 2 and 10 kb.

The present study is the first to characterize GC events in barley, along with a few in other crop species (Li et al. 2015; Si et al. 2015; Gardiner et al. 2019). The poor documentation of GCs in plants beyond studies in *Arabidopsis* is because phenotypic screens can barely detect GC events. In addition, the detection of GC events at the molecular genetic level is also challenging because of their short length (Mancera et al. 2009; Mercier et al. 2015). This makes GC detection very sensitive to the marker density and GC rate, which also depend on the recombination rate, the tract length of the repair intermediates where GCs occur, and the polymorphism

density (Wijnker et al. 2013). Moreover, in most flowering plants, gametes do not remain grouped after meiosis, making it difficult to observe the expected 3:1 allelic proportion of GCs (Sun et al. 2012). In the present study, the average number of genome-wide GC events per RIL across populations was 58, 251, and 6,521 for long (2–10 kb), medium (20 bp–2 kb), and short (2–20 bp) GC-related marker block lengths, respectively. The SNPs analyzed per population could be translated to 0.00003, 0.00014, and 0.0039 GC events per site per RIL per generation for the different types of GC-related marker block lengths. Marker blocks shorter than 20 bp are expected to contain a high number of false-positive GCs since they were predominantly called based on two markers only. Therefore, if considering the GC-related marker blocks of long (2–10 kb) and medium (20 bp–2 kb) length only, the detected 0.00017 GCs per site per RIL per generation is on the same order of magnitude as that observed in studies using a similar approach for GC detection (Yang et al. 2012; Gardiner et al. 2019). However, the observed GC rate in our study was two orders of magnitude greater than that reported in tetrad analysis studies performed in *Arabidopsis* (Lu et al. 2012; Sun et al. 2012; Wijnker et al. 2013) and sequencing of rice F2 populations (Si et al. 2015). This disagreement might be explained not only by the less precise GC detection method used in our study but also by the occurrence of nonallelic sequence alignments caused by SVs inflating the number of false-positive gene conversions (Qi et al. 2014; Si et al. 2015). It is also worth noting that the reported GC rate is the frequency of GCs generated from NCO and CO events combined. In the present study, we did not attempt

to estimate the rate of NCO in barley because these NCOs are only traceable after gamete formation when they lead to GC, and it was not possible to precisely differentiate between CO and NCO conversion tracts with the employed marker resolution. In addition, the DSB rate in barley is not known enough to estimate the NCO rate from the detected CO events.

The present study is the first to report genome-wide recombination hotspots at high-resolution in barley. On average, across the three investigated populations, of the 80 hotspots per chromosome, only 12 were found in the pericentromeric region, and the rest were found in the distal regions of the barley chromosomes (Figure S10 and Table S8). In addition, while the three investigated populations always shared one parent, the proportion of shared hotspots between two and three populations was 10% and 1%, respectively, of the total hotspots detected in a given population (Figure S11A). This observation is in good agreement with previous studies in *Arabidopsis* showing that recombination hotspots were cross-specific (Salomé et al. 2012).

In contrast, in the case of the GC hotspot windows, the overlap among the three populations was more than 15% (Figure S12). Moreover, GC hotspots were found to have high overlap not only with GC hotspots of the same population but also with windows that are hotspots in other populations (Table S10). Thus, GC hotspots might be considered fingerprints of population-specific silenced COs that result in NCO in regions with high DSB rates in the genome of the species. This observation suggested that although the CO rate and distribution present extensive intraspecies

variation, such DSBs might be highly conserved within the species. However, this requires further research.

Additionally, in barley, recombination hotspots alternate in the genome with coldspots. For example, in domains where CO rates are significantly lower than the genome average (Figure S10), as observed in previous studies in other species (Mercier et al. 2015). Indeed, by dividing the genome into 10 kb genomic windows, hotspot windows were found to be adjacent to coldspot windows in 42.5% of the cases in the distal regions of the barley chromosomes (Table S15). This continuous intermittence in the recombination rate might explain the above-mentioned large difference in CO events found between the high- and low-resolution analyses on these barley populations.

To avoid calling recombination coldspots in the pericentromeric and telomeric regions of the chromosomes, which are long regions depleted from recombination as seen in previous studies (Boideau et al. 2022), in the present study, coldspots were identified only in the distal region of the chromosomes by employing a long physical distance margin between regions. The majority of the detected 10 kb coldspot windows were located in coldspot regions with an average length of 322 kb (Table S7). Each population shared 60% of its coldspot windows and more than 16% with the other two populations, thus demonstrating a greater conservation of coldspots than hotspots in barley (Figure S11B). Such differential conservation between recombination hotspots and coldspots might be linked to the different genomic features determining their occurrence.

The genomic features associated with recombination rate variation in barley

The present study is the first comprehensive evaluation of the genomic features associated with differences in recombination rates among barley populations. On a scale of 1 Mbp windows, the recombination rate was found to be positively correlated with sequence divergence among parental inbred lines, gene density, and SV load on the barley chromosomes and was negatively correlated with the methylation level (Figure 1).

The results of the present study are in line with earlier studies in plants in which recombination was found to be positively associated with genetic divergence among homologous chromosomes (Yang et al. 2012; Marand et al. 2019; Blackwell et al. 2020). Although a negative association was reported in some studies (Saintenac et al. 2011; Gion et al. 2016; Bouchet et al. 2017; Serra et al. 2018; Jordan et al. 2018; Gutierrez-Gonzalez et al. 2019), the contradiction might be explained by a sigmoid relationship between both variables, meaning that recombination has a positive correlation with genetic divergence until a level after which the high polymorphism among homologs suppresses COs due to the increase in mismatches, as recently reported in *Arabidopsis* (Blackwell et al. 2020; Hsu et al. 2022). In this respect, the few observed 1 Mbp windows with a significant negative association were found to have parental sequence divergence at the relative maximum level, which appeared to be associated with an extensive SV load (Figure S5A and B).

The SV load was not identified as a determining factor for the differences in the recombination rate among populations, presumably because the employed resolution of 1 Mbp was too coarse to detect differences among populations, as most of the analyzed SVs were smaller (Figure 2). Additionally, the positive correlation between the recombination rate and the SV load on a broad scale might be explained by the accumulation of DNA repair errors in highly recombining regions throughout the evolutionary history of barley (Figures 1A and 1C). Genomic regions with a high rate of DSBs are expected to have a historically increased mutation rate produced by COs and GCs, which elevates the allelic diversity at such regions among genotypes as demonstrated in humans (Arbeithuber et al. 2015; Halldorsson et al. 2019).

In addition to the positive correlation shown between recombination and gene density on a broad scale, in the present study, 10 kb hotspot windows were found to be associated with regions of high gene density (Table 1), as repeatedly reported in previous studies in grasses (Rodgers-Melnick et al. 2015; Darrier et al. 2017; Bouchet et al. 2017; Jordan et al. 2018; Gardiner et al. 2019; Marand et al. 2019; Casale et al. 2022), and other plant families (Paape et al. 2012; Choi et al. 2013; Silva-Junior and Grattapaglia 2015; Gion et al. 2016; Wang et al. 2016; Apuli et al. 2020). Furthermore, the hotspot windows were located in proximity (< 20 kb apart) but did not overlap with the genes (Figure 4). This finding in barley is in line with previous observations in *Arabidopsis*, maize, and rice showing an increased CO frequency toward gene promoters and terminators (Choi et al. 2013; Wijnker

et al. 2013; Li et al. 2015; Marand et al. 2019; Sun et al. 2019), similar to that observed in budding yeast (*Saccharomyces cerevisiae*) (Pan et al. 2011).

In the present study, the genetic effects were calculated as the proportion of the sum of the *GRE* of both parents for a given population that was not explained by methylation, assuming parental sequence divergence and SV load as part of the *SRE* (Casale et al. 2022). Such genetic effects were shown to be the factor explaining the greatest proportion of differences in recombination rates among barley populations. Here, we hypothesize that such genetic effects are the product of the expression of genes related to the recombination machinery being in part modulated by the methylation level, explaining a portion of the uneven distribution of the recombination hotspots along the barley chromosomes. The wide distribution of the observed genetic effects along the chromosomes is in line with previous studies in wheat (Jordan et al. 2018) and barley (Casale et al. 2022) reporting that differences in the genome recombination rate among populations are explained by multiple loci with small effects.

A negative correlation was observed on a broad scale between recombination rate and the extent of methylation in the CpG and CHG sequence contexts (Figure 1B). This is in accordance with previous studies in other plant species showing that COs occurred in euchromatic regions while heterochromatic regions were depleted of COs and that hypomethylation at CpG sites increased the genome-wide CO rate (Melamed-Bessudo and Levy 2012; Colomé-Tatché et al. 2012; Salomé et al. 2012; Yelina et al. 2012; Mirouze et al. 2012; Wijnker et al. 2013; Rodgers-Melnick

et al. 2015; Marand et al. 2019; Apuli et al. 2020; Boideau et al. 2022; Fernandes et al. 2024). This association was confirmed at high-resolution by observing the relationship between 10 kb coldspot windows and increased methylation in the CpG and CHG sequence contexts (Figure 3). Furthermore, compared with those in the nonhotspot windows, the methylation levels in the CpG and CHG sequence contexts in both the distal and pericentromeric regions of the barley chromosomes decreased in the hotspot windows. This result is in line with previous findings in maize showing a strong relationship between the occurrence of hotspots and decreased CpG and CHG methylation but no association with CHH methylation (Rodgers-Melnick et al. 2015). In this respect, it was suggested that increased recombination was associated with increased CHH methylation in regions with high CpG-related methylation levels but with decreased recombination where CpG methylation was low (Rodgers-Melnick et al. 2015).

The presence of SVs was shown to suppress COs in previous studies in *Arabidopsis* (Rowan et al. 2019; Fernandes et al. 2024) and other plant species (Rodgers-Melnick et al. 2015; Shen et al. 2019). In the present study, we were able to detect a decreased overlap and a longer distance between CO breakpoints and SVs compared to a random distribution, thus indicating the negative association of SVs with the occurrence of COs in barley (Tables S13 and S14). The type and size of the SVs were not found to be related to such effects, which is in line with previous findings in *Arabidopsis* (Rowan et al. 2019). Moreover, this is the first study showing the joint effect of methylation and the accumulation of SVs in determining genomic

regions deprived of recombination outside the pericentromeric region (Fernandes et al. 2024) and the variation in this effect within the genomic region (Figure 3 and Tables S11 and S12). In the distal telomeric region of the barley chromosomes, most 10 kb coldspot windows were found to be associated with increases in both the recombination level and the SV load. However, in the distal proximal region, increased methylation was found to be associated with most of the 10 kb coldspot windows, but an increased SV load was found to be associated with coldspot windows with no increased methylation. Interestingly, the effects of both methylation and SVs on the occurrence of coldspots were noticeable not only when comparing coldspot windows with other windows located far away in the same genomic region but also when comparing coldspots with their neighboring windows. This indicated a marked local effect of methylation and SVs on recombination suppression (Figure 4). In addition, the differences in both methylation level and SV load among barley populations were found to be responsible for the differences in the localization of coldspot windows among such populations. The parental inbred lines of the analyzed populations were found to differ in the methylation level in the genomic windows identified as coldspots in their offspring populations (Table 2), indicating that the inheritance of high methylation from only one parent was sufficient to prevent recombination in a particular region of the genome.

In a previous study on the same barley populations in which methylation was not separated from the genetic effects of genotypes, the effect of individual parental inbred lines was shown to be the major determinant of the recombination rate of

the respective biparental offspring populations (Casale et al. 2022). The increased methylation at genomic regions leading to coldspots might be an important part of the genetic effect of the parents on the recombination rate, which was negatively correlated with methylation in the present study.

In contrast to the above described association of methylation and SVs with the occurrence of recombination coldspots, no significant differences between parental inbred lines were found in the methylation level of hotspot windows, indicating that parents must have equally low methylation at a genomic segment to allow a recombination hotspot. The employed window resolution of 10 kb, which is longer than the length reported earlier for recombination hotspots in other plant species (Choi and Henderson 2015), might be the reason for the lack of detection of SV effects on hotspots and the lack of differences in methylation between hotspots and their neighboring windows.

Our findings demonstrate that the recombination landscape in barley is highly predictable. Most of the recombination occurs in multiple short highly recombining sections in the distal regions of the chromosomes. These recombination hotspots are located in proximity to genes and where the levels of methylation and SV load are low enough to allow CO concretion. In this sense, such hotspots alternate with long regions deprived of recombination because of increased methylation or the accumulation of SVs preventing CO from occurring. Therefore, local differences in the recombination rate among barley populations can be explained to a considerable extent by differences in the methylation level and the accumulation of SVs at

multiple locations within the genome. Such differences are highly inheritable and can be determined by the effect of only one parent in a cross. However, our analyses suggest that in addition to these two genomic features, additional differences in the recombination machinery must exist, which forms the basis for what we designated genetic effect.

METHODS

Identification of the genomic features associated with recombination rate variation among barley populations

The recombination rates of 45 biparental barley populations, referred to as double round-robin populations, were obtained from Casale et al. (2022). These have been derived from genotyping the populations with the 50K Illumina Infinium iSelect SNP genotyping array (Bayer et al. 2017). The recombination rates were recalculated based on the Morex v3 reference genome sequence (Mascher et al. 2021) at 1 Mbp genomic windows. The pericentromeric region of each chromosome was defined as the continuous region surrounding the centromere for which the average recombination rate across the 45 DRR populations was 5-fold lower than the respective chromosome average across populations in 1 Mbp genomic windows. The regions of the chromosome that did not belong to the pericentromeric region were designated in the following as distal regions.

Whole-genome bisulfite DNA sequencing data for the 23 DRR parental inbred lines was obtained by extracting DNA from inbred lines from a mix of tissues, including the whole seedling plant, the leaf, and the apex, at stage 47 on the Zadoks scale. DNA library preparation was performed with NEBNext® Ultra™ II (New England Biolabs, Inc., USA), and bisulfite conversion was performed with the EZ DNA Methylation-Lightning Kit (Zymo Research, USA). The resulting 150 bp paired-end libraries were sequenced with Illumina HiSeq™ 2000 and NovaSeq™ (Il-

lumina, Inc., USA). The raw reads were trimmed with Trim Galore! (Krueger et al. 2023), mapped against the Morex v3 reference genome with Bismark (Krueger and Andrews 2011), and aligned with Bowtie 2 (Langmead and Salzberg 2012). For quality control, SNPs were called with Bis-SNP (Liu et al. 2012) and compared with single nucleotide variation (SNV) data generated by DNA sequencing of the respective inbred lines (Weisweiler et al. 2022). The level of methylation in cytosine positions present in the methylated sequence contexts CpG, CHG, and CHH, was calculated as the percentage of the methylated reads per position. For each DRR population, the methylation level at each sequence context in a given genomic window was calculated as the average among the respective parental inbred lines' methylation level values for the methylated cytosine positions in that window, weighted by the number of methylated cytosine positions corresponding to each parent. The average methylation level across the three sequence contexts in a given genomic 1 Mbp window was calculated as the average among the calculated methylation levels for such contexts in the window, weighted by the number of methylated cytosine positions corresponding to each of the contexts in the window. The difference in methylation level between the two parental inbred lines of a population at each 1 Mbp genomic window was calculated for the three methylated sequence contexts and their average.

The gene density in 1 Mbp windows was calculated as the physical fraction spanned by the coding sequence of genes in each window. The locations of genes and intergenic regions on the barley chromosomes were obtained from the Morex v3

reference sequence. The genetic divergence among parents of the DRR populations was calculated from single nucleotide variant (SNV) data derived from genome-wide sequencing (Weisweiler et al. 2022).

The SNVs were also used to calculate the general recombination effects (*GRE*) of the parental inbred lines as described by Casale et al. (2022). In the next step, the proportion of the sum of the GRE of both parents for a given population that was not explained by the average methylation in each 1 Mbp window was estimated using linear regression. This residual was assumed to represent the genetic effects on recombination in a given genomic window. The specific recombination effect (*SRE*) for a given parental combination was not taken into account to estimate genetic effects because it was previously described to cause only a minor effect on the recombination rate of a given biparental barley population (Casale et al. 2022).

The SVs such as inversions, insertions, deletions, duplications, and translocations, between the parental inbred lines of the DRR populations and the Morex reference genome were obtained from Weisweiler et al. (2022). The SVs were categorized by size (50–299 bp, 0.3–4.9 kb, 5–49 kb, 50–249 kb, 0.25–1 Mbp, and >1 Mbp), except for translocations whose length was not determined (Table S1). The physical length fraction spanned by SVs in every 1 Mbp genomic window across the genome was estimated for all SV categories and sizes. Furthermore, the total SV span fraction generated by the sum of all SV categories and sizes in each 1 Mbp window was calculated (hereafter referred to as SV load).

Pearson’s correlation between all pairs of the abovementioned genomic features

was calculated by genomic window and population. To identify which of the features better explained the recombination rate variation among the 45 DRR populations at each 1 Mbp genomic window in the barley chromosomes, a stepwise regression approach was used. The procedure keeps for each 1 Mbp window the subset of genomic features that explain differences among the recombination rates of the DRR populations with the highest Akaike information criterion (AIC). The fraction of 1 Mbp windows of the barley genome in which a given genomic feature was retained in the model provides an estimation of the feature's importance across the entire genome. Moreover, the direction of the correlation between the analyzed features and the recombination rate provides a notion of the impact of the feature on either promoting or repressing recombination. The model included the total SV load, parental sequence divergence, genetic effects, average methylation level, and difference among parental inbred lines at the average methylation level. In addition, an extended model was constructed by breaking down the methylation-related variables into the respective methylated sequence contexts and the SV load into the different SV types.

***Investigation of the genomic features associated with the
recombination rate in barley at high-resolution***

Plant material, genotyping, and data cleaning

From the 45 DRR population set, three populations (HvDRR13, HvDRR27, and HvDRR28) were selected for high-density genotyping using an mRNAseq approach

as described by Arlt et al. (2023). These populations are the product of a triangle cross among three parental inbred lines (HOR8160, SprattArcher, and Unumli-Arpa). The respective 64, 92, and 79 RILs were cultivated at the S7 generation in petri dishes in a randomized incomplete block design where the parental inbred lines were included as controls. The blocks were harvested 7 days after planting with less than 2 hours difference between the first and last sample. The whole seedling was utilized for mRNA extraction. For mRNA sequencing (RNA-Seq), the RIL-specific library was constructed using the VAHTS Universal V6 RNA-seq Library Prep Kit for Illumina (Vazyme, China), and RIL-specific barcodes were used. The pooled libraries were sequenced on the DNBSEQ-G400 platform (MGI Tech Co., Ltd., China) by BGI Genomics (Beijing, China), generating 1.42 billion 150 bp paired-end reads. The reads were trimmed with Trimmomatic (Bolger et al. 2014) and aligned to the Morex v3 reference sequence using HISAT2 (Kim et al. 2019). Variant calling was performed using BCFtools (Li et al. 2009). The obtained variants were selected based on their intersection with the reported SNVs from the parental inbred lines (Weisweiler et al. 2022). Furthermore, a 12 K subset of the SNP markers reported previously (Casale et al. 2022) for the same parental inbred lines and RILs was added to the total set at genomic positions not present in the RNAseq dataset. On a population basis, SNPs carrying nonparental alleles were set to missing data, and SNPs with 100% missing data or monomorphic parental alleles were discarded. Missing data for genotypes at polymorphic positions were reconstructed using Beagle (Browning et al. 2018) with default parameters.

Detection of recombination events

A recombination event in a given RIL haplotype was called when a block of SNP alleles inherited from one parental inbred switched to a block of SNP alleles belonging to the other parent (i.e., parental allele phase change). The recombination breakpoints were determined as the midpoint of the region between both blocks (i.e., the CO interval). The blocks comprising fewer than three SNPs were considered false positive CO events and were discarded. Then, blocks shorter than 10 kb were considered GC events, while blocks longer than 10 kb were considered to be produced by CO (Yang et al. 2012; Gardiner et al. 2019). To enable comparisons with earlier studies (Yang et al. e.g. 2012; Gardiner et al. e.g. 2019), GC-related blocks were grouped into long (2–10 kb), short (20 bp–2 kb), and very short (2–19 bp) GC blocks, while the CO-related segments were grouped into short (10–500 kb), medium (500 kb–3 Mbp), and long (> 3 Mbp) CO blocks. The longest block length threshold that kept every major parental allele phase change (3 Mbp) was defined visually by graphical genotypes on a 500 kb scale from 0.5 to 20 Mbp (Figure S2). The CO block length categories were considered different CO layers, and further analyses were performed on a multilayer basis. Individuals with a CO count falling outside the 3-fold interquartile range of their respective population were assumed to be outliers and were discarded from further analyses.

The pericentromeric and distal regions of each chromosome were defined as explained above at 10 kb genomic windows for each of the three analyzed popula-

tions independently. The distal regions were defined as the chromosome segments between the pericentromeric region and the telomeres of the chromosomes.

In each 10 kb genomic window of the genome, the physical fraction spanned by the CO intervals determined in all RILs of a population was aggregated to calculate the accumulated CO probability per window on a population basis. The accumulated CO probabilities per window were normalized per chromosome and per population. The CO hotspot windows in a given population were defined as the windows with a CO probability $> 99\%$. The GC hotspot windows were determined in the same way as the CO hotspot windows. The windows located in the distal regions with a CO probability of zero were considered coldspot windows. The coldspot windows that were located beside other coldspot windows were combined into coldspot regions. To avoid calling coldspot windows near the pericentromeric region and telomeres, only windows located away from such regions were considered coldspot windows. This distance was granted by introducing an arbitrary margin with a length of 12.5% of the physical length of the respective distal region. To control for spurious associations generated by the variation of the features along the chromosomes, the windows from the distal regions were grouped into those close to the telomere (distal telomeric) and those close to the pericentromeric region (distal proximal).

Investigation of the association between genomic features and recombination rate

The fraction of each window spanned by SVs, the gene density, the methylation level at the sequence contexts CpG, CHG, and CHH, their average, and the parental difference for these contexts were calculated as described above for each 10 kb genomic window of the barley chromosomes in the three analyzed populations HvDRR13, HvDRR27, and HvDRR28. In addition, the 10 kb windows neighboring coldspot and hotspot regions (any genomic length spanned by contiguous coldspot or hotspot 10 kb windows) were grouped by their relative position in the range from -40 kb to +40 kb around the respective coldspot or hotspot in the mentioned chromosome regions. The statistical comparison among window groups of any kind for the mentioned features was performed with the Mann–Whitney U test with Bonferroni correction for multiple testing.

The observed overlaps of the coldspot and hotspot regions with genes and intergenic regions in the chromosomes of the three analyzed populations were statistically compared with random overlaps simulated with regioneR (Gel et al. 2016) by running 1,000 permutations on a by-chromosome basis where the respective pericentromeric regions were masked.

The observed overlap between CO intervals and the genomic regions spanned by SVs was assessed as described above with regioneR for insertions/deletions, duplications, and inversions. In addition, the mean distance between the CO breakpoints

and their closest SV in the genome was compared with the equivalent of random simulated COs in 10 Mbp genomic windows, each with a probability of CO occurrence according to the recombination rate per chromosome in the DRR populations reported by Casale et al. (2022). The significant differences among the means of the observed and simulated CO-SV distances were evaluated with the Mann–Whitney U test.

DATA ACCESS

The methylation data of the 23 DRR population parental lines and the raw mRNA sequencing data of the RILs from populations HvDRR13, HvDRR27, and HvDRR28 (and their parental lines) are both available at NCBI, BioProject accessions PRJNA1100572 and 1088431, respectively. Employed scripts are available from the authors upon request.

COMPETING INTEREST STATEMENT

The authors declare that they have no competing interests.

ACKNOWLEDGEMENTS

Computational infrastructure and support were provided by the Centre for Information and Media Technology at Heinrich Heine University Düsseldorf. The authors give thanks to the IPK for providing the seeds of the diversity panel. We thank our former colleagues Andrea Lossow, Nicole Kliche-Kamphaus, Nele Kaul,

Isabelle Scheibert, Marianne Haperscheid, George Alskief, Florian Esser as well as our present colleagues Konstantin Shek and Stefanie Krey for their excellent technical assistance in creating and maintaining the DRR populations.

FUNDING

This research is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy (EXC 2048/1, Project ID: 390686111) and core funding of HHU and JKI.

AUTHORS' CONTRIBUTIONS

FC contributed to the design of the project, analysed the data and performed the analyses related to meiotic recombination, and performed all statistical analyses. CA created the libraries, sequenced, and processed the RNA-Seq data. MK analysed the bisulfite-sequenced DNA data. Both CA and MK contributes equally to this work as second authors. JL created the segregating populations. JE and TH created the bisulfite libraries for DNA sequencing. BS designed and coordinated the project. FC and BS wrote the manuscript. All authors read and approved the final manuscript.

Table 1: The observed overlaps among the recombination coldspot and hotspot 10 kb windows with genes and the intergenic regions in the distal region of the barley chromosomes, and their comparison with the overlaps generated under a random distribution of such regions in the analyzed double round-robin (DRR) populations HvDRR13, HvDRR27, and HvDRR28. The random distribution of the genomic regions was simulated with 1,000 permutations.

Recombination region	Genetic region	Population	Observed overlaps	Random overlaps
Coldspot windows	Genes	HvDRR13	23989	37029
		HvDRR27	19545	26369
		HvDRR28	24832	37091
	Intergenic regions	HvDRR13	77006	129894
		HvDRR27	73553	87290
		HvDRR28	78332	135060
Hotspot windows	Genes	HvDRR13	1057***	237
		HvDRR27	739***	146
		HvDRR28	1264***	282
	Intergenic regions	HvDRR13	823	840
		HvDRR27	684***	486
		HvDRR28	1057	1031

*** $P < 0.001$ for $H_0: \mu_{Observed} \leq \mu_{Random}$; Z-test.

Table 2: Comparison of the methylation level at sequence contexts CpG, CHG, and CHH in the 10 kb windows identified as recombination hotspots and coldspots in the double round-robin (DRR) populations HvDRR13, HvDRR27, and HvDRR28 for the respective parental inbred lines. The coldspot and hotspots windows were calculated on the basis of the recombination rate of the DRR populations. For a given population and given methylated sequence context, the significant difference ($\alpha = 0.016$) in the Wilcoxon's rank sum test among genomic window groups are indicated with different letters.

Population	Coldspot windows				Hotspot windows			
	Offspring	HOR8160	Unumli-Arpa	SprattArcher	Offspring	HOR8160	Unumli-Arpa	SprattArcher
HvDRR13								
CpG	0.880 c	0.875 b	-	0.883 a	0.666 a	0.661 a	-	0.668 a
CHG	0.581 c	0.571 b	-	0.590 a	0.334 a	0.322 a	-	0.345 a
CHH	0.020 c	0.020 b	-	0.021 a	0.029 a	0.028 a	-	0.030 a
HvDRR27								
CpG	0.894 b	-	0.891 a	0.892 b	0.673 a	-	0.667 a	0.678 a
CHG	0.595 c	-	0.587 a	0.602 b	0.346 a	-	0.330 a	0.362 a
CHH	0.019 b	-	0.019 a	0.020 b	0.029 ab	-	0.027 a	0.031 b
HvDRR28								
CpG	0.880 c	0.875 b	0.881 a	-	0.645 a	0.643 a	0.644 a	-
CHG	0.573 c	0.570 b	0.574 a	-	0.319 a	0.319 a	0.315 a	-
CHH	0.019 a	0.019 b	0.019 a	-	0.028 a	0.028 a	0.027 a	-

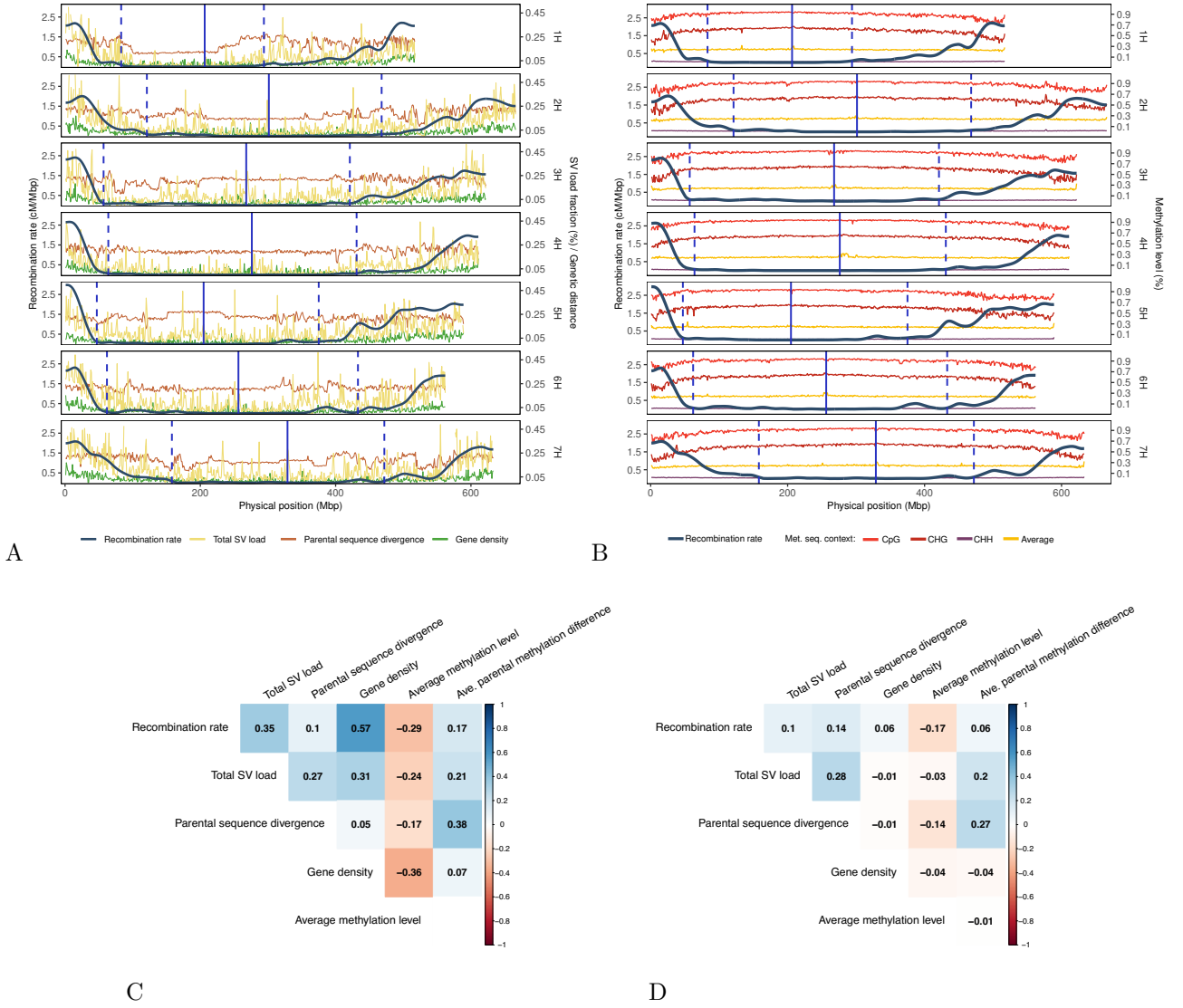
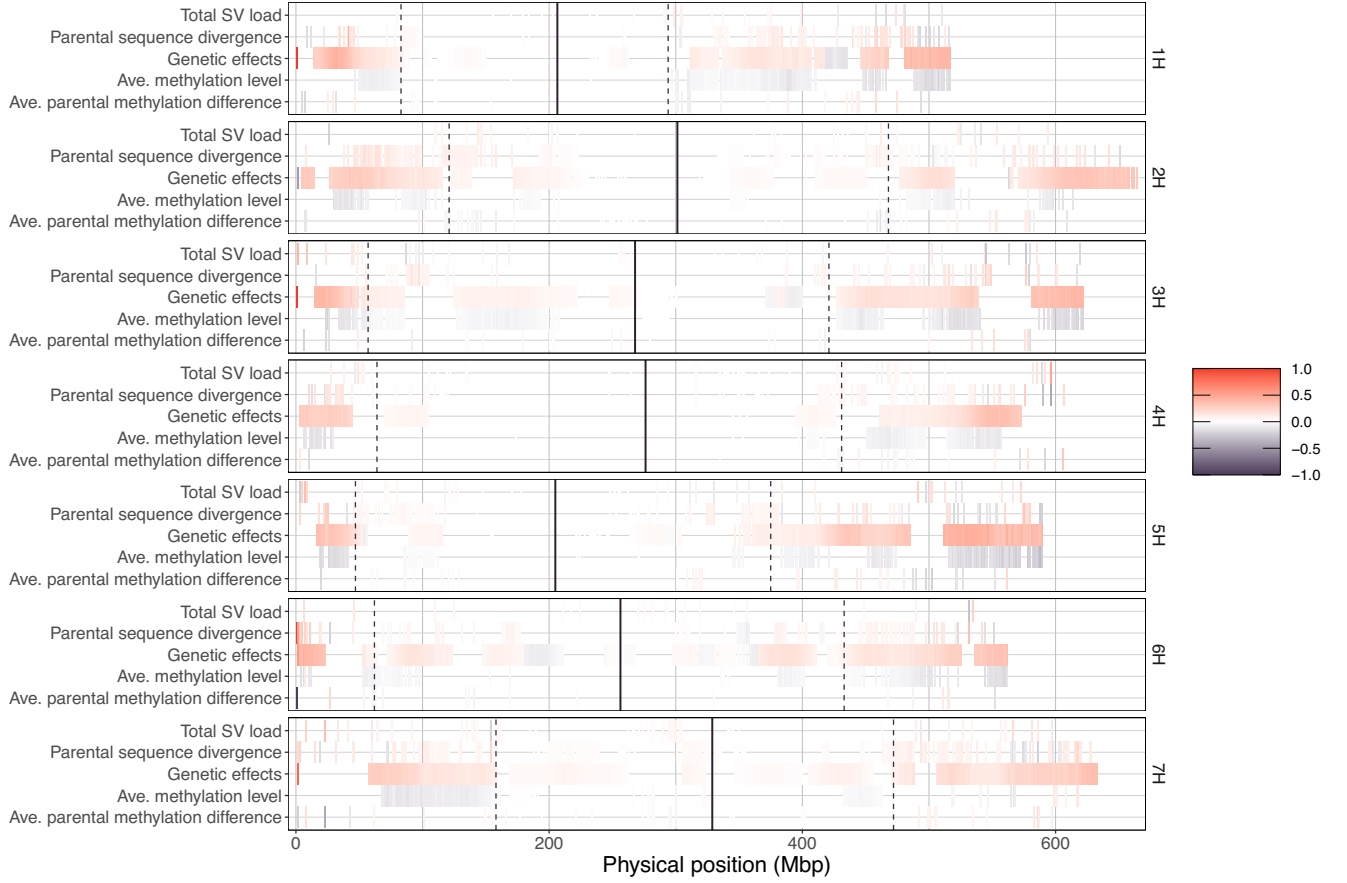


Fig. 1: (A) Distribution of recombination rate, total structural variants' load fraction, gene density, and parental sequence divergence between the respective parental inbred lines on average across the 45 double round-robin (DRR) populations in 1 Mbp windows across the seven barley chromosomes. The total structural variants (SV) load fraction represents the portion spanned of a given 1 Mbp genomic window by the sum of insertions, deletions, inversions, and duplications. (B) Distribution of recombination rate and methylation level in 1 Mbp windows for the sequence contexts CpG, CHG, CHH, and their mean, on average across the seven barley chromosomes and 45 DRR populations. The methylation level values for a given methylation sequence context represent the percentage of methylated reads of such context present in a 1 Mbp window averaged across the 45 DRR populations. The vertical blue solid line indicates the position of the centromere in the Morex v3 reference genome and the vertical blue dashed lines indicate the pericentromeric region calculated across the 45 DRR populations. (C-D) Correlation matrix between recombination rate, total SV load fraction, parental sequence divergence, gene density, average methylation level, and parental difference in methylation level, across 1 Mbp windows in the distal (C) and the pericentromeric (D) regions of the barley chromosomes for the average across the 45 DRR populations. Pearson's correlation coefficients are indicated with a color gradient from -1 (red) to 1 (blue).



Genomic feature	Chromosome region			
	Distal		Pericentromeric	
	+	-	+	-
Total SV load	0.05	0.02	0.04	0.01
Parental sequence divergence	0.24	0.04	0.19	0.01
Genetic effects	0.76	0.01	0.40	0.06
Average methylation level	0.00	0.43	0.01	0.12
Ave. parental methylation difference	0.04	0.03	0.02	0.02

Fig. 2: The genomic features explaining the variation in the recombination rate among the 45 double round-robin (DRR) populations in 1 Mbp windows across the seven barley chromosomes. The genomic features were selected for each window by a stepwise regression procedure. The standardized regression coefficients are indicated by a color gradient from -1 (purple) to 1 (red). The studied genomic features included sequence divergence among parental inbred lines, genetic effects, total structural variants (SV) load, average methylation level across the sequence contexts CpG, CHG, and CHH, and the difference in the methylation level among the parental inbred lines for the average of such sequence contexts. The vertical solid line indicates the position of the centromere in the reference genome, and the vertical dashed lines indicate the pericentromeric region calculated across the 45 DRR populations. The analysis was performed separately for the distal and pericentromeric regions of the barley chromosomes, respectively. The fraction of 1 Mbp windows of the barley genome in which the genomic features were found to be significantly associated with the recombination rate variation across the 45 DRR populations are displayed in the table. The fractions of 1 Mbp windows positively and negatively associated with the recombination rate were indicated with the "+" or the "-" signs, respectively.

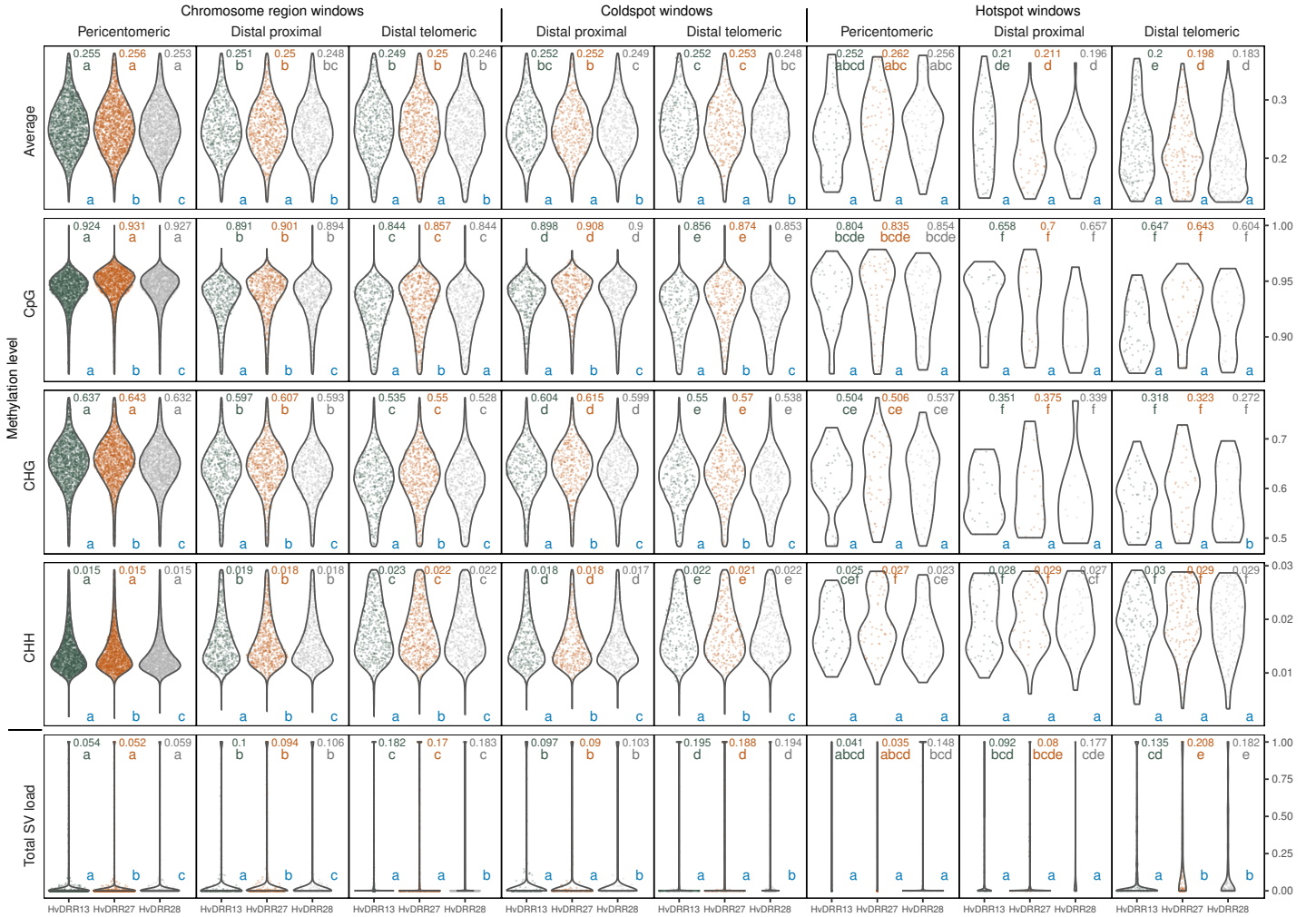


Fig. 3: Distribution of the methylation level for the methylated sequence contexts CpG, CHG, CHH, and their average, and the total structural variants (SV) load fraction in 10 kb genomic windows grouped by their location in different chromosomal regions -pericentromeric, distal proximal, and distal telomeric-, and recombination rate -total, coldspots, and hotspots- of the barley chromosomes in the analyzed double round-robin (DRR) populations HvDRR13 (green), HvDRR27 (orange), and HvDRR28 (grey). For each DRR population, the methylation level at each sequence context in a given genomic window was calculated as the average among the respective parental inbred lines' methylation level values for the methylated cytosine positions in that window, weighted by the number of methylated cytosine positions corresponding to each parent. The displayed dots for chromosome regions and coldspots are a random subset of 1% of the total windows in each window group. For each genomic feature, the distribution's mean of each window group from a given population is indicated with the related population color at the top-right of the respective plot. Significant differences ($\alpha = 0.001$) in the Wilcoxon's rank sum test among such means across window groups is indicated with different letters below the respective mean and sharing the population color. For a given genomic feature in each window grouping category, significant differences ($\alpha = 0.008$) in the Wilcoxon's rank sum test among populations are indicated with different blue letters at the bottom of each panel.

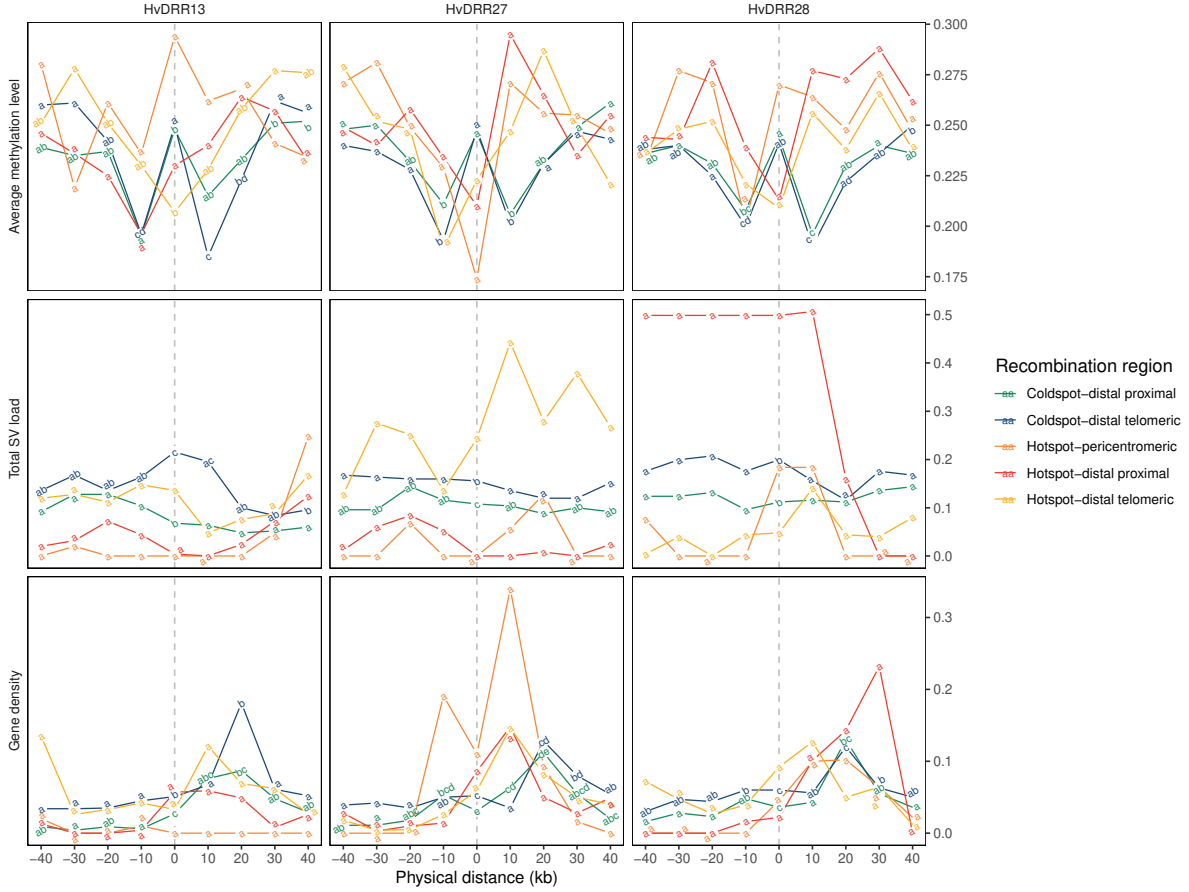


Fig. 4: The mean values of the average methylation level among the methylated sequence contexts CpG, CHG, and CHH, total structural variants (SV) load, and gene density in the genomic windows grouped by the physical positions in the range from -40 kb to +40 kb around the coldspot and hotspot genomic regions in the pericentromeric, distal proximal, and distal telomeric region of the chromosomes in the analyzed double round-robin (DRR) populations HvDRR13, HvDRR27, and HvDRR28. The vertical dashed lines indicate the relative location of either the coldspot or the hotspot windows, respectively. The significant difference ($\alpha = 0.001$) in the Wilcoxon's rank sum test among the window groups corresponding to the different 10 kb physical position neighboring the respective hotspot or coldspot region in a particular chromosome region is indicated with different letters.

REFERENCES

- Apuli RP, Bernhardsson C, Schiffthaler B, Robinson KM, Jansson S, Street NR, and Ingvarsson PK. 2020. Inferring the genomic landscape of recombination rate variation in european aspen (*Populus tremula*). *G3: Genes, Genomes, Genetics* **10**: 299–309.
- Arbeithuber B, Betancourt AJ, Ebner T, and Tiemann-Boege I. 2015. Crossovers are associated with mutation and biased gene conversion at recombination hotspots. *Proceedings of the National Academy of Sciences* **112**: 2109–2114.
- Arlt C, Wachtmeister T, Köhrer K, and Stich B. 2023. Affordable, accurate and unbiased RNA sequencing by manual library miniaturization: A case study in barley. *Plant Biotechnology Journal* **21**: 2241–2253.
- Barton NH and Charlesworth B. 1998. Why sex and recombination? *Science* **281**: 1986–1990.
- Baudat F and Massy BD. 2007. Regulating double-stranded dna break repair towards crossover or non-crossover during mammalian meiosis. *Chromosome research* **15**: 565–577.
- Bauer E, Falque M, Walter H, Bauland C, Camisan C, Campo L, Meyer N, Ranc N, Rincenc R, Schipprack W, et al.. 2013. Intraspecific variation of recombination rate in maize. *Genome Biology* **14**.
- Bayer MM, Rapazote-Flores P, Ganai M, Hedley PE, Macaulay M, Plieske J, Ramsay L, Russell J, Shaw PD, Thomas W, et al.. 2017. Development and evaluation of a barley 50k iselect SNP array. *Frontiers in Plant Science* **8**: 1792.
- Blackwell AR, Dlużewska J, Szymanska-Lejman M, Desjardins S, Tock AJ, Kbiri N, Lambing C, Lawrence EJ, Bieluszewski T, Rowan B, et al.. 2020. MSH2 shapes the meiotic crossover landscape in relation to interhomolog polymorphism in Arabidopsis. *The EMBO Journal* **39**: 1–22.
- Blary A and Jenczewski E. 2019. Manipulation of crossover frequency and distribution for plant breeding. *Theoretical and Applied Genetics* **132**: 575–592.
- Boideau F, Richard G, Coriton O, Huteau V, Belser C, Deniot G, Eber F, Falentin C, de Carvalho JF, Gilet M, et al.. 2022. Epigenomic and structural events preclude recombination in Brassica napus. *New Phytologist* **234**: 545–559.
- Bolger AM, Lohse M, and Usadel B. 2014. Trimmomatic: A flexible trimmer for illumina sequence data. *Bioinformatics* **30**: 2114–2120.
- Bouchet S, Olatoye MO, Marla SR, Perumal R, Tesso T, Yu J, Tuinstra M, and Morris GP. 2017. Increased Power To Dissect Adaptive Traits in Global Sorghum

- Diversity Using a Nested Association Mapping Population. *Genetics* **206**: 573–585.
- Browning BL, Zhou Y, and Browning SR. 2018. A one-penny imputed genome from next-generation reference panels. *American Journal of Human Genetics* **103**: 338–348.
- Burt A. 2000. Perspective: Sex, recombination, and the efficacy of selection - was weismann right? *Evolution* **54**: 337–351.
- Casale F, Van Inghelandt D, Weisweiler M, Li J, and Stich B. 2022. Genomic prediction of the recombination rate variation in barley – A route to highly recombinogenic genotypes. *Plant Biotechnology Journal* **20**: 676–690.
- Choi K and Henderson IR. 2015. Meiotic recombination hotspots - a comparative view. *Plant Journal* **83**: 52–61.
- Choi K, Zhao X, Kelly KA, Venn O, Higgins JD, Yelina NE, Hardcastle TJ, Ziolkowski PA, Copenhaver GP, Franklin FCH, et al.. 2013. Arabidopsis meiotic crossover hot spots overlap with H2A.Z nucleosomes at gene promoters. *Nature Genetics* **45**: 1327–1338.
- Colomé-Tatché M, Cortijo S, Wardenaar R, Morgado L, Lahouz B, Sarazin A, Etcheverry M, Martin A, Feng S, Duvernois-Berthet E, et al.. 2012. Features of the arabidopsis recombination landscape resulting from the combined loss of sequence variation and dna methylation. *Proceedings of the National Academy of Sciences of the United States of America* **109**: 16240–16245.
- Darrier B, Rimbert H, Balfourier F, Pingault L, Josselin AA, Servin B, Navarro J, Choulet F, Paux E, and Sourdille P. 2017. High-resolution mapping of crossover events in the hexaploid wheat genome suggests a universal recombination mechanism. *Genetics* **206**: 1373–1388.
- Dreissig S, Mascher M, Heckmann S, and Purugganan M. 2019. Variation in recombination rate is shaped by domestication and environmental conditions in barley. *Molecular Biology and Evolution* **36**: 2029–2039.
- Dreissig S, Maurer A, Sharma R, Milne L, Flavell AJ, Schmutzer T, and Pillen K. 2020. Natural variation in meiotic recombination rate shapes introgression patterns in intraspecific hybrids between wild and domesticated barley. *New Phytologist* **228**: 1852–1863.
- Esch E, Szymaniak JM, Yates H, Pawlowski WP, and Buckler ES. 2007. Using crossover breakpoints in recombinant inbred lines to identify quantitative trait loci controlling the global recombination frequency. *Genetics* **177**: 1851–1858.

- Fernandes JB, Naish M, Lian Q, Burns R, Tock AJ, Rabanal FA, Wlodzimierz P, Habring A, Nicholas RE, Weigel D, et al.. 2024. Structural variation and dna methylation shape the centromere-proximal meiotic crossover landscape in arabidopsis. *Genome Biology* **25**: 30.
- Gardiner LJ, Wingen LU, Bailey P, Joynson R, Brabbs T, Wright J, Higgins JD, Hall N, Griffiths S, Clavijo BJ, et al.. 2019. Analysis of the recombination landscape of hexaploid bread wheat reveals genes controlling recombination and gene conversion frequency. *Genome Biology* **20**: 69.
- Gel B, Díez-Villanueva A, Serra E, Buschbeck M, Peinado MA, and Malinverni R. 2016. RegioneR: An R/Bioconductor package for the association analysis of genomic regions based on permutation tests. *Bioinformatics* **32**: 289–291.
- Gion JM, Hudson CJ, Lesur I, Vaillancourt RE, Potts BM, and Freeman JS. 2016. Genome-wide variation in recombination rate in Eucalyptus. *BMC Genomics* **17**: 1–12.
- Giraut L, Falque M, Drouaud J, Pereira L, Martin OC, and Mézard C. 2011. Genome-wide crossover distribution in arabidopsis thaliana meiosis reveals sex-specific patterns along chromosomes. *PLoS Genetics* **7**: e1002354.
- Gutierrez-Gonzalez JJ, Mascher M, Poland J, and Muehlbauer GJ. 2019. Dense genotyping-by-sequencing linkage maps of two Synthetic W7984×Opata reference populations provide insights into wheat structural diversity. *Scientific Reports* **9**: 1–15.
- Hall JC. 1972. Chromosome segregation influenced by two alleles of the meiotic mutant c(3)G in Drosophila melanogaster. *Genetics* **71**: 367–400.
- Halldorsson BV, Palsson G, Stefansson OA, Jonsson H, Hardarson MT, Eggertsson HP, Gunnarsson B, Oddsson A, Halldorsson GH, Zink F, et al.. 2019. Characterizing mutagenic effects of recombination through a sequence-level genetic map. *Science* **363**.
- Henderson IR. 2012. Control of meiotic recombination frequency in plant genomes. *Current Opinion in Plant Biology* **15**: 556–561.
- Higgins JD, Perry RM, Barakate A, Ramsay L, Waugh R, Halpin C, Armstrong SJ, and Franklin FCH. 2012. Spatiotemporal asymmetry of the meiotic program underlies the predominantly distal distribution of meiotic crossovers in barley. *Plant Cell* **24**: 4096–4109.
- Hsu YM, Falque M, and Martin OC. 2022. Quantitative modelling of fine-scale variations in the Arabidopsis thaliana crossover landscape. *Quantitative Plant Biology* **3**: e3.

- Jordan KW, Wang S, He F, Chao S, Lun Y, Paux E, Sourdille P, Sherman J, Akhunova A, Blake NK, et al.. 2018. The genetic architecture of genome-wide recombination rate variation in allopolyploid wheat revealed by nested association mapping. *Plant Journal* **95**: 1039–1054.
- Kim D, Paggi JM, Park C, Bennett C, and Salzberg SL. 2019. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature Biotechnology* **37**: 907–915.
- Kim S, Plagnol V, Hu TT, Toomajian C, Clark RM, Ossowski S, Ecker JR, Weigel D, and Nordborg M. 2007. Recombination and linkage disequilibrium in *Arabidopsis thaliana*. *Nature Genetics* **39**: 1151–1155.
- Krueger F and Andrews SR. 2011. Bismark: a flexible aligner and methylation caller for bisulfite-seq applications. *Bioinformatics* **27**: 1571–1572.
- Krueger F, James F, Ewels P, Afyounian E, Weinstein M, Schuster-Boeckler B, Hulselmans G, and sclamons. 2023. Trimalore v0.6.10.
- Langmead B and Salzberg S. 2012. Fast gapped-read alignment with Bowtie 2. *Nature methods* **9**: 357–359.
- Lawrence EJ, Griffin CH, and Henderson IR. 2017. Modification of meiotic recombination by natural variation in plants. *Journal of Experimental Botany* **68**: 5471–5483.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, and Durbin R. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**: 2078–2079.
- Li X, Li L, and Yan J. 2015. Dissecting meiotic recombination based on tetrad analysis by single-microspore sequencing in maize. *Nature Communications* **6**.
- Liu Y, Siegmund K, Laird P, and Berman B. 2012. Bis-SNP: combined DNA methylation and SNP calling for bisulfite-seq data. *Genome biology* **13**: R61.
- Lu P, Han X, Qi J, Yang J, Wijeratne AJ, Li T, and Ma H. 2012. Analysis of *Arabidopsis* genome-wide variations before and after meiosis and meiotic recombination by resequencing *landsberg erecta* and all four products of a single meiosis. *Genome Research* **22**: 508–518.
- Mancera E, Bourgon R, Brozzi A, Huber W, and Steinmetz M. 2009. High-resolution mapping of meiotic crossovers and noncrossovers in yeast. *Nature* **454**: 479–485.
- Marand AP, Zhao H, Zhang W, Zeng Z, Fang C, and Jianga J. 2019. Historical meiotic crossover hotspots fueled patterns of evolutionary divergence in rice. *Plant Cell* **31**: 645–662.

- Martini E, Diaz RL, Hunter N, and Keeney S. 2006. Crossover homeostasis in yeast meiosis. *Cell* **126**: 285–295.
- Mascher M, Wicker T, Jenkins J, Plott C, Lux T, Koh CS, Ens J, Gundlach H, Boston LB, Tulpová Z, et al.. 2021. Long-read sequence assembly: A technical evaluation in barley. *Plant Cell* **33**: 1888–1906.
- McMullen MD, Kresovich S, Villeda HS, Bradbury P, Li H, Sun Q, Flint-Garcia S, Thornsberry J, Acharya C, Bottoms C, et al.. 2009. Genetic properties of the maize nested association mapping population. *Science* **325**: 737–740.
- Melamed-Bessudo C and Levy AA. 2012. Deficiency in dna methylation increases meiotic crossover rates in euchromatic but not in heterochromatic regions in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America* **109**.
- Mercier R, Mézard C, Jenczewski E, Macaisne N, and Grelon M. 2015. The molecular biology of meiosis in plants. *Annual Review of Plant Biology* **66**: 297–327.
- Mirouze M, Lieberman-Lazarovich M, Aversano R, Bucher E, Nicolet J, Reinders J, and Paszkowski J. 2012. Loss of DNA methylation affects the recombination landscape in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America* **109**: 5880–5885.
- Muller HJ. 1932. Some genetic aspects of sex. *The American Naturalist* **66**: 118–138.
- Mézard C. 2006. Meiotic recombination hotspots in plants. *Biochemical Society Transactions* **34**: 531–534.
- Nachman MW. 2002. Variation in recombination rate across the genome: Evidence and implications. *Current Opinion in Genetics and Development* **12**: 657–663.
- Paape T, Zhou P, Branca A, Briskine R, Young N, and Tiffin P. 2012. Fine-scale population recombination rates, hotspots, and correlates of recombination in the *Medicago truncatula* genome. *Genome Biology and Evolution* **4**: 726–737.
- Pan J, Sasaki M, Kniewel R, Murakami H, Blitzblau HG, Tischfield SE, Zhu X, Neale MJ, Jasin M, Socci ND, et al.. 2011. A hierarchical combination of factors shapes the genome-wide topography of yeast meiotic recombination initiation. *Cell* **144**: 719–731.
- Peck JR. 1994. A ruby in the rubbish: Beneficial mutations, deleterious mutations and the evolution of sex. *Genetics* **137**: 597–606.
- Qi J, Chen Y, Copenhaver GP, and Ma H. 2014. Detection of genomic variations and dna polymorphisms and impact on analysis of meiotic recombination and

- genetic mapping. *Proceedings of the National Academy of Sciences of the United States of America* **111**: 10007–10012.
- Ritz KR, Noor MA, and Singh ND. 2017. Variation in recombination rate: Adaptive or not? *Trends in Genetics* **33**: 364–374.
- Rodgers-Melnick E, Bradbury PJ, Elshire RJ, Glaubitz JC, Acharya CB, Mitchell SE, Li C, Li Y, and Buckler ES. 2015. Recombination in diverse maize is stable, predictable, and associated with genetic load. *Proceedings of the National Academy of Sciences of the United States of America* **112**: 3823–3828.
- Rowan BA, Heavens D, Feuerborn TR, Tock AJ, Henderson IR, and Weigel D. 2019. An ultra high-density *Arabidopsis thaliana* crossover. *Genetics* **213**: 771–787.
- Saintenac C, Faure S, Remay A, Choulet F, Ravel C, Paux E, Balfourier F, Feuillet C, and Sourdille P. 2011. Variation in crossover rates across a 3-mb contig of bread wheat (*triticum aestivum*) reveals the presence of a meiotic recombination hotspot. *Chromosoma* **120**: 185–198.
- Salomé PA, Bomblies K, Fitz J, Laitinen RA, Warthmann N, Yant L, and Weigel D. 2012. The recombination landscape in *Arabidopsis thaliana* f2 populations. *Heredity* **108**: 447–455.
- Serra H, Choi K, Zhao X, Blackwell AR, Kim J, and Henderson IR. 2018. Inter-homolog polymorphism shapes meiotic crossover within the *Arabidopsis* RAC1 and RPP13 disease resistance genes. *PLoS Genetics* **14**.
- Shen C, Wang N, Huang C, Wang M, Zhang X, and Lin Z. 2019. Population genomics reveals a fine-scale recombination landscape for genetic improvement of cotton. *Plant Journal* **99**: 494–505.
- Si W, Yuan Y, Huang J, Zhang X, Zhang Y, Zhang Y, Tian D, Wang C, Yang Y, and Yang S. 2015. Widely distributed hot and cold spots in meiotic recombination as shown by the sequencing of rice F2 plants. *New Phytologist* **206**: 1491–1502.
- Silva-Junior OB and Grattapaglia D. 2015. Genome-wide patterns of recombination, linkage disequilibrium and nucleotide diversity from pooled resequencing and single nucleotide polymorphism genotyping unlock the evolutionary history of *Eucalyptus grandis*. *New Phytologist* **208**: 830–845.
- Sun H, Rowan BA, Flood PJ, Brandt R, Fuss J, Hancock AM, Michelmore RW, Huettel B, and Schneeberger K. 2019. Linked-read sequencing of gametes allows efficient genome-wide analysis of meiotic recombination. *Nature Communications* **10**: 1–9.

- Sun Y, Ambrose JH, Haughey BS, Webster TD, Pierrie SN, Muñoz DF, Wellman EC, Cherian S, Lewis SM, Berchowitz LE, et al.. 2012. Deep genome-wide measurement of meiotic gene conversion using tetrad analysis in *Arabidopsis thaliana*. *PLoS Genetics* **8**: e1002968.
- Szostak JW, Orr-Weaver TL, Rothstein RJ, and Stahl FW. 1983. The double-strand-break repair model for recombination. *Cell* **33**: 25–35.
- Wang J, Street NR, Scofield DG, and Ingvarsson PK. 2016. Natural selection and recombination rate variation shape nucleotide polymorphism across the genomes of three related populus species. *Genetics* **202**: 1185–1200.
- Weisweiler M, Arlt C, Wu PY, Inghelandt DV, Hartwig T, and Stich B. 2022. Structural variants in the barley gene pool: precision and sensitivity to detect them using short-read sequencing and their association with gene expression and phenotypic variation. *Theoretical and Applied Genetics* **135**: 3511–3529.
- Wijnker E, James GV, Ding J, Becker F, Klasen JR, Rawat V, Rowan BA, de Jong DF, de Snoo CB, Zapata L, et al.. 2013. The genomic landscape of meiotic crossovers and gene conversions in *Arabidopsis thaliana*. *eLife* **2**: e01426.
- Yang S, Yuan Y, Wang L, Li J, Wang W, Liu H, Chen JQ, Hurst LD, and Tian D. 2012. Great majority of recombination events in *Arabidopsis* are gene conversion events. *Proceedings of the National Academy of Sciences of the United States of America* **109**: 20992–20997.
- Yelina NE, Choi K, Chelysheva L, Macaulay M, de Snoo B, Wijnker E, Miller N, Drouaud J, Grelon M, Copenhaver GP, et al.. 2012. Epigenetic remodeling of meiotic crossover frequency in *arabidopsis thaliana* DNA methyltransferase mutants. *PLoS Genetics* **8**: e1002844.

SUPPORTING INFORMATION

Table S1: Number of structural variants (SV) segregating in the analyzed double round-robin (DRR) populations HvDRR13 HvDRR27, and HvDRR28, classified by type (inversions, deletions, insertions, duplications, and translocations) and size (50—299 bp, 0.3—4.9 kb, 5—49 kb, 50—249 kb, 0.25—1 Mbp, and >1 Mbp) on average per recombinant inbred line (RIL).

SV	SV size	Populations		
		HvDRR13	HvDRR27	HvDRR28
inversions	50—299 bp	2.73	4.02	6.79
	0.3—4.9 kb	2.41	3.68	6.26
	5—49 kb	3.40	5.29	8.82
	50—249 kb	3.49	5.28	8.85
	0.25—1 Mbp	1.73	2.56	4.15
deletions	50—299 bp	194.41	310.01	520.13
	0.3—4.9 kb	147.63	235.56	391.03
	5—49 kb	124.13	193.98	322.59
	50—249 kb	5.73	8.39	13.82
	0.25—1 Mbp	0.35	0.58	1.00
insertions	50—299 bp	104.02	158.99	263.67
	0.3—4.9 kb	11.83	17.06	28.49
duplications	50—299 bp	31.81	50.33	84.38
	0.3—4.9 kb	41.54	63.29	106.15
	5—49 kb	106.02	157.11	261.18
	50—249 kb	9.89	14.72	24.33
	0.25—1 Mbp	1.22	1.78	3.03
translocations	-	467.43	366.05	400.85

Table S2: The fraction of 1 Mbp windows of the barley genome in which the genomic features were found to be significantly associated with the recombination rate variation across the 45 double round-robin (DRR) populations. The analysis was performed separately for the distal and pericentromeric regions of the barley chromosomes, respectively. The studied genetic features included parental sequence divergence, genetic effects, and span fraction of structural variants (SV) such as insertions, deletions, inversions, and duplications. The epigenetic features included the methylated level for the sequence contexts CpG, CHG, and CHH, as well as, the difference in the methylation level among the parental inbred lines for such sequence contexts. The significant correlations between the recombination rate and the mentioned features were determined by a stepwise regression procedure. The fraction of 1 Mbp windows positively and negatively associated with the recombination rate were indicated with the "+" or the "-" signs, respectively.

Genomic feature	Chromosome region			
	Distal		Pericentromeric	
	+	-	+	-
Extended model				
Insertions	0.05	0.03	0.04	0.03
Deletions	0.07	0.04	0.06	0.03
Inversions	0.03	0.02	0.02	0.01
Duplications	0.04	0.03	0.05	0.03
Parental sequence divergence	0.26	0.05	0.20	0.02
Genetic effects	0.76	0.01	0.40	0.06
Methylation level: CpG	0.03	0.29	0.04	0.08
Methylation level: CHG	0.02	0.12	0.02	0.07
Methylation level: CHH	0.02	0.18	0.02	0.07
Parental methylation difference: CpG	0.04	0.08	0.04	0.04
Parental methylation difference: CHG	0.08	0.03	0.03	0.03
Parental methylation difference: CHH	0.08	0.02	0.06	0.03

Table S3: The quality metrics of the filtered high-density RNAseq single nucleotide polymorphism (SNP) markers in the analyzed double round-robin (DRR) populations HvDRR13, HvDRR27, and HvDRR28.

Population	Molecular level	No. markers	Density (mrk/Mbp)	Inter-SNP distance		
				Median (bp)	Mean (bp)	Max (Mbp)
HvDRR13	Genome-wide	214,112	51	135	19,587	11.99
	Chromosomes					
	1H	20,220	39	122	25,525	6.03
	2H	36,824	55	137	18,074	3.47
	3H	43,542	70	147	14,272	3.13
	4H	16,158	26	153	37,761	11.99
	5H	39,233	67	138	14,979	2.95
	6H	27,092	48	135	20,733	5.48
HvDRR27	Genome-wide	259,540	62	130	16,162	5.45
	Chromosomes					
	1H	33,233	64	137	15,530	4.50
	2H	44,013	66	120	15,121	4.91
	3H	48,897	79	144	12,709	2.90
	4H	28,722	47	144	21,243	4.76
	5H	34,566	59	125	17,002	4.24
	6H	34,024	61	132	16,509	5.45
HvDRR28	Genome-wide	244,353	58	132	17,167	7.17
	Chromosomes					
	1H	32,732	63	137	15,768	5.22
	2H	40,850	61	123	16,292	3.52
	3H	37,402	60	137	16,615	3.29
	4H	28,759	47	150	21,202	5.79
	5H	44,375	75	138	13,253	2.96
	6H	28,936	52	138	19,412	7.17
	7H	31,299	49	111	20,194	4.60

Table S4: The central tendency metrics of the length (bp) and the number of markers per block (N markers/block) of the crossover (CO) related parental marker blocks. The marker blocks are presented for three different thresholds (>10 kb, >500 kb, and >3 Mbp) on average per recombinant inbred lines (RIL) in the analyzed double round-robin (DRR) populations HvDRR13, HvDRR27, and HvDRR28. Pearson's correlation coefficient (r) between block length and N markers per block is indicated as well as the respective significance.

CO layer	Block length (bp)			N markers/block			r	Significance
	median	mean	sd	median	mean	sd		
>3 Mbp	27,540,588	111,572,731	171,044,519	3,282	5,962	7,339	0.84	***
>500 kb	7,033,905	44,161,151	94,042,919	336	2,395	4,260	0.78	***
>10 kb	886,623	14,799,456	48,144,287	50	817	2,024	0.75	***

*** $P < 0.001$; Pearson's correlation.

Table S5: The summary of the crossover (CO) and gene conversion (GC) events classified by their related marker block size (COs: >10 kb, >500 kb, and >3 Mbp; GCs: 2—20 bp, 20 bp—2 kb, and 2—10 kb, respectively) on average per recombinant inbred line (RIL) in the analyzed double round-robin (DRR) populations HvDRR13, HvDRR27, and HvDRR28.

Population	Molecular level	Crossovers (COs)			Gene conversions (GCs)		
		>3 Mbp	>500 kb	>10 kb	2—10 kb	20 bp—2 kb	2—20 bp
HvDRR13	Genome-wide	32	96	283	51	225	6,584
	Chromosomes						
	1H	4.92	11.43	30.84	6.73	32.71	739
	2H	4.24	17.57	55.25	8.48	42.00	1,176
	3H	3.92	13.78	46.59	13.94	53.75	1,186
	4H	5.56	14.90	33.00	4.52	20.60	615
	5H	4.52	14.83	48.95	6.63	27.92	1,164
	6H	2.95	10.02	29.49	4.43	19.86	780
HvDRR27	Genome-wide	32.76	100.27	318.95	73.75	311.16	7,687
	Chromosomes						
	1H	4.31	12.96	39.35	10.52	41.54	946
	2H	5.96	17.38	55.06	14.99	68.24	1,370
	3H	4.16	16.20	53.39	15.21	57.47	1,363
	4H	3.21	7.65	28.26	4.69	21.69	822
	5H	6.22	18.39	51.08	5.19	18.55	1,091
	6H	3.89	12.76	40.02	9.49	38.15	943
HvDRR28	Genome-wide	26.40	64.35	205.99	49.13	216.64	5,292
	Chromosomes						
	1H	2.99	7.47	22.22	6.12	28.47	659
	2H	4.45	12.13	36.19	9.19	44.36	907
	3H	4.63	10.23	34.51	9.79	33.73	850
	4H	2.69	6.32	18.56	3.91	14.63	582
	5H	3.73	8.18	32.28	5.87	26.47	916
	6H	2.77	8.47	24.63	4.44	17.27	607
HvDRR28	7H	5.14	11.54	37.59	9.81	51.71	773

Table S6: The central tendency measures for crossover (CO) intervals' length classified by related marker block size (>10 kb, >500 kb, and >3 Mbp) on average per recombinant inbred lines (RIL) of the analyzed double round-robin (DRR) populations HvDRR13, HvDRR27, and HvDRR28.

CO layer	median	mean	sd
>3 Mbp	151,632	695,585	1,455,871
>500 kbp	145,897	619,208	1,279,607
>10 kbp	41,959	345,952	893,055

Table S7: The number of 10 kb windows corresponding to recombination coldspots and hotspots, and the number of coldspot regions in the barley chromosomes in the analyzed double round-robin (DRR) populations HvDRR13, HvDRR27, and HvDRR28.

Recombination region	Chromosome	HvDRR13	HvDRR27	HvDRR28
Hotspot windows	1H	45	40	27
	2H	82	65	73
	3H	83	73	80
	4H	42	26	53
	5H	53	56	76
	6H	41	36	45
	7H	57	42	80
	Genome-wide	403	338	434
Coldspot windows	1H	12355	10351	10847
	2H	9574	7387	11192
	3H	9099	9468	11388
	4H	14549	14213	15751
	5H	8254	9880	10340
	6H	8831	13829	9362
	7H	7051	8639	6804
	Genome-wide	69713	73767	75684
Coldspot regions	1H	179	329	310
	2H	313	484	385
	3H	335	420	389
	4H	207	341	287
	5H	357	437	404
	6H	224	379	204
	7H	280	353	313
	Genome-wide	1895	2743	2292
Average length (kb)		367	269	330
Length range (kb)		10–17,580	10–12,200	10–12,010

Table S8: The number of 10 kb windows corresponding to the total, coldspots, and hotspots in the pericentromeric, distal proximal, and distal telomeric regions of the chromosomes in the analyzed double round-robin (DRR) populations HvDRR13, HvDRR27, and HvDRR28.

Recombination region	Chromosome region	Population		
		HvDRR13	HvDRR27	HvDRR28
Total windows	Pericentromeric	248993	220993	242993
	Distal proximal	63990	74490	66240
	Distal telomeric	63988	74488	66238
Coldspot windows	Distal proximal	39915	41844	43047
	Distal telomeric	29798	31923	32637
Hotspot windows	Pericentromeric	66	81	100
	Distal proximal	85	67	71
	Distal telomeric	252	190	263

Table S9: The observed overlaps among the gene conversion (GC) hotspot and crossover (CO) coldspot and hotspot 10 kb windows in the distal region of the barley chromosomes, and their statistical comparison with the overlaps generated under a random distribution of such regions in the analyzed double round-robin (DRR) populations HvDRR13, HvDRR27, and HvDRR28. The random distribution of the genomic regions was simulated with 1,000 permutations.

Recombination region		Population	Observed overlaps	Random overlaps
GC hotspot windows	CO coldspot windows	HvDRR13	91	443
		HvDRR27	71	323
		HvDRR28	138	483
	CO hotspot windows	HvDRR13	54***	3
		HvDRR27	38***	1
		HvDRR28	60***	3

*** $P < 0.001$ for $H_0: \mu_{Observed} \geq \mu_{Random}$; Z-test.

Table S10: The proportion of gene conversion (GC) hotspots of 10 kb windows overlapping with crossover (CO) hotspots, their neighboring windows at 10 and 20 kb away, CO hotspots in the two other studied populations, and the rest of the windows in the pericentromeric and the distal regions of the chromosomes in the analyzed double round-robin (DRR) populations HvDRR13, HvDRR27, and HvDRR28. For a given population, the significant difference ($\alpha = 0.0083$) in the Chi-squared test among the overlap of two given window regions with the GC hotspot windows are indicated with different letters.

Window region	Population		
	HvDRR13	HvDRR27	HvDRR28
CO hotspot	0.147 a	0.122 a	0.119 a
CO hotspot +10 kb neighbor	0.019 b	0.019 b	0.025 b
CO hotspot +20 kb neighbor	0.010 b	0.013 b	0.012 b
CO hotspot in the other populations	0.075 c	0.086 a	0.095 a
Pericentromeric	0.000 d	0.000 c	0.000 c
Distal	0.002 e	0.002 d	0.002 d

Table S11: Total structural variants (SV) load fraction of the coldspot 10 kb windows in the distal telomeric region of the barley chromosomes which were below and above the critical value in the comparison of the methylation level between such coldspots and the total genomic windows of the distal proximal region of the chromosomes for the methylated sequence contexts CpG and CHG in the three analyzed populations HvDRR13, HvDRR27, and HvDRR28.

Population	CpG		CHG	
	Below critical	Above critical	Below critical	Above critical
HvDRR13	0.222 ***	0.181	0.229 ***	0.164
HvDRR27	0.235 ***	0.165	0.222 ***	0.158
HvDRR28	0.222 ***	0.178	0.224 ***	0.164

*** $P < 0.001$; Kruskal-Wallis-Test.

Table S12: Methylation level of the coldspot 10 kb genomic windows in the distal proximal region of the barley chromosomes which were below and above the critical value in the comparison of the total structural variants (SV) load fraction between such coldspots and the total windows in their respective subregion in the chromosomes of the three analyzed populations HvDRR13, HvDRR27, and HvDRR28.

Population	CpG		CHG	
	Below critical	Above critical	Below critical	Above critical
HvDRR13	0.898 ***	0.898	0.604 ***	0.602
HvDRR27	0.909 ***	0.902	0.615 ***	0.609
HvDRR28	0.901 ***	0.897	0.600 ***	0.593

*** $P < 0.001$; Kruskal-Wallis-Test.

Table S13: Observed overlaps among the crossover (CO) intervals (related to >3 Mbp marker blocks) and the genomic regions spanned by structural variants (SV) along the distal regions of the barley chromosomes, and their statistical comparison with the overlaps generated under a random distribution of such regions in the double round-robin (DRR) populations HvDRR13 HvDRR27, and HvDRR28. The SVs were classified by type (inversions, deletions, insertions, duplications and translocations). The random distribution of the genomic regions was simulated with 1,000 permutations.

SV type		Population		
		HvDRR13	HvDRR27	HvDRR28
	Total CO intervals	1692	2371	1891
Deletions	Total SVs	773	808	841
	Obs. overlaps	3670***	4663***	4576***
	Perm. overlaps	7705	13911	9616
Duplications	Total SVs	773	808	841
	Obs. overlaps	2249***	2693***	2447***
	Perm. overlaps	2866	4583	3382
Insertions	Total SVs	773	808	841
	Obs. overlaps	1141***	1428***	1287***
	Perm. overlaps	1852	3139	2206
Inversions	Total SVs	773	808	841
	Obs. overlaps	262	257***	297
	Perm. overlaps	282	441	330

*** $P < 0.001$ for $H_0: \mu_{Observed} \geq \mu_{Random}$; Z-test.

Table S14: Physical distance between crossover (CO) breakpoints and their closest structural variant (SV) compared with the same distance for simulated COs across the genome of the three analyzed populations HvDRR13, HvDRR27, and HvDRR28. The SVs were classified by type (inversions, deletions, insertions, duplications, and translocations) and size (50—299 bp, 0.3—4.9 kb, 5—49 kb, 50—249 kb, 0.25—1 Mbp, and >1 Mbp).

SV type	SV size	Observed distance			Simulated random distance		
		mean	median	sd	mean	median	sd
Inversions	50—299 bp	17,099,530***	6,671,607	27,149,096	7,555,888	4,082,929	9,701,340
	0.3—4.9 kb	17,214,528***	7,919,230	25,013,784	8,350,472	4,881,894	9,698,011
	5—49 kb	13,582,916***	7,082,586	17,761,643	7,568,365	4,025,608	10,906,124
	50—249 kb	15,582,683***	5,935,824	23,080,540	7,183,728	2,968,177	11,953,339
	0.25—1 Mbp	30,009,390***	13,862,023	51,763,750	22,406,932	7,977,245	52,897,364
Deletions	50—299 bp	544,971***	137,030	1,205,795	232,554	83,471	478,947
	0.3—4.9 kb	804,234***	189,082	1,754,699	306,130	101,326	677,696
	5—49 kb	1,231,131***	190,110	3,546,213	337,269	107,321	996,528
	50—249 kb	12,709,060***	3,625,545	25,218,392	3,985,361	1,875,750	6,368,258
Insertions	50—299 bp	1,046,414***	206,000	2,882,087	358,385	140,706	826,689
	0.3—4.9 kb	6,830,746***	1,578,779	16,618,799	1,790,073	913,399	3,024,435
Duplications	50—299 bp	1,545,476***	644,083	2,505,797	754,287	391,311	1,138,254
	0.3—4.9 kb	1,495,159***	646,717	2,393,655	793,538	395,685	1,227,684
	5—49 kb	663,015***	258,878	1,235,982	354,688	186,860	537,661
	50—249 kb	5,279,236***	2,299,383	7,738,683	2,443,219	1,339,239	3,225,454
	0.25—1 Mbp	57,759,853***	22,418,461	92,387,936	51,499,104	18,058,220	95,936,696
Translocations	NA	94,117***	44,497	137,792	75,258	39,326	109,027

*** $P < 0.001$ for $H_0: \mu_{Observed} = \mu_{Random}$; Mann–Whitney U test.

Table S15: Proportion of either hotspot or coldspot 10 kb genomic windows in the averaged neighboring windows located 10 kb upstream or downstream of the hotspot and coldspot genomic regions in the distal proximal and telomeric subregions of the barley chromosomes in the three analyzed populations HvDRR13, HvDRR27, and HvDRR28.

Population	Window type	Distal subregion	Hotspot	Coldspot
HvDRR13	Coldspot	Proximal	0.04	0.00
		Telomeric	0.05	0.00
	Hotspot	Proximal	0.00	0.49
		Telomeric	0.00	0.34
HvDRR27	Coldspot	Proximal	0.02	0.00
		Telomeric	0.02	0.00
	Hotspot	Proximal	0.00	0.37
		Telomeric	0.00	0.33
HvDRR28	Coldspot	Proximal	0.04	0.00
		Telomeric	0.05	0.00
	Hotspot	Proximal	0.00	0.58
		Telomeric	0.00	0.45

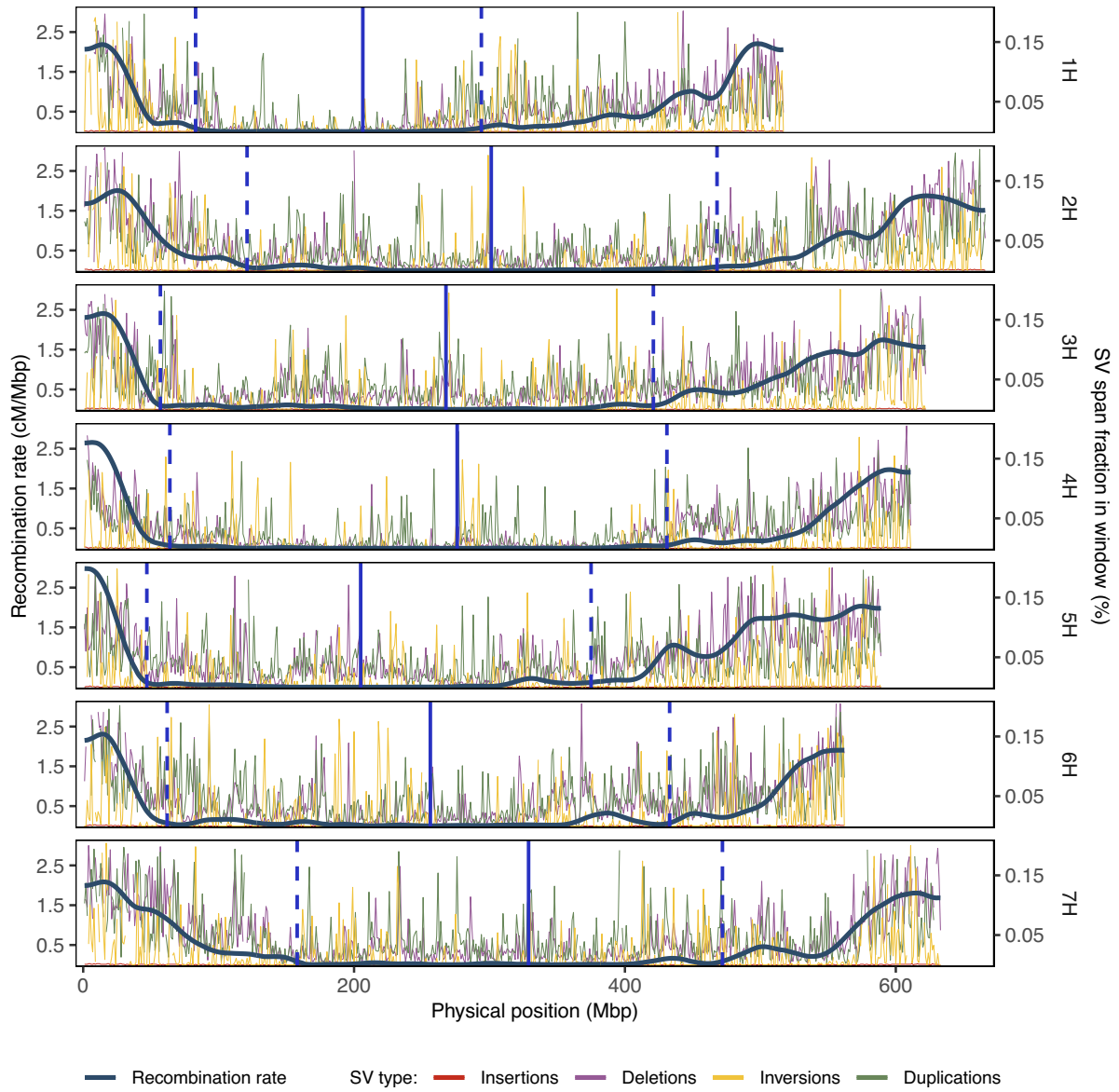


Fig. S1: Distribution of recombination rate and structural variants (SV): insertions, deletions, inversions, and duplications, across the seven barley chromosomes. The displayed values represent the spanned fraction of a 1 Mbp genomic window by the sum of all SVs of the same kind. The values were averaged across the 45 double round-robin (DRR) populations in 1 Mbp windows across the seven barley chromosomes. The vertical blue solid line indicates the position of the centromere in the Morex v3 reference genome and the vertical blue dashed lines indicate the pericentromeric region calculated across the 45 DRR populations.

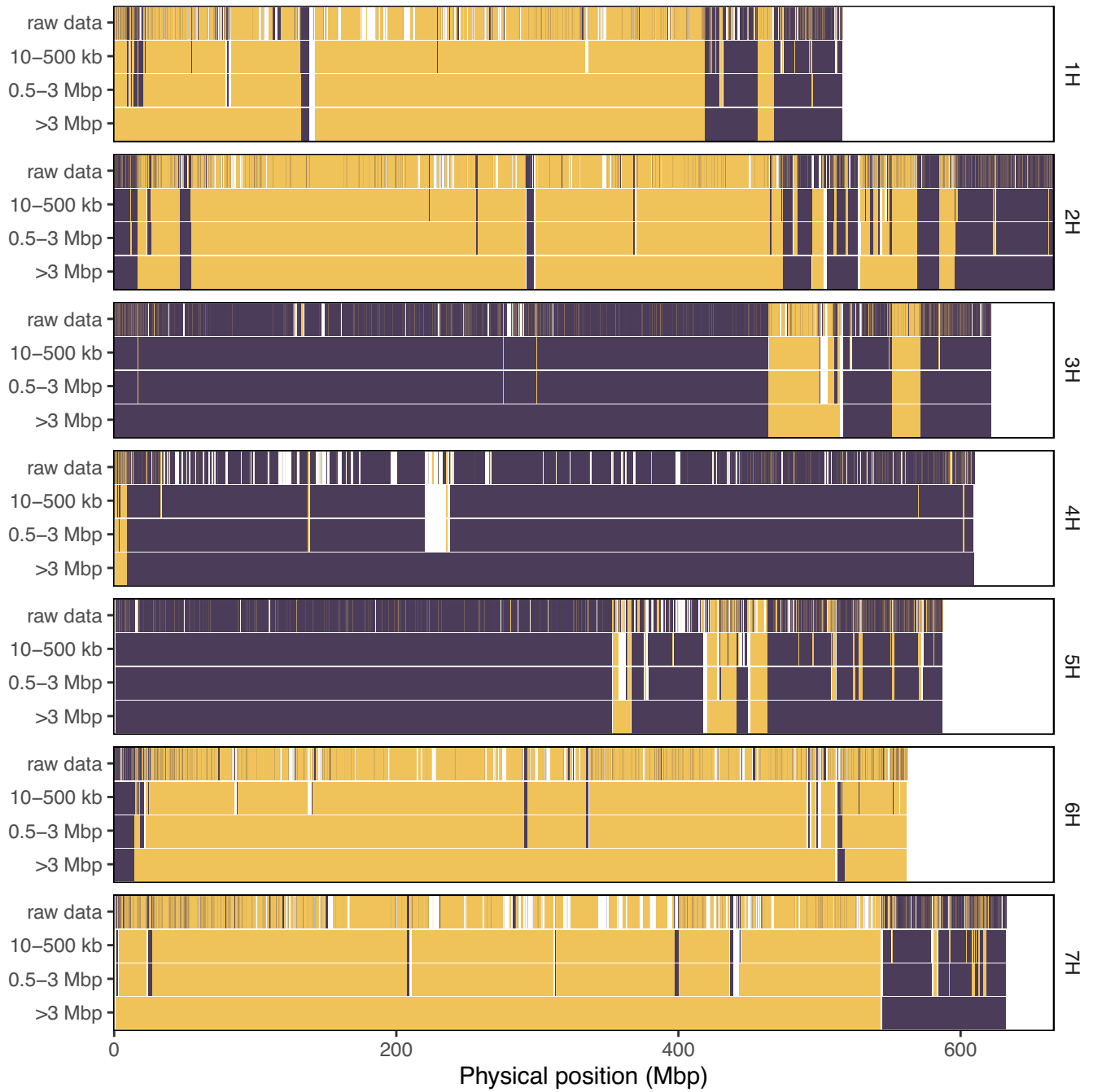
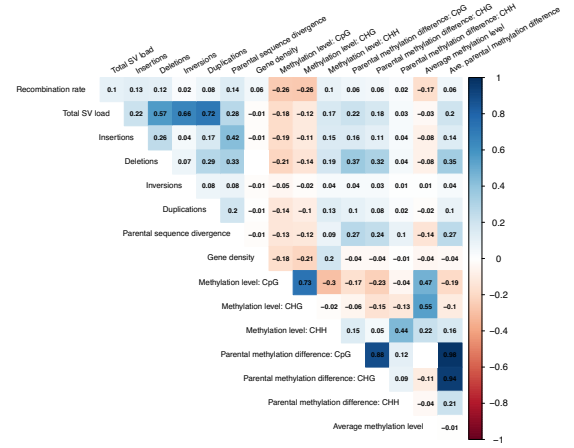
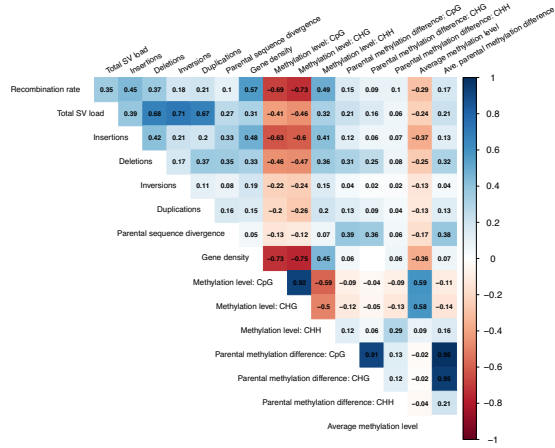


Fig. S2: Example of the graphical genotype for one recombinant inbred line (RIL). The four graphical genotypes for each of the seven barley chromosomes show, from the raw data, the different three crossover (CO) related parental marker block categories considered in this study: 10-500 kb, 0.5-3 Mbp, and >3 Mbp, respectively.



[A]

[B]

Fig. S3: Correlation matrix between recombination rate, the total structural variants (SV) load and such of the different types –insertions, deletions, inversions, and duplications– separately, the parental sequence divergence, gene density, methylation level and the parental difference in methylation at the sequence contexts CpG, CHG, CHH, and their average, across 1 Mbp windows in the distal (A) and the pericentromeric (B) regions of the barley chromosomes for the average among the 45 double round-robin (DRR) populations. Pearson's correlation coefficients are displayed and indicated with a color gradient from -1 (red) to 1 (blue).

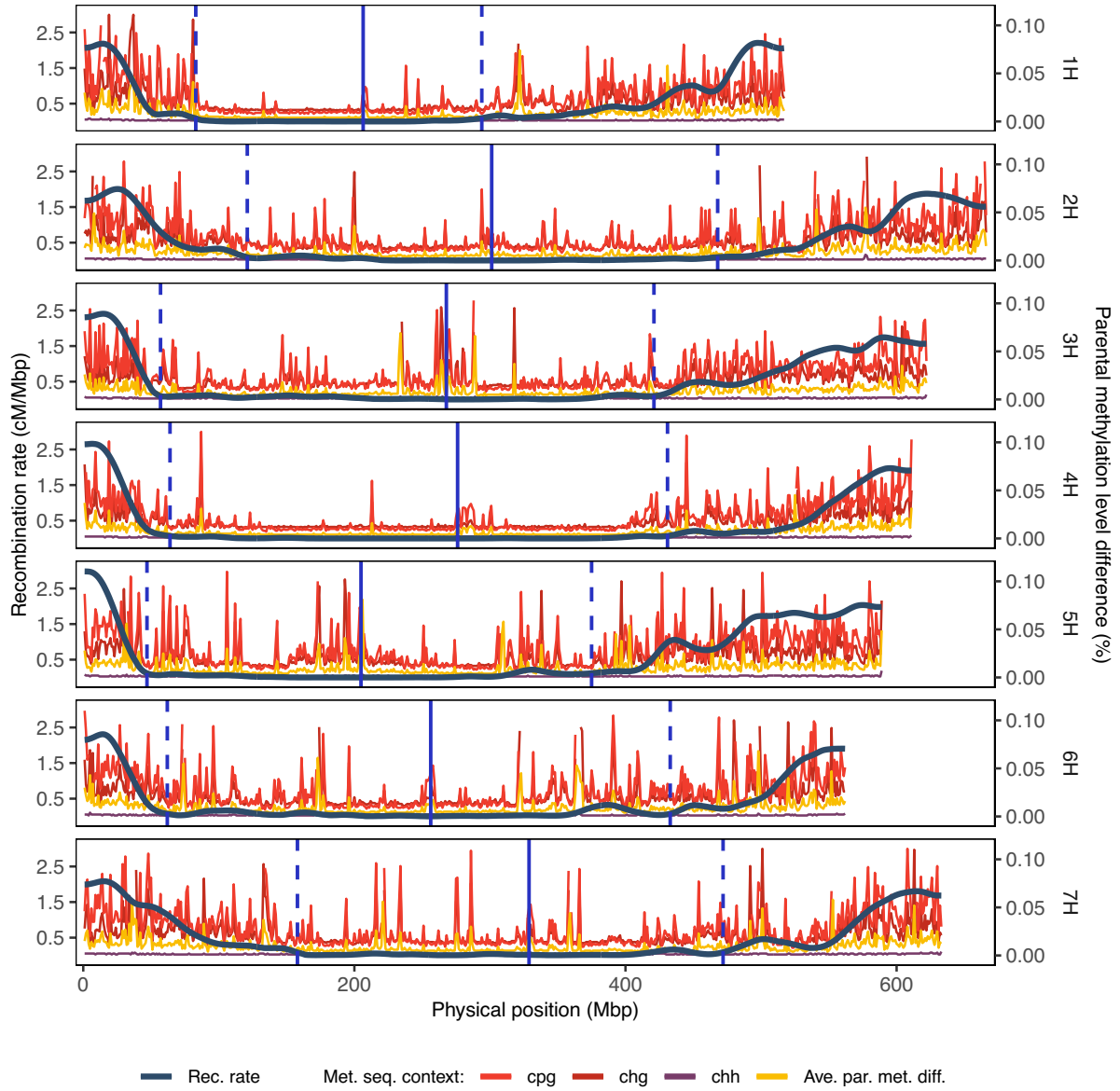


Fig. S4: Distribution of recombination rate and the parental methylation level difference in 1 Mbp windows at the methylated sequence contexts CpG, CHG, and CHH on average across the 45 DRR populations across the seven barley chromosomes. The parental methylation level difference values for a given methylation context represent the difference among the parental inbred lines of a given population on the percentage of methylated reads of such context present in a 1 Mbp window averaged among populations. The average among the three sequence contexts (Ave. par. met. level) is also displayed. The vertical blue solid line indicates the position of the centromere in the Morex v3 reference genome and the vertical blue dashed lines indicate the pericentromeric region calculated across the 45 DRR populations.

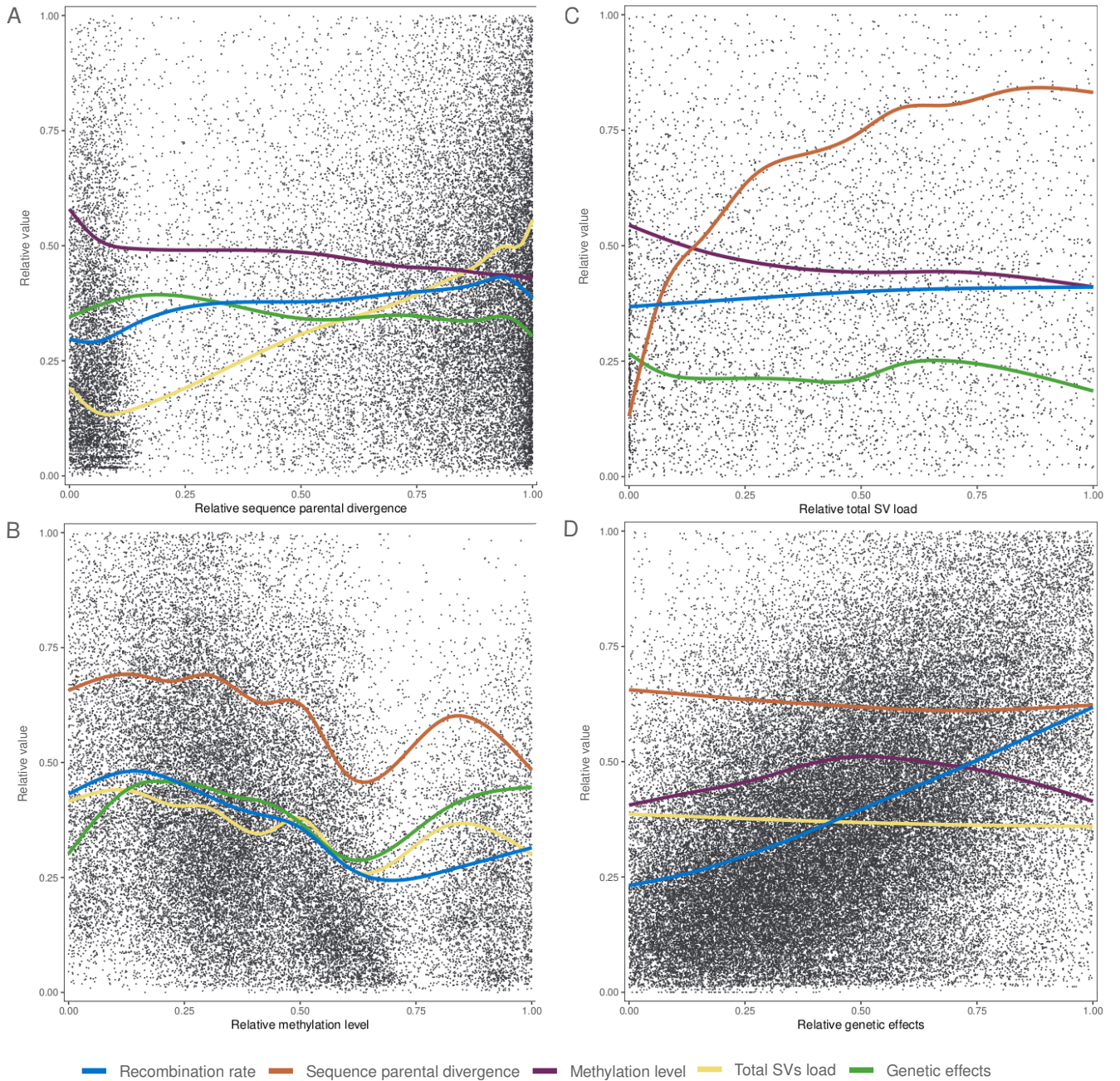


Fig. S5: Relative value of genomic features in the 1 Mbp genomic windows in which the explanatory feature was found significantly correlated with the variation in the recombination rate among the 45 double round-robin (DRR) populations. The significance of the correlation was determined by a stepwise regression procedure. The studied genomic features included the recombination rate (lightblue), sequence divergence among parental inbred lines (orange), genetic effects (green), total structural variants (SV) load (yellow), and average methylation level across the sequence contexts CpG, CHG, and CHH (purple). The explanatory features are (A) sequence parental divergence, (B) total SV load, (C) average methylation level, and (D) genetic effects. The dots correspond to the relative value of the recombination rate.

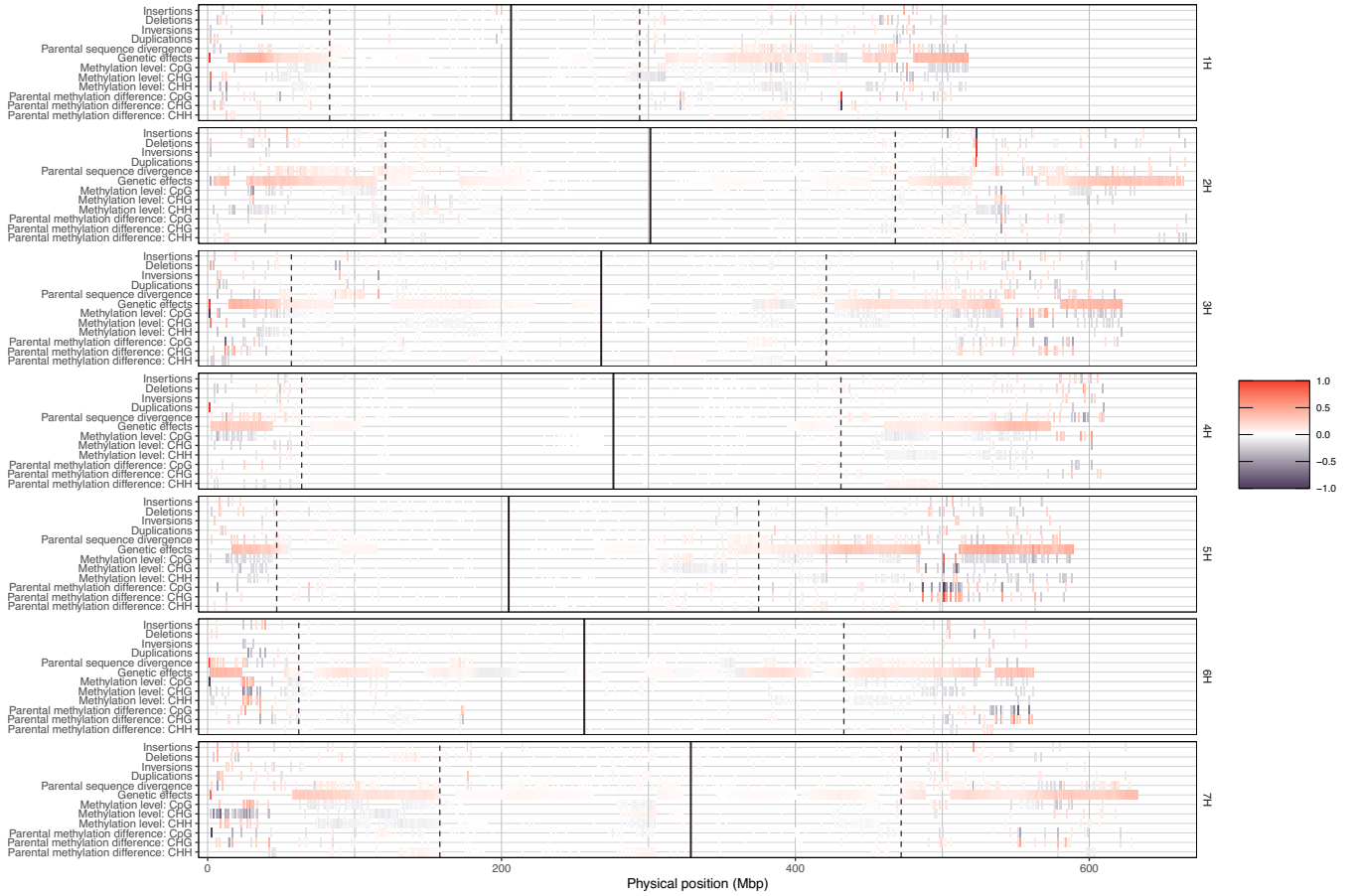


Fig. S6: Summary genomic features explaining the variation in the recombination rate among the 45 double round-robin (DRR) populations in 1 Mbp windows across the seven barley chromosomes. The significance of the correlation was determined by a stepwise regression procedure. The standardized regression coefficients are indicated by a color gradient from -1 (purple) to 1 (red). The studied genomic features included sequence divergence among parental inbred lines, genetic effects, span fraction of structural variants (SV) -such as insertions, deletions, inversions, and duplications-, methylation level -for the sequence contexts CpG, CHG, and CHH-, and the difference in the methylation level among the parental inbred lines for such sequence contexts. The vertical solid line indicates the position of the centromere in the reference genome, and the vertical dashed lines indicate the pericentromeric region calculated across the 45 DRR populations.

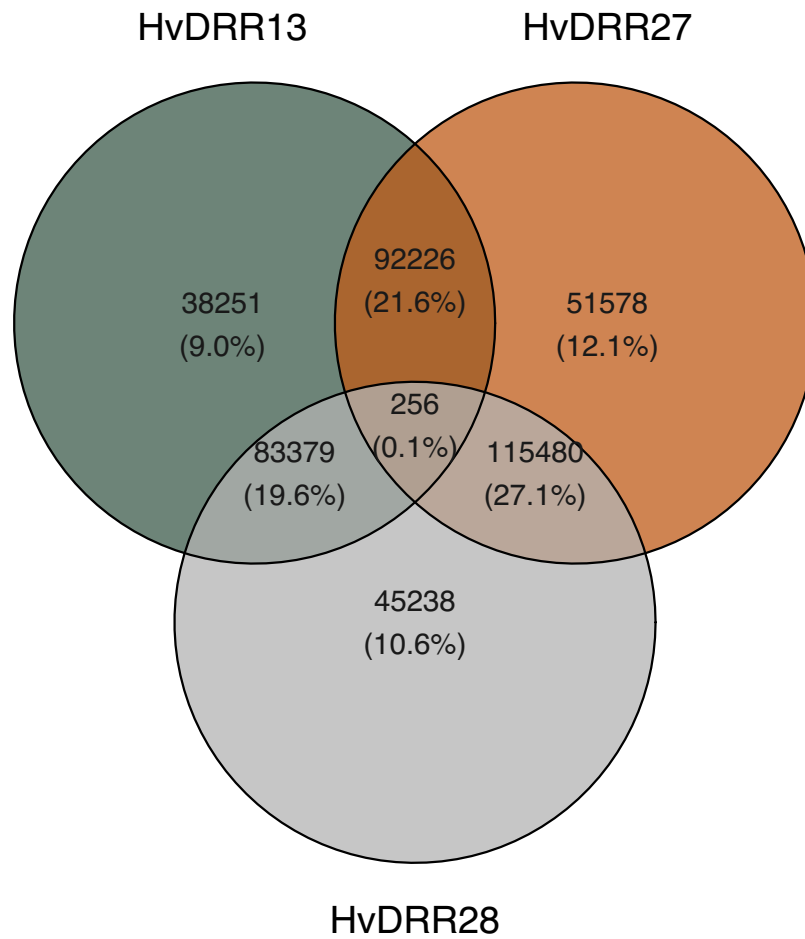


Fig. S7: Venn diagram for the number of polymorphic single nucleotide polymorphism (SNP) markers detected by RNAseq in the analyzed double round-robin (DRR) populations HvDRR13, HvDRR27, and HvDRR28.

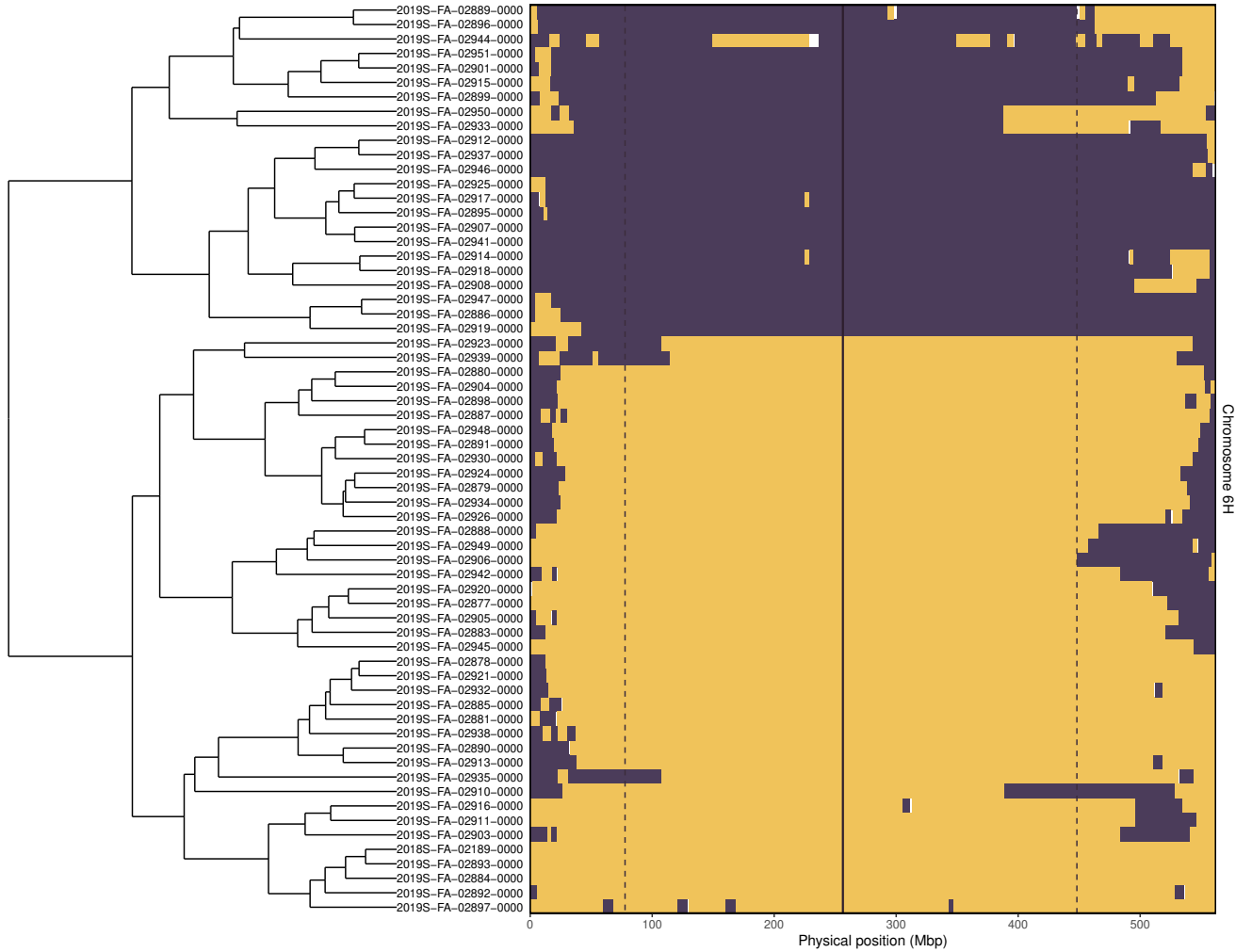


Fig. S8: Stacked graphical genotypes of the recombinant inbred lines (RILs) for chromosome 6H in the HvDRR13 population. Different colors indicate parental allelic blocks inherited from different parents. The RILs are ordered according to their hierarchical complete clustering based on Euclidean distances of recombination rates for this chromosome. The vertical solid line in the background indicates the position of the centromere in the Morex v3 reference genome and the vertical dashed lines indicate the pericentromeric region calculated across the 45 DRR populations.

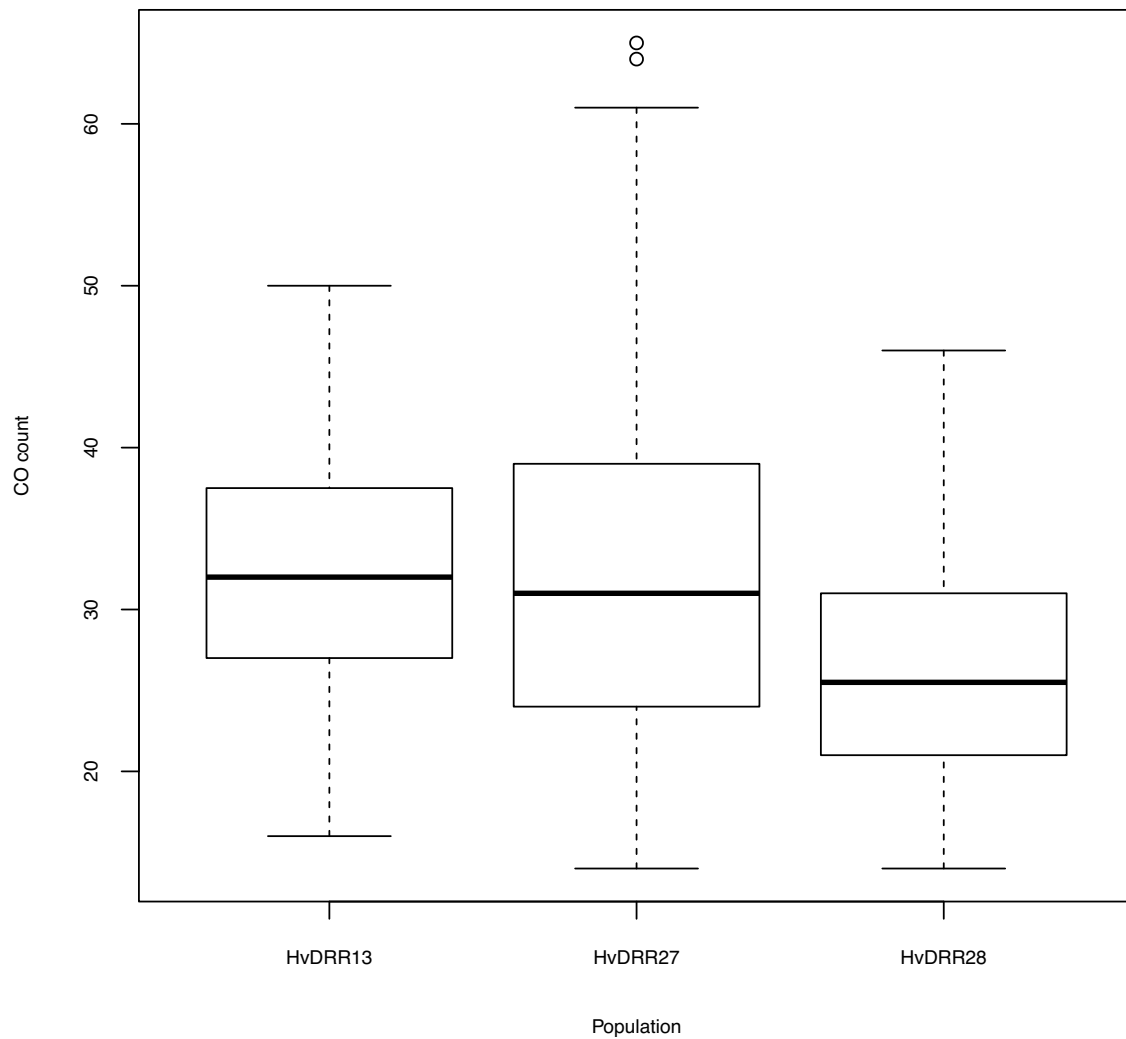


Fig. S9: Box-plot of the number of counted genome-wide crossovers (CO) with >3 Mbp marker blocks in the analyzed double round-robin (DRR) populations HvDRR13, HvDRR27, and HvDRR28.

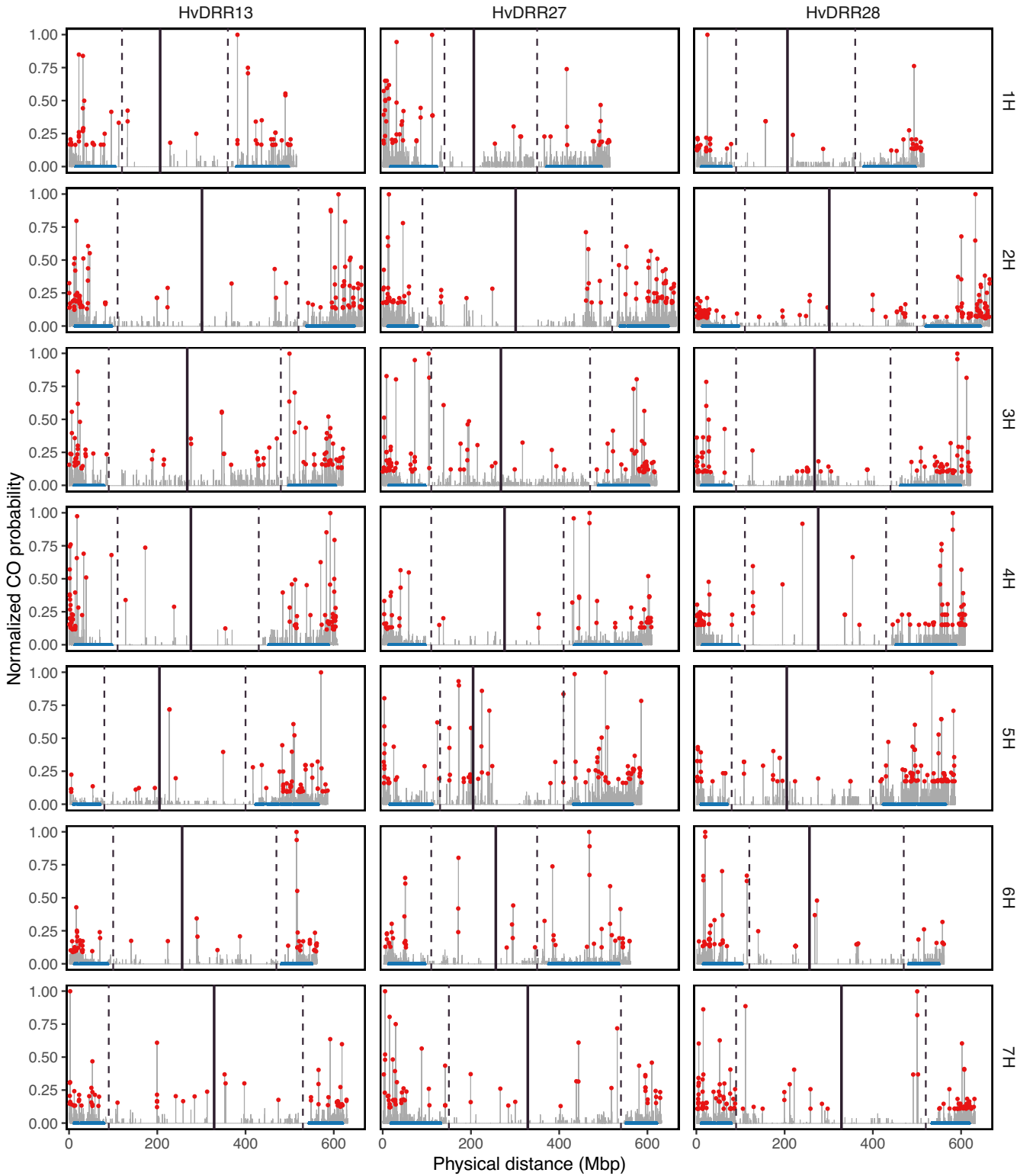


Fig. S10: Normalized accumulated crossover (CO) probability (grey line), hotspot (red dots), and coldspot (turquoise triangles) 10 kb genomic windows across the seven barley chromosomes in the analyzed double round-robin (DRR) populations HvDDR13, HvDDR27, and HvDDR28. The vertical solid line in the background indicates the position of the centromere in the v3 Morex reference genome and the vertical dashed lines indicate the pericentromeric region calculated for each chromosome and population.

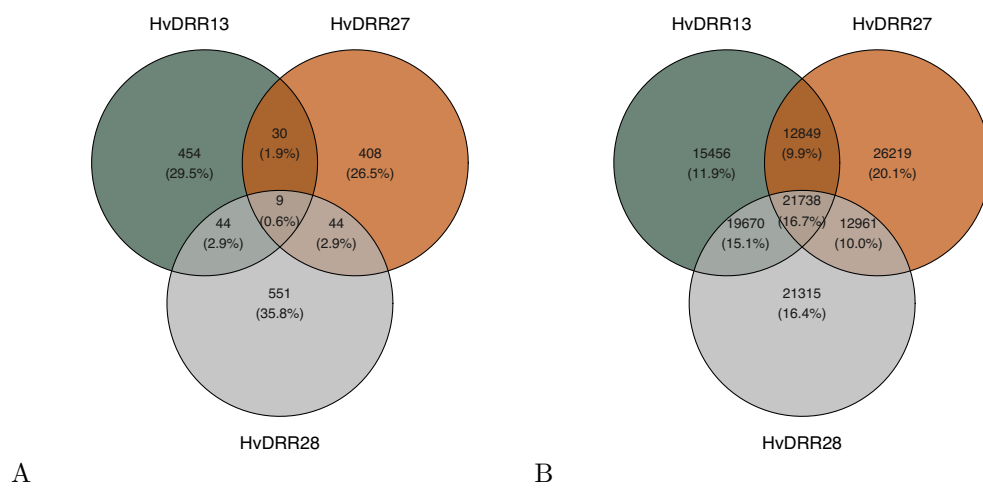


Fig. S11: Venn diagrams for the number of crossocers (CO) hotspot (A) and coldspot (B) 10 kb windows in the analyzed double round-robin (DRR) populations HvDRR13, HvDRR27, and HvDRR28.

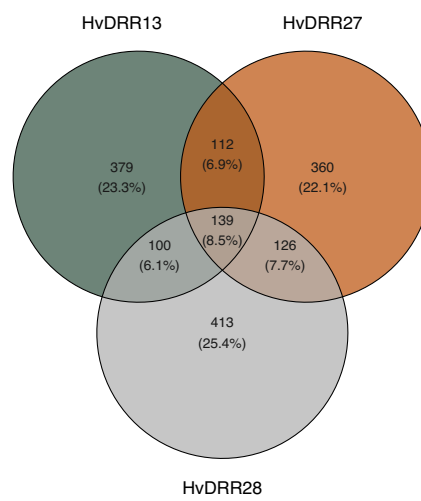


Fig. S12: Venn diagrams for the number of gene conversion (GC) hotspot 10 kb windows in the analyzed double round-robin (DRR) populations HvDRR13, HvDRR27, and HvDRR28.

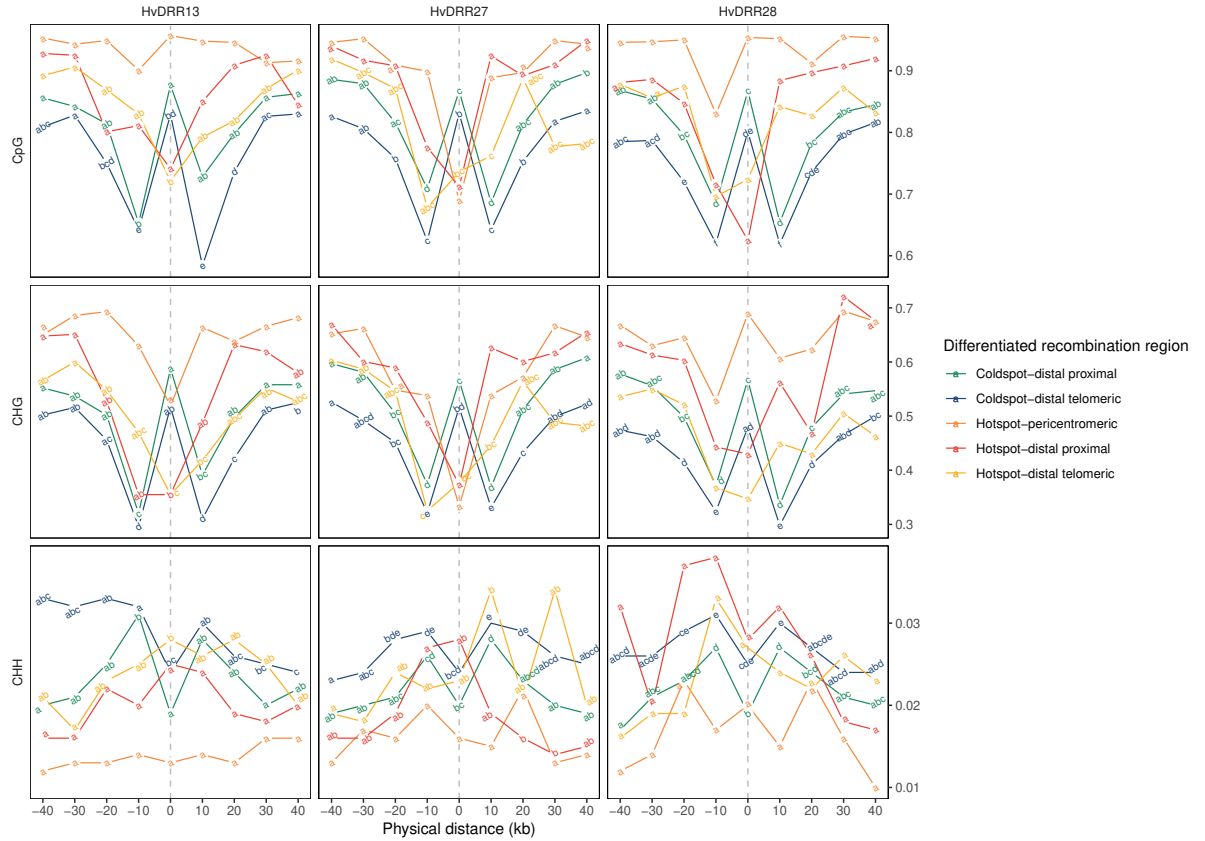


Fig. S13: The mean values of the methylation level at the methylated sequence contexts CpG, CHG, and CHH in the genomic windows grouped by the physical positions in the range from -40 kb to +40 kb around the coldspot and hotspot genomic regions in the pericentromeric, distal proximal, and distal telomeric region of the chromosomes in the analyzed double round-robin (DRR) populations HvDRR13, HvDRR27, and HvDRR28. The vertical dashed lines indicate the relative location of either the coldspot or the hotspot windows, respectively. The significant difference ($\alpha = 0.001$) in the Wilcoxon's rank sum test among the windows groups corresponding to the different 10 kb physical position neighboring the respective hotspot or coldspot region in a particular chromosome region is indicated with different letters.

5 Accurate recombination estimation from pooled genotyping and sequencing: a case study on barley

Authors:

Michael Schneider, **Federico Casale**, and Benjamin Stich.

Own contribution: Second author. I performed part of the data analyses and corrected the manuscript.

RESEARCH

Open Access



Accurate recombination estimation from pooled genotyping and sequencing: a case study on barley

Michael Schneider¹, Federico Casale¹ and Benjamin Stich^{1,2,3*}**Abstract**

Sexual reproduction involves meiotic recombination and the creation of crossing over between homologous chromosomes, which leads to new allele combinations. We present a new approach that uses the allele frequency differences and the physical distance of neighboring polymorphisms to estimate the recombination rate from pool genotyping or sequencing. This allows a considerable cost reduction compared to conventional mapping based on genotyping or sequencing data of single individuals. We evaluated the approach based on computer simulations at various genotyping depths and population sizes as well as applied it to experimental data of 45 barley populations, comprising 4182 RIL. High correlations between the recombination rates from this new pool genetic mapping approach and conventional mapping in simulated and experimental barley populations were observed. The proposed method therefore provides a reliable genetic map position and recombination rate estimation in defined genomic windows.

Keywords: Recombination rate, Pool sequencing, Population genetics, Genetic map, Breeding value estimation

Introduction

Sexual reproduction involves meiotic recombination and the creation of crossing over between homologous chromosomes, which leads to new allele combinations [1]. The resulting phenotypic diversity is the basis of evolution and human selection [2]. Meiotic recombination is therefore essential in various research fields such as medicine, animal and plant breeding, conservational and evolutionary genomics [2–8]. Especially in breeding, the response to selection is strongly associated with the recombination rate. Therefore, increased recombination can enhance breeding and selection efficiency [9]. Besides, a high recombination rate could foster the

dissociating of phenotypic and genetic variation [10] and affect reproductive barriers.

The exchange mentioned above between homologous chromosomes was first reported by T.H. Morgan, who identified novel allele combinations after crossing two *Drosophila melanogaster* strains [11–13]. Since then, incredible progress has been made in uncovering the molecular mechanisms of meiotic recombination [14, 15]. Furthermore, interest increases in understanding the effect of environmental factors on the recombination rate (RR) or the inter- and intraspecies variation of RR (e.g. [15–18]).

Detecting differences in RR among environmental conditions, genetic backgrounds, or species requires the genotypic characterization of a representative number of genotypes of each treatment. The most frequently applied genotyping approach in this context is using SNP arrays. However, the main limitation of such approaches is that the costs increase linearly with the number of evaluated

*Correspondence: benjamin.stich@hhu.de

³ Cluster of Excellence on Plant Sciences, From Complex Traits Towards Synthetic Modules, Universitätsstraße 1, 40225 Düsseldorf, Germany
Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

genotypes. Furthermore, the number of loci typically genotyped with SNP arrays is limited to a few thousand variants [19–23]. This limits the resolution of the resulting genetic map, which hinders, e.g., studies on populations with a long history of natural or artificial selection [24]. Sequencing strategies like genotyping by sequencing [25, 26], exome capture [27, 28], whole-genome resequencing [29, 30], or RNA sequencing [31, 32] are useful to increase the genome-wide variant density and coverage. However, such approaches applied to individual genotypes have the same limitations as mentioned above for SNP array genotyping – the costs increase linearly with the number of studied genotypes.

The progress of sequencing techniques allowed the estimation of recombination events from linked read gamete sequencing [33]. Although this approach revealed promising results, the high experimental effort and associated costs might prevent its implementation in extensive recombination screening studies.

Our study proposes an alternative approach to overcome the burden of either high costs, low variant densities, or low genotype count. The proposed method allows the estimation of the RR from pooled genotype samples. In this situation, any user-defined quantity of genotypes can be pooled without increasing the monetary costs of genotyping or sequencing. Our approach uses the allele frequency differences and the physical distance of neighboring polymorphisms to estimate the RR, an idea initially proposed for situations with a linked locus under selection that causes a fitness differential [34].

The objectives of our study were

- i. to assess the accuracy of estimated genetic maps and RR from pool genotyping based on computer simulations,
- ii. describe a best practice guideline for accurate RR extraction from pool genotyping and sequencing, and
- iii. apply the RR estimation on experimental populations of barley

Results

Raw pool genetic map (PGM) calculation from simulated populations

We simulated 1260 F_2 populations with various genotyping depths and population sizes. The simulations were performed based on a consensus genetic map calculated for 4182 recombinant inbred lines from 45 barley HvDRR populations [18].

The genome-wide SNP allele frequency observed in the simulated populations deviated from the expected 0.5 (Supplementary Figure 1). The average deviation

was highest in small populations (50 genotypes – 0.04, standard deviation = 0.03). It decreased exponentially to a genome-wide average of 0.003 (sd = 0.002) for the populations consisting of 10,000 genotypes.

Based on the allele frequency deviation of pairs of physically neighboring SNPs and their physical distance, we estimated the raw pool genetic map (PGM) and calculated the PGM recombination rate (RR_{PGM}) for 50 MB windows across the genome (Fig. 1). The average correlation coefficient of the RR derived from the consensus genetic map ($RR_{\text{consensus}}$) and the RR_{PGM} was $r = 0.894$ across all genotyping depths. The lowest correlation was observed when only 500 markers were used for genotyping the population ($r = 0.819$, Table 1). Generally, a continuous increase in the correlation between RR_{PGM} to $RR_{\text{consensus}}$ was observed with increasing genotyping depth, where a maximum Pearson correlation of 0.994 was observed for a genotyping depth of 42,077 (Table 1). Despite the above described high correlation coefficients between $RR_{\text{consensus}}$ and RR_{PGM} , we observed that the average PGM to consensus genetic map position ratio was 0.0093, indicating a significant underestimation of the overall PGM length and RR_{PGM} (Fig. 2). Additionally, the PGM's standard deviation across all samples was 0.01–1.1 times the average genetic map length ratio. Therefore, we investigated the effect of the genotyping depth and the population size on the length of the PGM and the accuracy of the RR_{PGM} estimation. We observed a shorter PGM in those simulated samples with a low genotyping depth and an almost linear increase in map length with increasing genotyping depth (Fig. 3A). Analogously, the population size influenced the overall extent of the genetic map length and RR_{PGM} . We noticed a decrease in genetic map length with increasing population size (Fig. 3B). In contrast to the genotyping depth, no effect of the population size was observed on the correlation between RR_{PGM} to $RR_{\text{consensus}}$.

In order to obtain a PGM with a length as close as possible to that of the consensus map, correction approaches were investigated. We evaluated the use of two models that included effects for the genotyping depth and population size: a linear and a non-linear model. While the linear model revealed a log-likelihood of 10,332 and an AIC of 20,672, the non-linear model resulted in a log-likelihood and AIC values 30 and 26% lower than that of the linear model, respectively. This was accompanied by a Pearson correlation between the RR_{PGM} and the $RR_{\text{consensus}}$ of 0.635 for the linear model and 0.998 for the non-linear model (Supplementary Figure 2). Therefore, we used the latter to correct the SNP's genetic position on the PGM to a non-linear adjusted pool genetic map position (nPGM).

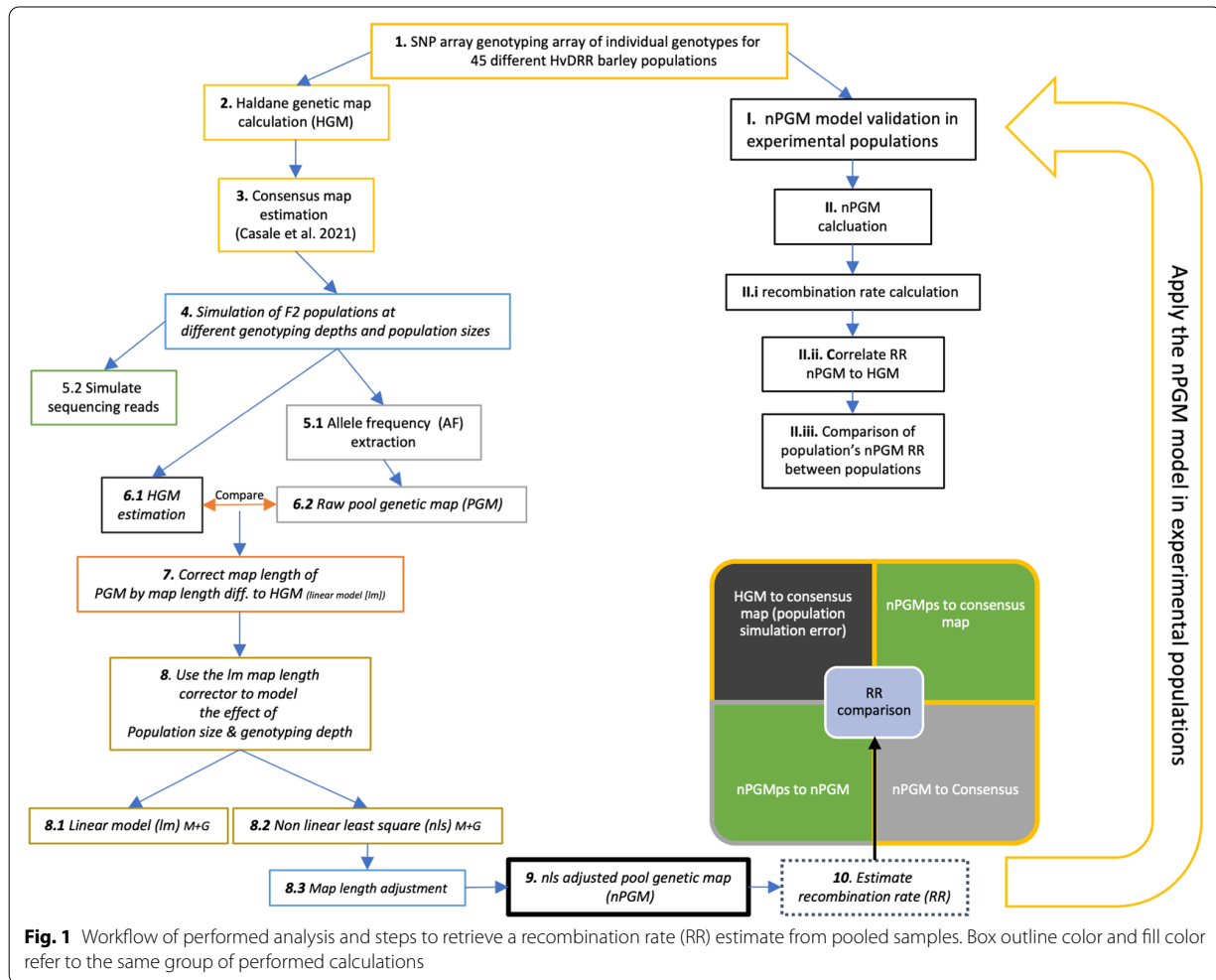


Table 1 Evaluation of the precision and accuracy of the adjusted pool genetic map derived recombination rate (RR_{nPGM}) in comparison to $RR_{consensus}$ on varying levels of the genotyping depth

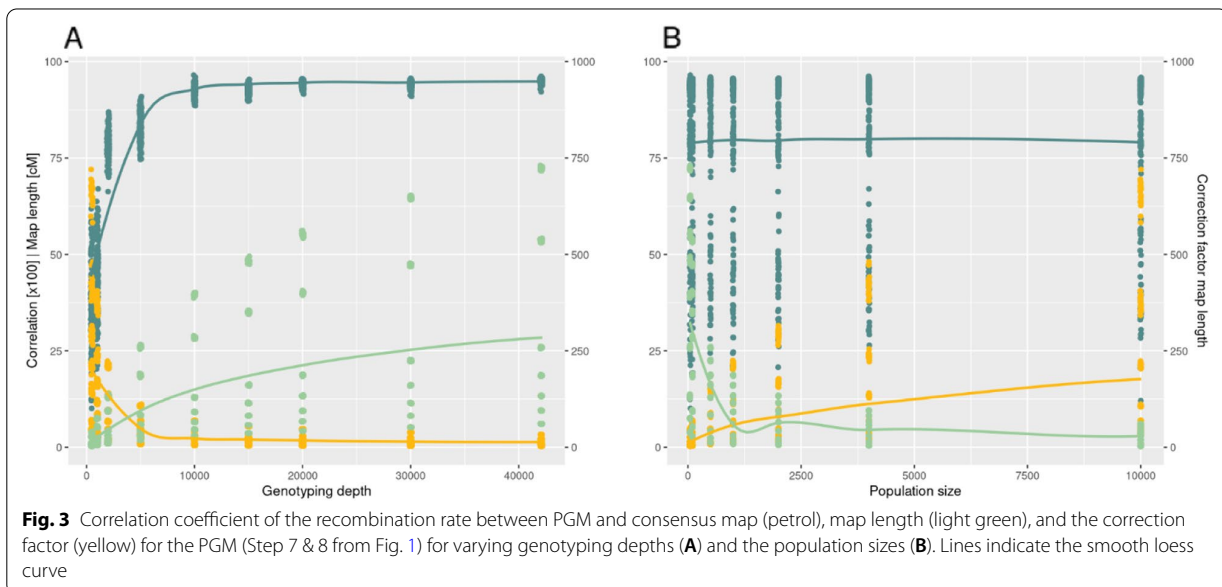
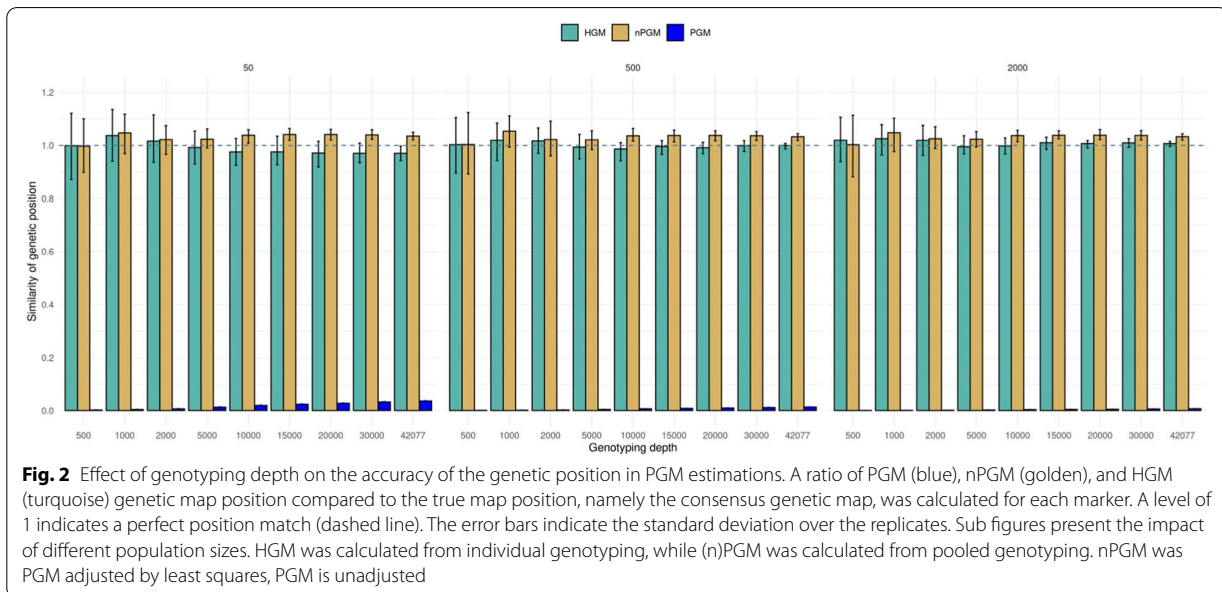
Genotyping depth	RMSE		Pearson correlation	
	average	SD	average	SD
500	0.8597	0.00027	0.387	0.1037
1000	0.4596	0.00029	0.446	0.0939
2000	0.2526	0.00028	0.787	0.0357
5000	0.1168	0.00026	0.835	0.0406
10,000	0.0673	0.00023	0.925	0.0150
15,000	0.0510	0.00020	0.928	0.0120
20,000	0.0405	0.00018	0.941	0.0080
30,000	0.0312	0.00015	0.941	0.0080
42,077	0.0250	0.00014	0.950	0.0070

SD Standard deviation, RMSE Root mean square error

Non-linear adjusted pool genetic map and derived recombination rate

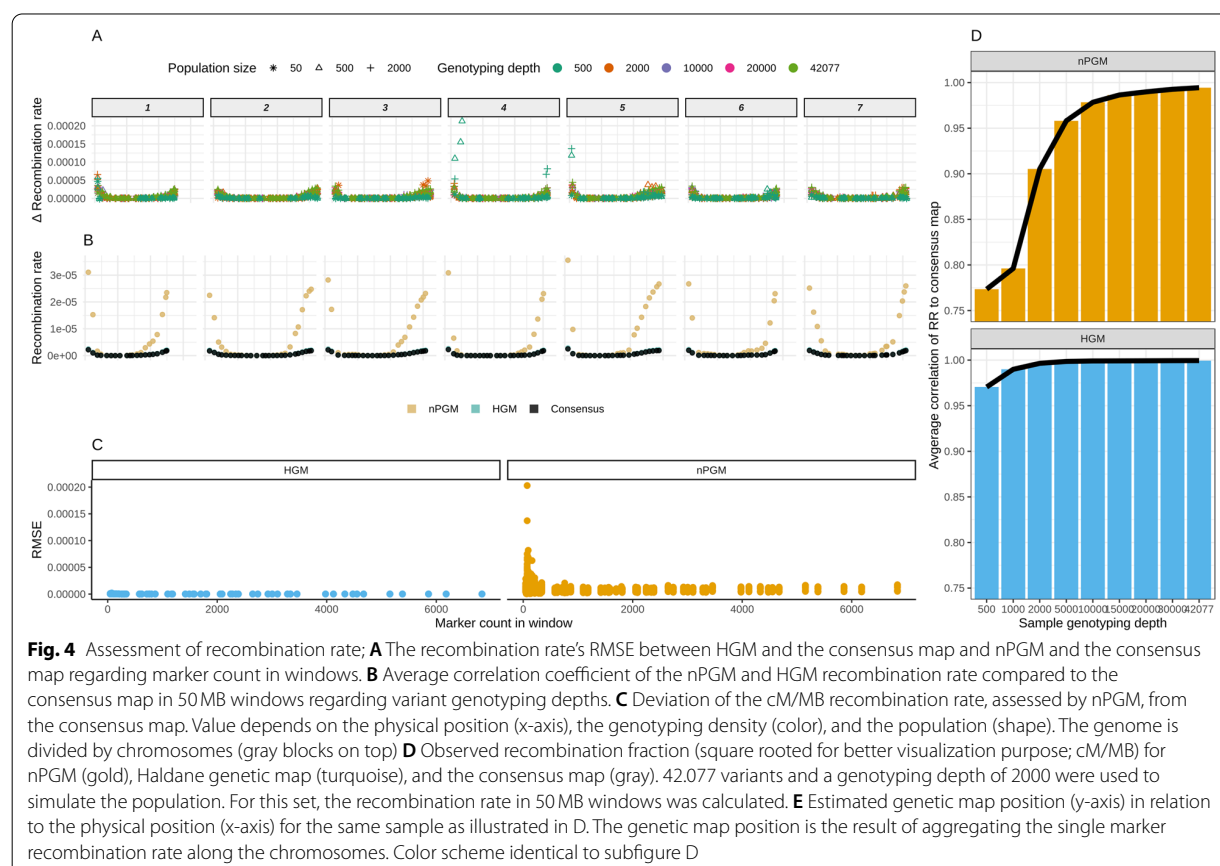
After utilizing the above described non-linear adjustment, the nPGM estimated genetic map positions deviated marginally from the consensus map (Fig. 2). Across all tested samples, each marker's average nPGM to consensus map position ratio was 1.03, which was very close to the ratio of the Haldane genetic map (HGM) to consensus map (ratio = 1.00). HGM is the genetic map recalculated from simulated samples by the Haldane mapping approach. In addition, nPGM resulted in a lower relative standard deviation across all population sizes and genotyping depth than PGM ($sd = 0.014$).

To compare the RR_{HGM} and RR_{nPGM} to the $RR_{consensus}$, we calculated the RR in genomic windows of 50 MB for all replicates of the simulated samples with a population size of 50, 500, and 2000 at all genotyping depths. We observed significant RR correlations between HGM and the consensus map across all tested SNP and genotype



levels (correlation test $p < 2 \times 10^{-16}$), with an overall Pearson correlation of 0.973. The correlation increased to 0.999 when excluding those samples with a genotyping depth below 10,000 markers (Fig. 4B). Similarly, we observed an average Pearson correlation of 0.913 between the RR_{nPGM} and $RR_{consensus}$ for those samples with a genotyping depth $\geq 10,000$ (Fig. 4B). Furthermore, we noted a significant effect of the number of markers in the 50MB windows on the correlation coefficient

and the RMSE in the RR_{HGM} and RR_{nPGM} estimations ($p < 0.0001$, Fig. 4A). The RMSE of nPGM decreased by a factor of four in genomic windows with more than 1000 markers compared to windows with less than 100 markers. In contrast, the RMSE decrease was only 1.17 times for HGM for the same comparison. Analogously, samples characterized by a low genotyping depth resulted in a lower SNP density in genomic windows and, thus, resulted in an increased deviation of RR_{nPGM} (Fig. 4C).



In the last step, we compared the absolute recombination rates on the chromosomal scale among the three approaches (Fig. 4D). The RR was highly similar between HGM and the consensus map throughout the entire genome ($r > 0.98$). In analogy, we observed high similarities in the pericentromeric regions when comparing nPGM and the consensus map. Nevertheless, the non-pericentromeric regions revealed a more pronounced deviation of RR – especially on the long chromosomal arms. However, the Pearson correlation coefficient between the RR_{nPGM} and $RR_{consensus}$ remained high with $r_{non-pericentromeric} = 0.782$ vs. $r_{pericentromeric} = 0.915$ for samples with a genotyping depth $\geq 10,000$ across all chromosomes and replicates. However, the deviations of the RR, estimated from nPGM compared to that of the consensus map, only minorly altered the marker's genetic map position (Fig. 4E).

nPGM estimation in experimental populations

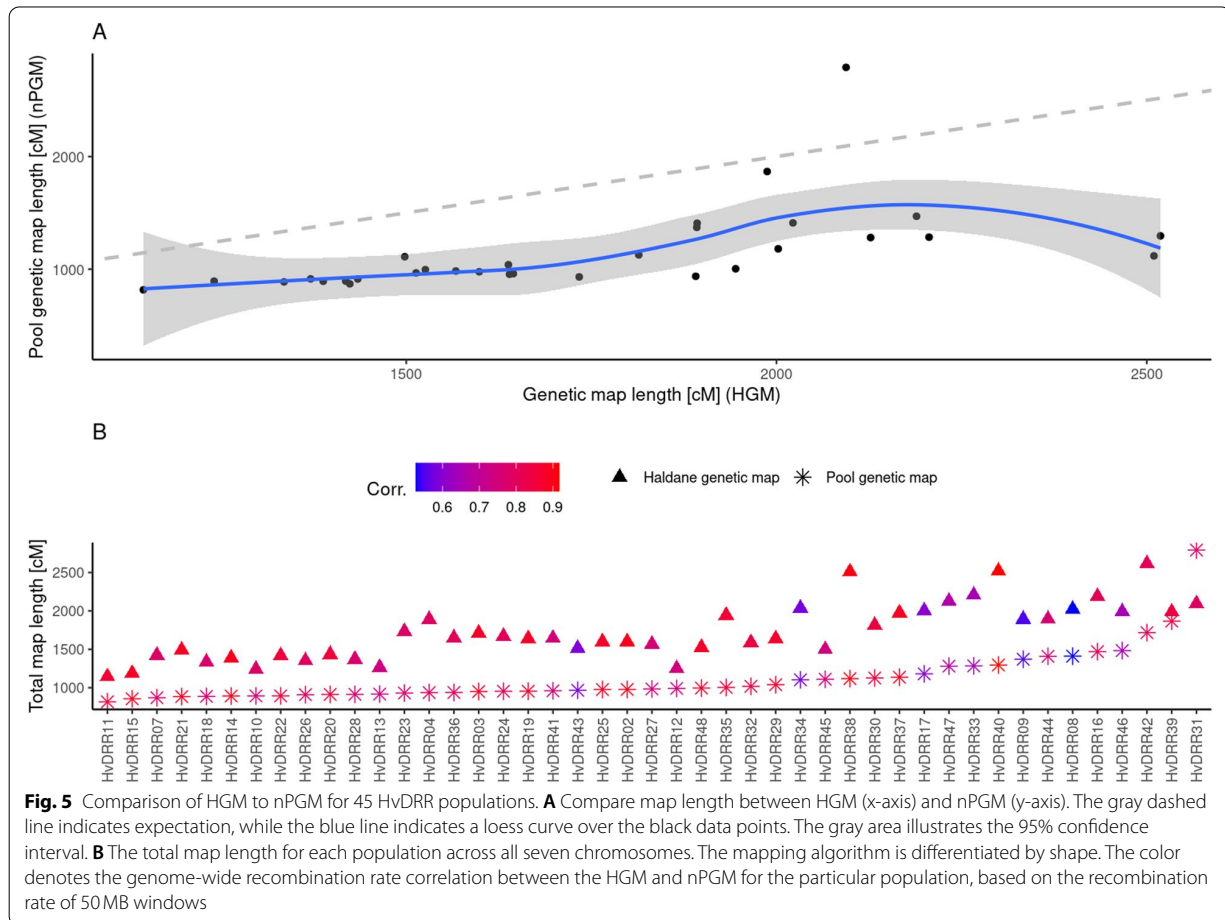
In addition to simulations, we were interested in using the nPGM approach in experimental populations. Therefore, we applied the nPGM strategy to a set of 45 segregating spring barley populations [33], characterized by an

average of 87 recombinant inbred lines (RIL) per population and an average number of 1639 polymorphic SNPs. Pooled genotyping information for all populations was derived from the available genotyping data of individual RIL, and the nPGM and nPGM-derived RR were calculated and compared to HGM-derived values.

Across all 45 populations, an average Pearson correlation of 0.829 was observed between the RR_{HGM} and the RR_{nPGM} in 50MB windows (95% Confidence interval $r = 0.37:0.95$, correlation test $p < 5 \times 10^{-10}$; Fig. 5B).

We observed a similar range of map length across all populations for the nPGM approach (90% confidence interval 873:1670 cM) compared to the HGM (90% confidence interval 1242:2449 cM) (Fig. 5A). Nevertheless, the overall map length was, on average, across all populations, 635 cM longer in HGM than PGM (Fig. 5B). Spearman rank sum correlation between HGM and nPGM revealed a high correlation of 0.83, whereas the Pearson correlation was 0.61.

To evaluate whether the accuracy of the nPGM approach is sufficient to detect differences among the RR_{HGM} and RR_{nPGM} , we used the genome-wide RR_{nPGM} to estimate a general recombination effect (GRE) for each



of the 23 parental inbreds, as was proposed by [18]. This step revealed considerable variations in the GRE among parental inbreds, indicating that some inbreds result in a higher RR_{nPGM} in their progenies than others (Suppl. Figure 4). The direct comparison of the GRE, calculated from RR_{nPGM_i} with the RR_{HGM} GRE from Casale et al. (2021), revealed a rank-sum correlation of 0.877, indicating high similarities (Pearson correlation = 0.803). In the group of the ten genotypes with the highest GRE, nine matched between nPGM and HGM. Similarly, eight of ten genotypes with the lowest GRE were identical between nPGM and HGM.

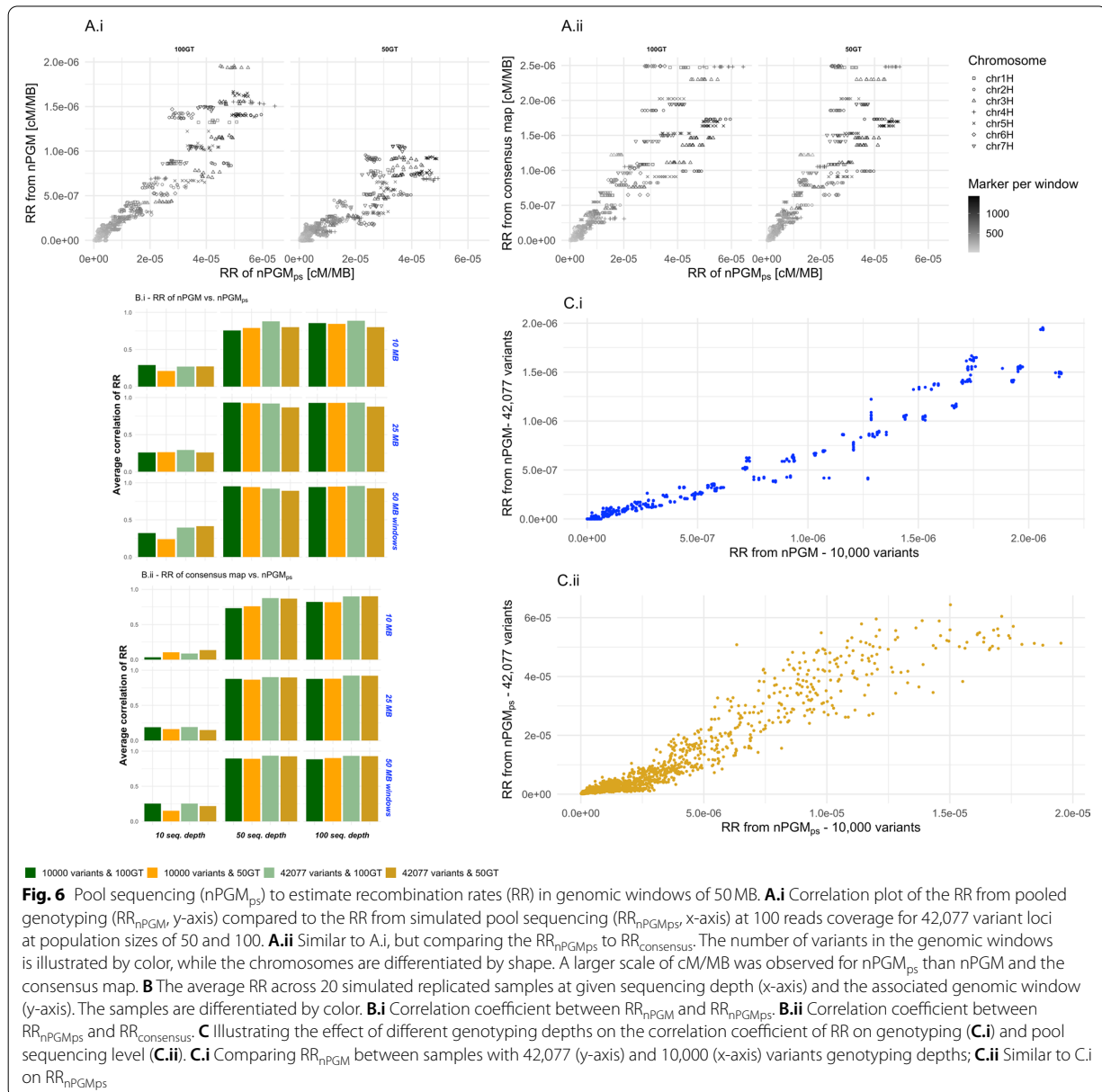
Effect of sequencing bias on nPGM accuracy

Next, we evaluated the genetic map estimation from pooled sequencing data using simulated reads for 42,077 and 10,000 SNPs and three sequencing depths. Limitations of the simulation software did not allow reliable simulations with more than 100 genotypes; thus, we evaluated population sizes of 50 and 100. After simulation, variant calling, and allele frequency estimation, the

genetic map of the simulated pool sequencing was calculated ($nPGM_{ps}$), and the corresponding RR_{nPGMps} was assessed.

While the population size and genotyping depth did not significantly ($P_{popSize} = 0.21$; $P_{genotypingDepth} = 0.56$) affect the RR_{nPGMps} estimation accuracy, the sequencing depth and the genomic window size significantly ($P_{seqDepth} < 0.0001$; $P_{genWindow} < 0.001$) impacted the accuracy. When RR_{nPGM} and RR_{nPGMps} were compared based on a shallow sequencing depth of 10 reads per locus, we observed a low Pearson correlation of 0.26 between them (Fig. 6B.i). However, the correlation coefficient increased to 0.88 and 0.9 for sequencing depth of 50 and 100 reads/locus (Fig. 6A.i & B.i).

The correlation coefficients between RR_{nPGMps} and RR_{nPGM} , estimated in genomic windows of 10 MB, were about 90% lower than that observed for 50 MB windows. This was caused by the low SNP density in the 10 MB windows. Generally, the highest correlation was observed for a read depth of 100 in 50 MB windows ($r = 0.94$, Suppl. Figure 5).



Similar patterns of Pearson correlation coefficients were observed when comparing RR_{nPGMps} and $RR_{consensus}$. The highest correlation was detected for genomic windows of 50 MB and a sequencing depth of 100 ($r=0.912$, Fig. 6A.ii & B.ii). In contrast to the comparison between RR_{nPGM} and RR_{nPGMps} , where the correlation was higher for 10,000 than 42,077 variant loci, a higher correlation coefficient ($r=0.918$) was observed for the scenario with 42,077 variant loci compared to that with 10,000 variant loci ($r=0.863$) when considering the comparison between RR_{nPGMps} and $RR_{consensus}$.

The assessment of the absolute levels of the RR_{nPGMps} based on the different population sizes (Fig. 6C) revealed no influence of this parameter (Fig. 6C.i, 1:1 ratio to RR_{nPGM}). However, the genotyping depth strongly impacted the absolute value of the RR_{nPGMps} (Fig. 6C.ii, 4:1 ratio to RR_{nPGM}).

Similarly, we observed a mean overestimation of RR_{nPGMps} to RR_{PGM} of 16.7 times (Fig. 6A, axis scales). This observation indicates that the sequencing procedure likely adds extra allele frequency deviations.

Discussion

Need for cheap RR estimation

The accurate and cost-efficient RR estimation of populations, lines, species, or genetic material that experienced different environmental cues is a technique many research fields could benefit from [35–39]. Commonly used approaches to estimate recombination rates require both the haplotype and allele frequency [40]. This information is typically derived from genotyping or sequencing of single individuals. However, crossing-overs per chromosome and meiosis are typically limited to one to four [28]. Therefore, many individuals must be genotyped to obtain accurate recombination rate estimations [41].

Concept of pool-based RR estimation and its evaluation using computer simulations

This study describes an approach for RR estimation that does not require genotyping or sequencing of single individuals – without considerable sacrificing accuracy. The method relies on two sources of information: (i) the allele frequency at each polymorphic locus and (ii) the physical position of these loci. We followed the idea that the allele frequency difference of two neighboring polymorph loci indicates a crossing-over [34]. Thus, our approach does not require collecting haplotype information by genotyping single individuals.

The required allele frequency variations across the genome are caused by the combination of migration,

high correlation coefficient ($r > 0.9$) between RR_{PGM} and $RR_{\text{consensus}}$ (Fig. 3). Nevertheless, the actual PGM map length and the extent of RR_{PGM} were (i) underestimated and (ii) depended on the genotyping depth and population size (Figs. 2 & 3). Exemplarily, the average map length in the simulated samples with only 500 SNPs was 20 times shorter than the map length of samples with 42,077 SNPs (Fig. 2). This can be explained thereby that with increasing genotyping depth, undetected recombination in maps with fewer loci will be observed, which increases the recombination rate. From this observation, we concluded that it is crucial to integrate the number of polymorph loci in the RR_{PGM} estimation. Especially when only a few polymorph loci are available, the variation in genotyping depth between two populations might affect the comparison. In addition, we also observed an effect of the population size (Fig. 3) on RR_{PGM} . Fewer genotypes resulted in a higher deviation of the actual allele frequency, which resulted in a higher RR_{PGM} estimate than expected. In analogy to the genotyping depth, this might not be relevant in weakly unbalanced experimental designs, but an adjustment might prevent the overestimation of RR_{PGM} .

Therefore, a linear and a non-linear model were examined to adjust the extent of RR_{PGM} by considering the genotyping depth and the population size. The non-linear least square model performed superior to the linear model (Suppl. Figure 2). The final model implemented in the further comparisons was:

$$n = 7958.92 * e^{-0.5401 * \log_2 \sum \text{SNPs}} * e^{0.3491 * \log_2 \sum \text{Genotypes}} + \frac{691.0495}{\sqrt{\sum \text{SNPs}}}$$

selection, drift, or gene flow [42]. However, even unintended selection or drift can result in traceable allele frequency deviations in populations (Suppl. Figure 1).

The extent of cross-overs can be quantified based on this concept, but the effective RR cannot be derived from allele frequency variations alone. For example, recombination between two loci with a distance of one Mb is much less likely than between two loci separated by 10 Mb. Suppose the allele frequency variation in both situations is identical. In this case, the recombination likelihood in the small interval is much lower; therefore, the local RR must be higher than that of the big interval. Accordingly, we scaled the allele frequency deviation by the \log_{10} of the physical distance of the considered loci to calculate parameters related to the local RR (Eq. 1).

The first objective of our study was to propose an accurate and reliable method for genetic map position and RR estimation in defined genomic windows. For a genotyping depth above 10,000 markers, we observed a

Using this model's result, multiplied with the outcome of eq. 1, provides an unbiased estimation of the recombination rate. This adjustment of PGM to nPGM resulted in genetic maps having the same map extension as HGM (Fig. 2), regardless of the genotyping depth ($\sum \text{SNPs}$) or population size ($\sum \text{Genotypes}$). Furthermore, we could show that the correlation coefficient between RR_{nPGM} and $RR_{\text{consensus}}$ was only slightly lower than the correlation coefficient between $RR_{\text{consensus}}$ and RR_{HGM} (Fig. 4A & C). Especially when the genotyping depth was high, the correlation coefficients were almost identical. An even higher correlation coefficient between RR_{nPGM} and RR_{HGM} was observed than between RR_{nPGM} and $RR_{\text{consensus}}$. This can be explained thereby that the simulation of populations introduced a measurable simulation error. These observations indicated that RR estimation from pooled samples is possible with high accuracy at dramatically reduced costs.

In addition to the correlation of the recombination rates, we evaluated the accuracy of the RR_{nPGM} estimation on a genome-wide scale. This analysis indicated an overestimation of the RR_{nPGM} in non-pericentromeric regions of the genome (Fig. 4D). This deviation is presumably caused by the different variant distribution of the SNP array in the non-pericentromeric compared to the pericentromeric region. Therefore, the observed overestimation of $RR_{non-pericentromeric}$ regions is only problematic if the RR is compared between different genotyping approaches.

Pool-based RR estimation in experimental populations

The 45 HvDRR populations were characterized by varying genotyping depth (deviation of lowest to highest – 5.79x) and population sizes (variation of smallest to largest – 3.76x). Therefore we applied the nPGM approach to adjust for genotyping depth and population size. Although we used the nPGM model described above, we observed a map length that was, across all populations, about 33% lower compared to the HGM reported by Casale et al. (2021) (Fig. 5). This observation can be explained thereby that the experimental populations are RIL populations, while the model underlying the nPGM approach was established based on simulated F2 populations. The total number of recombination events accumulated in the gametes of a RIL, after endless selfing generations, was about twice the number of such events in an F2 population [43]. Therefore, if the absolute value of the map positions is of interest, then the model underlying nPGM approach needs to be derived de novo for the population type under consideration. However, in analogy to the results of the simulations, the map length variations did not affect the correlation of the RR_{HGM} to RR_{nPGM} , which was $r > 0.8$ for 23 of 45 populations ($r_{Spearman} > 0.6$ for 41 populations, Fig. 5B). We explored potential reasons for these deviations and observed that the populations with a correlation of the RR_{HGM} to $RR_{nPGM} < 0.6$ were characterized by a median inter-marker distance that was about 40% lower than that of the other populations (Suppl. Figure 6). We tested this effect for statistical significance in a linear model and retained a significant effect of the genome-wide median inter marker distance and standard deviation on the Pearson correlation of RR_{nPGM} to RR_{HGM} ($p_{Median} < 0.003$; $p_{Sd} < 0.002$). Similarly, we observed the same effect on the Spearman correlation ($p_{Median} = 0.001$; $p_{Sd} = 0.0015$). Contrary, no genotyping depth or population size effect was observed ($p_{GD} = 0.34$; $p_{PS} = 0.33$). We conclude from this observation that a skewed distribution of genomic marker distances can significantly affect the RR_{nPGM} estimation. One possibility to overcome this problem is to

sample the loci such that all loci with a distance below 10,000bp are omitted for further progression with the nPGM approach. However, this requires further research.

Subsequently, we were interested in comparing the general recombination estimate (GRE) derived from the nPGM approach with that from HGM. This parameter summarizes the RR of a parental genotype in combination with several parental genotypes and is highly relevant for breeders of all crops, exemplarily in introgression breeding [8]. Compared to the HGM-based GRE of Casale et al. (2021), the GRE calculated from nPGM resulted in almost the same ranking of the involved 23 parental inbreds (Suppl. Figure 4). Deviations in the ranking between nPGM and HGM-derived GRE might be due to discrepancies between the genetic and physical order of the underlying marker (Suppl. Figure 3), which either can be artifacts from the HGM approach or are structural variants in the genomes of some of the parental inbreds.

These observations together illustrated the validity and accuracy of RR estimates from nPGM also in experimental populations.

Pool-based RR estimation by sequencing

For the above-described results, we derived pool genotyping data from genotyping information of individuals as a starting point for evaluating our approach. This procedure results in the upper limit of the accuracy as it neglects the variation in allele frequency that is caused by its estimation in a pool. One possibility would be to consider this aspect in our simulations of genotyping with an SNP array. However, with today's sequencing costs [43], applying our method to datasets created from the sequencing of pooled samples is even more economically attractive. Therefore, we estimated the accuracy of recombination rate estimation from simulated pool sequencing samples. The correlation between $RR_{consensus}$ and RR_{nPGMps} was, at a coverage of 10 reads per locus, at a rather low level of about 0.3 (Fig. 6B.ii). However, we observed that increasing the read dept. from 10 to 50 reads per locus reduced the median variation of simulated pool sequencing compared to the RR from pool genotyping by 40% (Supple. Figure 5).

Similarly, $RR_{consensus}$ and RR_{nPGMps} correlation increased to 0.93 in 50 MB genomic windows at 50x coverage. A further increase of the read coverage to 100 did not result in similarly high additional precision, indicating that saturation was reached. The second aspect that was studied, in addition to the sequencing depth, was the size of the genomic window for which the RR was estimated. At a sequencing depth of 100 reads, the median error

was reduced by 50% when comparing the 10 to 25 MB genomic windows (RMSE 10 MB = 0.0004; 50 MB = 0.0002). The error was only further reduced by 2% comparing the RMSE of 25 and 50 MB windows (Supple. Figure 5). The choice of a reasonable window size for summarizing the RR is impacted by the number of variants present in a window. In our simulations, we assumed conservatively 10,000 and 42,077 genome-wide variant loci. However, a considerably higher number of polymorphic loci in most species will be identified when sequencing strategies are applied. For barley, e.g., more than 57 Mio. SNPs were collected in a variant database [44], indicating that more than 1350x variants than those used in our study are already known. Therefore, we expect that the window size can be considerably decreased down to less than 1 MB in future experimental studies. This in turn allows to increase the resolution.

Besides the high correlations of RR_{nPGMps} to RR_{nPGM} , we noticed a significant overestimation of RR_{nPGMps} compared to RR_{nPGM} (16.7 times higher, Fig. 6A). This observation might be due to the additional allele frequency variation between adjacent loci caused by sequencing errors. However, the above-described overestimation only matters when comparing the RR among different methods, like RR_{HGM} to RR_{nPGMps} .

Furthermore, we also observed a variation in the scale of RR between the genotyping depth levels. The extent of RR_{nPGMps} increased with the genotyping depth (Fig. 6C.ii). Higher genotyping depth might be associated with a smaller inter polymorphism distance and, therefore, might lead to a further increased overestimation of the RR_{nPGMps} . To generate a comparison on the same scale, a simple linear model correction for the RR_{nPGMps} might be suitable to compare it to other approaches' derived RR. Apart from this overestimation of the RR, the RR_{nPGMps} , RR_{nPGM} , and $RR_{consensus}$ indicated high similarities in the genomic window-base recombination estimation (Fig. 6B).

Comparison of the nPGM approach to other approaches of RR estimation

Generally, the observed accuracy of our approach of estimating the RR in pooled samples might overcome issues of related approaches, like high costs, and allow a high throughput screening for GRE.

Other attempts to solve the dilemma of high costs have been proposed earlier. For example, [45] proposed an ultra-low individual sequencing strategy, followed by an imputation step to recover none sequenced regions in the library. Nevertheless, the imputation might also

introduce errors in the recombination estimation, making accurate recombination estimation challenging.

Other approaches minimize the number of test samples by implementing Markov Coalescent models or machine learning strategies trained in different subsamples or even species [46, 47]. Few single genotypes need to be sequenced in these approaches to estimate the genetic map to retain haplotype information in the sample. This is based on applying genetic maps from related species might be a useful approach to estimate the RR, especially when few samples or no reference genome are available or costs should be reduced. In situations where no reference genome is available, our $nPGM_{ps}$ method cannot be performed and is inferior to these methods. Nevertheless, the RR might differ from one species to another [48], and our proposed $nPGM_{ps}$ approach allows differentiating populations of the same species with a much higher resolution.

Sun et al. (2019) showed that the unexpected breaks in linked read sequencing of F_1 plants' pollen could denote recombination events. While this approach is complex and costly to perform, generating a pooled sample with equal tissue contribution of each genotype from leaves or seeds underlying our method is technically easy. Furthermore, our method allows genotyping of undefined population sizes without cost inflations. The $nPGM_{ps}$ method does not demand more than 10 to 100 reads coverage per locus, while the pool-linked read sequencing of haploid cells requires ultra-high sequencing depth across all the pollen. Furthermore, pool sequencing prices can be further decreased when the sequencing depth is reduced [34, 49]. The only disadvantage of our method is that it is not based on the F_1 generation as the approach of Sun et al. (2019) but requires the establishment of at least the F_2 generation. However, that is possible for most species without big space limitations and is more than balanced by the considerably lower costs.

Implementation of the (n)PGM_{ps} approach in other genetic materials or species

In order to generate a genome-wide genetic map for a species of interest using the PGM approach, the following prerequisites have to be fulfilled. First, a reference sequence must be available to align short reads to annotated positions. Second, a pool sequencing strategy has to be chosen that ideally allows to remove duplicated reads (unlike restriction-site based genotyping by sequencing) and is unbiased regarding the expression level (like RNAseq). This is because such sequencing procedures can bias the accurate allele frequency estimation and therefore are less suitable for pool sequencing [50] and RR_{nPGMps} estimation.

Consequently, we propose whole-genome sequencing as the most convenient method to generate high-confidence allele frequency estimates (Table 2 – 1.1-1.3). Furthermore, a sequencing depth of approximately 100x or higher will result in sufficiently accurate allele frequency estimations. Nevertheless, a 100x coverage is associated with high monetary costs, especially for crop species with large genomes, so one might want to sequence a pool on a lower coverage level (exemplarily 10x, Table 2 – 2.1-2.7). SNP allele frequency aggregation to a haplotype (window) frequency is suitable for increasing precision in such cases. The haplotype creation can either be based on defined genomic windows [49] or on genomic features, like genes [51]. However, it must be pointed out that such haplotype aggregations or generally lower counts of detected variant loci (like GBS) will reduce the RR resolution (cf. Table 2 – 1.4).

Finally, in case the absolute map length of the PGM is of interest, a genetic map of the variants under consideration is required to scale the observed RR. This step is required, as the presented model cannot accommodate the entire variety of sequencing-induced allele frequency deviations and, thus, was not included in the model fitting. Instead, we propose identifying

the typical genetic map extension size in the species of interest and performing a linear scaling of the genetic map position and recombination rate according to Table 2, 3.3.

Beyond the relative RR and map length estimations, this case study presented a method to overcome variations in genotyping depth and the population size by exploiting computer simulations. We recommend the map length adjustment by genotyping depth and population size only in cases where the populations to compare are characterized by highly different numbers of genotypes, the polymorphism count is highly variant, or sequencing depth varies.

Conclusion

This case study presents a method that allows a cost-efficient estimation of genetic maps and the recombination rate in genomic windows. Our approach exploits the allele frequency and the physical position information. Furthermore, based on computer simulations and experimental data, we have shown that the proposed method allows an accurate assessment of RR. Finally, we have explained how to apply the procedure for other species and discussed potential pitfalls. The functions presented

Table 2 Best practice guideline to estimate the recombination rate from pooled sequencing data

Step	Task	1. Step – selection of genotyping approach			
1.1	Select sequencing method	Suitable approaches are WGS, Exome capture, RNAseq & GBS			
1.2	Sample treatment	WGS	Exome Capture	RNAseq	GBS
1.3.a	General instructions	ensure equal DNA / tissue contribution of each genotype in pool			
1.3.b	Specific instructions	-	-	Identical time point of sampling & same tissue	Restriction site digestion results in inability to remove duplicates. Allele frequency potentially biased in small genomic windows
1.4	Resolution of RR and genetic map	many polymorph loci (millions) → high resolution	lot of polymorph loci (hundred thousand) → high resolution	lot of polymorph loci (hundred thousand) → high resolution	few polymorph loci (ten thousands) → high resolution

2. Step – accurate allele frequency estimation		
2.1	Target seq. Coverage	$\geq 100\times$ (high coverage) ≤ 20 (low coverage)
2.2	Additional Information	optimal (genotyping parental lines) mandatory - sequencing parental lines (for haplotype reconstruction)
2.3	Alignment processing	remove ambiguous reads => reads with multiple loci (secondary supplementary alignments); IMPORTANT STEP Applicable to all model species and crops (precision is more relevant than retaining many reads!)
2.4	Duplicate removal	yes yes (not recommended for GBS)
2.5	Variant calling	call the allelic depth; if available, use a variant database and use only those called variants which are reported in the database
2.6	Variant filtering	remove monomorphic and low quality SNPs (QUAL ≥ 250)
2.7	Aggregating SNPs to Haplotypes	useful to annotate variants to genes, <u>not</u> mandatory essential to compensate for the low sequencing coverage and/or low confidence in accurate SNP allele frequency calling - mandatory

3. Step – Recombination estimation (nPGM)		
3.1	Window size selection	High genotyping depth $\sim \geq 100,000$ SNPs → ≤ 1 MB Low genotyping depth $\sim \leq 50,000$ SNPs → ≥ 5 or 10 MB
3.2	Adjustment of samples to each other	Providing population size (number of individuals pooled per pool sample) to popRR to derive adjusted recombination rates. Genotyping depth is referenced from derived variant table (vcf)
3.3	Genetic map position adjustment	Using a known consensus genetic map to scale the nPGM in relation to common genetic map length scaled nPGM RR = nPGM RR x [Consensus Map Length] / [popRR Map length] – not mandatory

in this publication can be obtained from GitHub <https://github.com/mischn-dev/popRR.git> for both *R* and *Julia* environments, using a filtered VCF file as input.

Methods

Consensus map-based population simulations

Our simulations were based on the consensus genetic map generated by Casale et al. (2021). In brief, 45 recombinant inbred line populations have been created by crossing 23 parental inbreds in a double round-robin design [52]. Each of the 4182 RIL was genotyped using a 50K SNP array [23], and the 45 genetic maps have been integrated. The resulting consensus map comprises 42,077 SNPs with a genetic and physical position [18] (Fig. 1, steps 1–3).

For the simulations, two virtual parental genotypes with different alleles for each of the 42,077 loci were generated and alphaSim [53] was used to derive F2 populations (F1 by crossing, F2 by selfing) with various populations sizes (50; 100; 500; 1000; 2000; 4000; 10,000 genotypes), and various genotyping depths (500; 1000; 2000; 5000; 10,000; 15,000; 20,000; 30,000; 42,077) across the entire genome. These simulations were repeated 20 times for each SNP–genotype count combination (i.e. in total 1260 populations, Fig. 1 – Step 4). For each replicate, a different set of SNPs (except the 42,077 sample) was sampled.

Recalculation of genetic maps from simulated populations based on Haldane's mapping function

To estimate the error introduced by the simulation process to the consensus map, we recalculated the Haldane genetic maps for all 20 replicates in populations with 50, 500, and 2000 genotypes. For computational reasons, we ordered the SNPs first by their physical position to realize a correct starting point. Subsequently, the Haldane genetic map (HGM) was calculated using the *qtl* package based on Haldane's mapping function at an error probability of 0.0001 [54]. Finally, the RR was calculated as the median centiMorgan per megabase pair [cM/MB] value in 50Mb windows across all variants in this window (Fig. 1, step 6.1).

Genetic maps from pooled samples

The alleles in a segregating population derived from two parental inbreds are expected to have a frequency of 0.5. However, due to selection or random sampling, the

allele frequency at a locus can deviate from this expected frequency. Notably, the deviating allele frequency is expected to attenuate distally toward the expected frequency due to increasing crossover events between the locus and gradually more distal loci [55]. Therefore, the allele frequency and its rate of change should be related to the genetic distance. The genetic map can be generated with as little as one library preparation since genome-wide allele frequency can be determined using whole-genome genotyping or sequencing of a pool of individuals from the population of interest [53]. Our study evaluates whether allele frequency differences across the genome can be used to estimate RRs and genetic maps, even in situations without substantial fitness differences.

We dismissed any individual genotype information after calculating the allele frequency at each SNP across all genotypes by pooling individuals' genotypic information (Fig. 1, step 5.1). We estimated the factor K_{M1M2} as:

$$K_{M1M2} = \frac{\Delta AF_{M1M2}}{\log_{10} \Delta Dist_{M1M2}} \quad (1)$$

where ΔAF_{M1M2} was the allele frequency deviation of the considered physically neighboring SNP pair (M1, M2) and $\log_{10} \Delta dist$ the decadic logarithm of the physical distance between them. The factor K_{M1M2} , which comprises the two SNPs' relative recombination rate, was added up along the chromosome to generate a pool genetic map (PGM, Fig. 1, step 6.2). As the absolute size of factor K_{M1M2} can not be interpreted, it needs to be scaled first. In the first step, we adjusted the PGM using the adj_{start} correction factor, which was calculated as the ratio between the length of the consensus map across all chromosomes (ML_{ref}) in cM and the sum of the PGM across all chromosomes (ML_{PGM}). An adjustment value adj_{start} was calculated separately for each simulated sample (Fig. 1, steps 7 & 8).

The above-described correction factor adj_{start} was used to estimate the effect of the genotyping depth (*Markers*) and population size (*Genotypes*) on the map length in order to realize in the next step a correction of the map length for these two factors. Therefore, we evaluated a simple linear model without intercept (Eq. 2; Fig. 1, step 8.1):

$$adj_{start} = a * \sum Markers + b * \sum Genotypes \quad (2)$$

and a non-linear least square model (nls, Eq. 3; Fig. 1, step 8.2):

$$adj_{start} = \alpha * e^{\beta * \log_2 \sum Markers} * e^{\gamma * \log_2 \sum Genotypes} + \frac{\theta}{\sqrt{\sum Markers}} \quad (3)$$

and compared them concerning AIC and log-likelihood to identify the best fitting model.

The nls model described above comprised four sub-transformations (α , β , γ , θ) and the SNP and genotype count were \log_2 transformed. For both models, the parameters were estimated across all simulated 1260 populations. Based on these estimates, the correction factor adj_{start} was calculated using each population's genotyping depth and population size (Fig. 1, step 8.3).

According to the observed log-likelihood and AIC, the nls model was used in all following analyses and multiplied to each SNP's K value to generate a corrected PGM estimate (nPGM), (Eq. 4; Fig. 1, step 9).

$$K' = K_{M1M2} * adj_{start(nls)} \quad (4)$$

RR estimation from adjusted pool genetic map (nPGM)

RR [cM/MB] was calculated from the nPGM for each SNP pair. Next, an average RR value was calculated for 50Mb windows, applying a sliding window approach (window size 50 MB, slide 0.5 x window size).

Finally, the RR of the simulated populations with 50, 500, and 2000 genotypes on all genotyping depths was compared between (i) nPGM (RR_{nPGM}) and the consensus map ($RR_{consensus}$) (ii) HGM (RR_{HGM}) and the consensus map (Fig. 1, step 10), and (iii) nPGM (RR_{nPGM}) and HGM (RR_{HGM}).

nPGM calculation in experimental populations

The previously described 45 HvDRR populations were used to estimate the nPGM in experimental populations and compare the RR_{nPGM} to the RR_{HGM} . The HGM was calculated as described by Casale et al. (2021). For the nPGM construction of each population, monomorphic SNPs and SNPs with identical or missing physical positions were omitted. In addition, SNPs with more than 10% missing information were omitted as well. Finally, the allele frequency was calculated, and the nPGM was derived from it, as was described above. The nPGM was used to estimate the RR_{nPGM} (Fig. 1, steps II & II.i). The RR_{nPGM} estimate accuracy was assessed by comparing it to RR_{HGM} .

Impact of sequencing error on the pool genetic map estimation accuracy

In the above-explained simulations, the allele frequency was calculated from the genotypic information of individual samples. However, the primary purpose of our nPGM approach was the recombination estimation from pool sequencing data. Therefore, based on the allele frequency of the individual genotyping simulations, we performed a pool sequencing simulation to estimate

the effect of both the sequencing and sampling error on the genetic map estimation accuracy using the nPGM approach (Fig. 1, step 11). Therefore, we selected four scenarios, characterized by a genotyping depth of 10,000 and 42,077 markers and a population size of 50 and 100 genotypes.

The *simReads* function of the Rsubread package [56] was used to simulate the sequencing data based on the allele frequency of the simulated populations and the barley reference genome (Barley Morex V2 pseudomolecules [57]; Fig. 1, step 5.2). *simReads* created a fastq file with a locus coverage of approximately 3000 reads per locus. From this set, three sequencing depths were sampled (10, 50 & 100 reads per locus) with ten replicates per combination of either 10,000 or 42,077 variants and 50 or 100 genotypes (*Sequencing depth x Genotyping depth x population size*).

In the next step, the subsets of simulated 100 bp long paired-end reads were aligned to the Barley Morex V2 pseudomolecules reference genome by *bwa mem* [58]. Following, the reads were filtered by omitting all reads with an alignment score below 60 using *samtools* [59]. Next, the variants were called from the aligned reads using *samtools 1.8 mpileup* and *bcftools 1.8 call* [60], where all reads with a variant quality below 30 were omitted.

Finally, the allele frequency and physical positions were extracted and based on eq. 5, a pool sequencing derived nPGM, named $nPGM_{ps}$, was calculated (Fig. 1, step 11.1). Next, we estimated from the $nPGM_{ps}$ the $RR_{nPGM_{ps}}$ and compared it in 10, 25, and 50 MB windows across the genome to the $RR_{consensus}$ and RR_{nPGM} . Furthermore, the two variant levels (10,000, 42,077) were compared to assess the effect of the genotyping depth in the pooling strategy.

Estimation of general recombination effect of parental inbreds

We calculated the general recombination effect of each of the 23 parental inbreds based on the nPGM, and compared it against the values reported by Casale et al. (2021). We used the same G-BLUP model to retain consistency in comparing both HGM and nPGM approaches. If not mentioned differently, all analyses were performed in R 4.0.2 [61] and Julia 1.6.2 [62].

Abbreviations

GRE: General recombination estimate; HGM: Haldane genetic map; HvDRR: Spring barley double round-robin populations; Lm: Linear model; MB: Megabase pairs; Nls: Non-linear least square model; nPGM: Non-linear adjusted pool genetic map; $nPGM_{ps}$: Non-linear adjusted pool genetic map derived from pool sequencing; ML: Map length; PGM: Pool genetic map; RIL: Recombinant inbred line; RR: Recombination rate; SNP: Single nucleotide polymorphism.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-022-08701-7>.

Additional file 1: Suppl. Figure 1. Deviation between observed and expected allele frequency (y-axis) for different numbers of genotypes per population (x-axis). The expected allele frequency value in an F2 population of infinite size is 0.5 and was set as the expected allele frequency. The observed allele frequency results from simulating a population with a given genotype count by AlphaSim. Each dot presents one simulated population. A total of 1260 populations were simulated. **Suppl. Figure 2.** Linear model (magenta) and non-linear least square (turquoise) models to predict the impact of a population's size and genotyping depth on the map extension (length). The model estimate on the y-axis is based on the pool genetic map estimation. Each point illustrates an individual population. The dashed line indicates the ideal fit. (step 8.1 & 8.2 in Fig. 1). **Suppl. Figure 3.** Marey map of genetic position (y-axis) against the physical position (x-axis) for all 45 experimental populations. The nPGM (coral) is compared against the HGM (blue). Chromosomes and populations are faceted. **Suppl. Figure 4.** The genome-wide general recombination effect for each parental inbred line, computed using a GBLUP model, based on the nPGM genome-wide RR observations. **Suppl. Figure 5.** The correlation plot of the RR from pooled genotyping (y-axis) compared to the RR from simulated pool sequencing (x-axis) at 100 reads coverage in 10 MB (A) and 50 MB (B) genomic windows. Four samples, differing in marker or genotype count, are indicated by the numbers 1 to 4 for both A and B. The number of variants in the genomic windows is indicated by color, while the chromosomes are differentiated by shape. **Suppl. Figure 6.** The effect of the median marker distance on the RR_{nPGM} to RR_{HGM} correlation coefficients across all HvDRR populations. A - the effect of median marker distance (bp) on the Pearson correlation. B - the effect of the median marker distance on the Spearman correlation. C - the genome-wide distribution of inter-marker distance (bp) for four HvDRR populations, characterized by a low (yellow, HvDRR08, HvDRR43) and a high (turquoise, HvDRR11, HvDRR43) RR_{nPGM} to RR_{HGM} Pearson correlation.

Acknowledgments

We thank Prof. Jens Léon for his suggestions and support in preparing this manuscript. Computational infrastructure and support were provided by the Center for Information and Media Technology (ZIM) at Heinrich Heine University Düsseldorf.

Authors' contributions

BS and MS conceptualized the research. MS and FC analyzed the data. MS, BS and FC wrote the manuscript. The author(s) read and approved the final manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Availability of data and materials

Code with example data sets and explanations to perform the analysis are provided at <https://github.com/mischn-dev/popRR.git>. HvDRR population information and genetic maps can be found in Casale et al. (2021).

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Institute of Quantitative Genetics and Genomics of Plants, Heinrich Heine University, 40225 Düsseldorf, Germany. ²Max Planck Institute for Plant Breeding Research, 50829 Köln, Germany. ³Cluster of Excellence on Plant Sciences, From Complex Traits Towards Synthetic Modules, Universitätsstraße 1, 40225 Düsseldorf, Germany.

Received: 6 April 2022 Accepted: 15 June 2022

Published online: 25 June 2022

References

- Roeder GS. Meiotic chromosomes: it takes two to tango; 1997.
- Comeron JM, Ratnappan R, Bailin S. The Many Landscapes of Recombination in *Drosophila melanogaster*. PLoS Genet. 2012;8:e1002905 [Cited 2021 Mar 18]. Public Library of Science; Available from: <https://dx.plos.org/10.1371/journal.pgen.1002905>.
- Jensen-Seaman MI, Furey TS, Payseur BA, Lu Y, Roskin KM, Chen CF, et al. Comparative recombination rates in the rat, mouse, and human genomes. Genome Res. 2004;528–38 [Cited 2021 Feb 25]. Available from: www.genome.org.
- Cao J, Schneeberger K, Ossowski S, Günther T, Bender S, Fitz J, et al. Whole-genome sequencing of multiple Arabidopsis thaliana populations. Nat Genet. 2011;43:956–65.
- Silva-Junior OB, Grattapaglia D. Genome-wide patterns of recombination, linkage disequilibrium and nucleotide diversity from pooled resequencing and single nucleotide polymorphism genotyping unlock the evolutionary history of Eucalyptus grandis. New Phytol. 2015;208:830–45 Blackwell Publishing Ltd.
- Pan Q, Li L, Yang X, Tong H, Xu S, Li Z, et al. Genome-wide recombination dynamics are associated with phenotypic variation in maize. New Phytol. 2016;210:1083–94 [Cited 2021 Oct 19]. John Wiley & Sons, Ltd. Available from: <https://onlinelibrary.wiley.com/doi/full/10.1111/nph.13810>.
- Kong A, Barnard J, Gudbjartsson DF, Thorleifsson G, Jonsdottir G, Sigurdardottir S, et al. Recombination rate and reproductive success in humans. Nat Genet. 2004;36:1203–6 Nature Publishing Group. [Cited 2021 Oct 19]. Available from: <https://www.nature.com/articles/ng1445>.
- Taagen E, Bogdanove AJ, Sorrells ME. Counting on crossovers: controlled recombination for plant breeding. Trends Plant Sci. 2020;25:455–65 Elsevier Current Trends.
- Battagin M, Gorjanc G, Faux AM, Johnston SE, Hickey JM. Effect of manipulating recombination rates on response to selection in livestock breeding programs. Genet Sel Evol. 2016;48:1–12 [Cited 2022 Jun 3]. BioMed Central Ltd. Available from: <https://gsejournal.biomedcentral.com/articles/10.1186/s12711-016-0221-1>.
- Semenov GA, Basheva EA, Borodin PM, Torgasheva AA. High rate of meiotic recombination and its implications for intricate speciation patterns in the white wagtail (*Motacilla alba*). Biol J Linn Soc. 2018;125:600–12 [Cited 2022 Jun 3]. Oxford Academic; Available from: <https://academic.oup.com/biolinnean/article/125/3/600/5106748>.
- Morgan TH. Sex limited inheritance in drosophila. Science (80-). 1910;32:120–2 [Cited 2021 Oct 19]. Available from: <http://www.espp.org>.
- Morgan TH. Random segregations versus coupling in Mendelian inheritance. Science (80-). 1911;34 Available from: <http://www.espp.org/foundations/genetics/classical/holdings/m/thm-1911a.pdf>.
- Morgan TH. An attempt to analyze the constitution of the chromosomes on the basis of sex-limited inheritance in drosophila. J Exp Zool. 1911;13:79.
- Petes TD. Meiotic recombination hot spots and cold spots. Nat Rev Genet. 2001;2:360–9 [Cited 2021 Oct 19]. Nature Publishing Group. Available from: <https://www.nature.com/articles/35072078>.
- Hunter N. Meiotic Recombination: The Essence of Heredity. Cold Spring Harb Perspect Biol. 2015;7:a016618 [Cited 2021 Oct 19]. Cold Spring Harbor Laboratory Press; Available from: <http://cshperspectives.cshlp.org/content/7/12/a016618.full>.
- Shen C, Li X, Zhang R, Lin Z. Genome-wide recombination rate variation in a recombination map of cotton. PLoS One. 2017;12:e0188682 Public Library of Science.

17. Schumer M, Xu C, Powell DL, Durvasula A, Skov L, Holland C, et al. Natural selection interacts with recombination to shape the evolution of hybrid genomes. *Science* (80-). 2018;360:656–60 American Association for the Advancement of Science.
18. Casale F, Van Inghelandt D, Weisweiler M, Li J, Stich B. Genomic prediction of the recombination rate variation in barley – A route to highly recombinogenic genotypes. *Plant Biotechnol J*. 2021; John Wiley & Sons, Ltd [Cited 2021 Dec 9]; Available from: <https://onlinelibrary.wiley.com/doi/full/10.1111/pbi.13746>.
19. Ganai MW, Durstewitz G, Polley A, Bérard A, Buckler ES, Charcosset A, et al. A large maize (*Zea mays* L.) SNP genotyping Array: development and germplasm genotyping, and genetic mapping to compare with the B73 reference genome. *PLoS One*. 2011;6:e28334 [Cited 2022 Feb 2]. Public Library of Science; Available from: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0028334>.
20. Sun C, Dong Z, Zhao L, Ren Y, Zhang N, Chen F. The wheat 660K SNP array demonstrates great potential for marker-assisted selection in polyploid wheat. *Plant Biotechnol J*. 2020;18:1354–60 [Cited 2022 Feb 2]. John Wiley & Sons, Ltd. Available from: <https://onlinelibrary.wiley.com/doi/full/10.1111/pbi.13361>.
21. Darrier B, Russell J, Milner SG, Hedley PE, Shaw PD, Macaulay M, et al. A comparison of mainstream genotyping platforms for the evaluation and use of barley genetic resources. *Front Plant Sci*. 2019;10:544 Frontiers Media S.A.
22. Lange TM, Heinrich F, Enders M, Wolf M, Schmitt AO. In silico quality assessment of SNPs—A case study on the axiom wheat genotyping arrays. *Curr Plant Biol*. 2020;21:100140 Elsevier.
23. Bayer MM, Rapazote-Flores P, Ganai M, Hedley PE, Macaulay M, Plieske J, et al. Development and evaluation of a barley 50k iSelect SNP array. *Front Plant Sci*. 2017;8:1792 [Cited 2021 Mar 10]. Frontiers Media S.A. Available from: <http://journal.frontiersin.org/article/10.3389/fpls.2017.01792/full>.
24. Turner TL, Bourne EC, Von Wettberg EJ, Hu TT, Nuzhdin SV. Population resequencing reveals local adaptation of *Arabidopsis lyrata* to serpentine soils. *Nat Genet*. 2010;42:260–3 Nature Publishing Group.
25. Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, et al. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One*. 2011;6:e19379 [Cited 2021 Oct 19]. Public Library of Science. Available from: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0019379>.
26. He J, Zhao X, Laroche A, Lu Z-X, Liu H, Li Z. Genotyping-by-sequencing (GBS), an ultimate marker-assisted selection (MAS) tool to accelerate plant breeding. *Front Plant Sci*. 2014;5:484 Frontiers.
27. Mascher M, Richmond TA, Gerhardt DJ, Himmelbach A, Clissold L, Sam-path D, et al. Barley whole exome capture: a tool for genomic research in the genus *Hordeum* and beyond. *Plant J*. 2013;76:494–505 [Cited 2021 Oct 19]. John Wiley & Sons, Ltd. Available from: <https://onlinelibrary.wiley.com/doi/full/10.1111/tpj.12294>.
28. Russell J, Mascher M, Dawson IK, Kyriakidis S, Calixto C, Freund F, et al. Exome sequencing of geographically diverse barley landraces and wild relatives gives insights into environmental adaptation. *Nat Genet*. 2016;48:1024–30 [Cited 2020 Jul 16]. Nature Publishing Group. Available from: <https://www.nature.com/articles/ng.3612>.
29. Lu K, Wei L, Li X, Wang Y, Wu J, Liu M, et al. Whole-genome resequencing reveals *Brassica napus* origin and genetic loci involved in its improvement. *Nat Commun*. 2019;10:1–12. [Cited 2021 Jun 15]. Nature Publishing Group. Available from: <https://doi.org/10.1038/s41467-019-09134-9>.
30. Wu D, Liang Z, Yan T, Xu Y, Xuan L, Tang J, et al. Whole-genome resequencing of a worldwide collection of rapeseed accessions reveals the genetic basis of ecotype divergence. *Mol Plant*. 2019;12:30–43 [Cited 2021 Jun 15]. Cell Press. Available from: <https://pubmed.ncbi.nlm.nih.gov/30472326/>.
31. Han Y, Gao S, Muegge K, Zhang W, Zhou B. Advanced Applications of RNA Sequencing and Challenges, vol. 9. London: SAGE PublicationsSage UK; 2015. p. 29–46. <https://doi.org/10.4137/BBIS28991>. [Cited 2021 Oct 19]. Available from: <https://journals.sagepub.com/doi/full/10.4137/BBIS28991>
32. Suárez-Vega A, Gutiérrez-Gil B, Klopp C, Tosser-Klopp G, Arranz JJ. Variant discovery in the sheep milk transcriptome using RNA sequencing. *BMC Genomics*. 2017;18:1–13 [Cited 2021 Oct 19]. BioMed Central; Available from: <https://link.springer.com/articles/10.1186/s12864-017-3581-1>.
33. Sun H, Rowan BA, Flood PJ, Brandt R, Fuss J, Hancock AM, et al. Linked-read sequencing of gametes allows efficient genome-wide analysis of meiotic recombination. *Nat Commun*. 2019;10:4310 Nature Publishing Group.
34. Wei KHC, Mantha A, Bachtrog D. The theory and applications of measuring broad-range and chromosome-wide recombination rate from allele frequency decay around a selected locus. *Mol Biol Evol*. 2020;37:3654–71 [Cited 2021 Feb 4]. NLM (Medline). Available from: <https://academic.oup.com/mbe/article/37/12/3654/5870837>.
35. Dreissig S, Maurer A, Sharma R, Milne L, Flavell AJ, Schmutzer T, et al. Natural variation in meiotic recombination rate shapes introgression patterns in intraspecific hybrids between wild and domesticated barley. *New Phytol*. 2020;228:1852–63 [Cited 2022 Feb 3]. John Wiley & Sons, Ltd. Available from: <https://onlinelibrary.wiley.com/doi/full/10.1111/nph.16810>.
36. Dreissig S, Mascher M, Heckmann S, Purugganan M. Variation in recombination rate is shaped by domestication and environmental conditions in barley. *Mol Biol Evol*. 2019;36:2029–39 [Cited 2022 Feb 3]. Oxford Academic. Available from: <https://academic.oup.com/mbe/article/36/9/2029/5519773>.
37. Serre D, Nadon R, Hudson TJ. Large-scale recombination rate patterns are conserved among human populations. *Genome Res*. 2005;15:1547–52 [Cited 2022 Feb 3]. Cold Spring Harbor Laboratory Press. Available from: <https://genome.cshlp.org/content/15/11/1547.full>.
38. Jordan KW, Wang S, He F, Chao S, Lun Y, Paux E, et al. The genetic architecture of genome-wide recombination rate variation in allopolyploid wheat revealed by nested association mapping. *Plant J*. 2018;95:1039–54 John Wiley & Sons, Ltd. John Wiley & Sons, Ltd. Available from: <https://onlinelibrary.wiley.com/doi/full/10.1111/tpj.14009>.
39. des Déserts AD, Bouchet S, Sourdis P, Servin B. Evolution of recombination landscapes in diverging populations of bread wheat. *Genome Biol Evol*. 2021;13 [Cited 2022 Feb 3] Oxford Academic. Available from: <https://academic.oup.com/gbe/article/13/8/evab152/6311266>.
40. Semagn K, Bjørnstad A, Ndjiondjop MN. Principles, requirements and prospects of genetic mapping in plants. *African J Biotechnol*. 2006;5:2569–87.
41. Ho WC, Zhang J. Evolutionary adaptations to new environments generally reverse plastic phenotypic changes. *Nat Commun*. 2018;9:1–11 [Cited 2020 Aug 3]. Nature Publishing Group. Available from: www.nature.com/naturecommunications.
42. Wright S. Evolution in mendelian populations. *Bull Math Biol*. 1990;52:241–95.
43. Bayle A, Droin N, Besse B, Zou Z, Boursin Y, Rissel S, et al. Whole exome sequencing in molecular diagnostics of cancer decreases over time: evidence from a cost analysis in the French setting. *Eur J Heal Econ*. 2021;22:855–64 [Cited 2022 Feb 4]. Springer Science and Business Media Deutschland GmbH. Available from: <https://link.springer.com/article/10.1007/s10198-021-01293-1>.
44. Tan C, Chapman B, Wang P, Zhang Q, Zhou G, Zhang XQ, et al. Barley-VarDB: a database of barley genomic variation. *Database*. 2020;2020 [Cited 2021 Dec 13]. Oxford Academic. Available from: <https://academic.oup.com/database/article/doi/10.1093/database/baaa091/6008688>.
45. Jensen SE, Charles JR, Muleta K, Bradbury PJ, Casstevens T, Deshpande SP, et al. A sorghum practical haplotype graph facilitates genome-wide imputation and cost-effective genomic prediction. *Plant Genome*. 2020;13:e20009 John Wiley and Sons Inc.
46. Adrien JR, Galloway JG, Kern AD. Predicting the Landscape of Recombination Using Deep Learning. *Mol Biol Evol*. 2020;37:1790 [Cited 2021 Oct 21]. Oxford University Press. Available from: <https://pmc/articles/PMC7253213/>.
47. Barroso GV, Puzović N, Dutheil JY. Inference of recombination maps from a single pair of genomes and its application to ancient samples. *PLoS Genet*. 2019;15:e1008449 Public Library of Science.
48. Smukowski CS, MAF N. Recombination rate variation in closely related species. *Hered*. 2011;107:496–508 [Cited 2021 Oct 21]. Nature Publishing Group. Available from: <https://www.nature.com/articles/hdy201144>.
49. Tilk S, Bergland A, Goodman A, Schmidt P, Petrov D, Greenblum S. Accurate allele frequencies from ultra-low coverage Pool-seq samples in evolve-and-resequence experiments. *G3 Genes, Genomes, Genet*. 2019;9:4159–68.

50. Schneider M, Shrestha A, Ballvora A, Léon J. High throughput crop genome genotyping by a combination of pool next generation sequencing and haplotype-based data processing. *Plant Methods*. 2021;49:1–17 Available from: <https://www.researchsquare.com/article/rs-415602/v1>.
51. Schneider M, Shrestha A, Ballvora A, Léon J. High-throughput estimation of allele frequencies using combined pooled-population sequencing and haplotype-based data processing. *Plant Methods*. 2022;18:1–18 [Cited 2022 Apr 4]. BioMed Central. Available from: <https://plantmethods.biomedcentral.com/articles/10.1186/s13007-022-00852-8>.
52. Stich B. Comparison of mating designs for establishing nested association mapping populations in maize and *Arabidopsis thaliana*. *Genetics*. 2009;183:1525–34 [Cited 2021 Dec 13]. Oxford Academic. Available from: <https://academic.oup.com/genetics/article/183/4/1525/6063168>.
53. Gaynor C, Gorjanc G, Hickey J. AlphaSimR: Breeding Program Simulations. 2020. Available from: <https://cran.r-project.org/package=AlphaSimR>
54. Broman KW, Wu H, Sen S, Churchill GA. R/qtl: {QTL} mapping in experimental crosses. *Bioinformatics*. 2003;19:889–90. Available from: <https://doi.org/10.1093/bioinformatics/btg112>.
55. Wei KHC, Mantha A, Bachtrog D. The theory and applications of measuring broad-range and chromosome-wide recombination rate from allele frequency decay around a selected locus. *Mol Biol Evol*. 2020;37:3654–71 [Cited 2021 Mar 11]. Oxford University Press. Available from: <https://academic.oup.com/mbe/article/37/12/3654/5870837>.
56. Liao Y, Smyth GK, Shi W. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res*. 2019;47:e47 Oxford University Press.
57. Monat C, Padmarasu S, Lux T, Wicker T, Gundlach H, Himmelbach A, et al. TRITEX: chromosome-scale sequence assembly of Triticeae genomes with open-source tools. *Genome Biol*. 2019;20:1–18 [Cited 2021 Oct 21]. BioMed Central. Available from: <https://genomebiology.biomedcentral.com/articles/10.1186/s13059-019-1899-5>.
58. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM; 2013. p. 1–3. Available from: <http://arxiv.org/abs/1303.3997>
59. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. *Bioinformatics*. 2009;25:2078–9 [Cited 2020 Feb 8]. Narnia. Available from: <https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btp352>.
60. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. The variant call format and VCFtools. *Bioinformatics*. 2011;27:2156–8 [Cited 2020 Sep 8]. Oxford Academic. Available from: <http://samtools.sourceforge.net>.
61. R Core Team. R: A Language and environment for statistical computing. Vienna; 2020. Available from: <https://www.r-project.org/>
62. Bezanson J, Edelman A, Karpinski S, Shah VB. Julia: A fresh approach to numerical computing. *SIAM Rev*. 2017;59:65–98 Society for Industrial and Applied Mathematics Publications.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

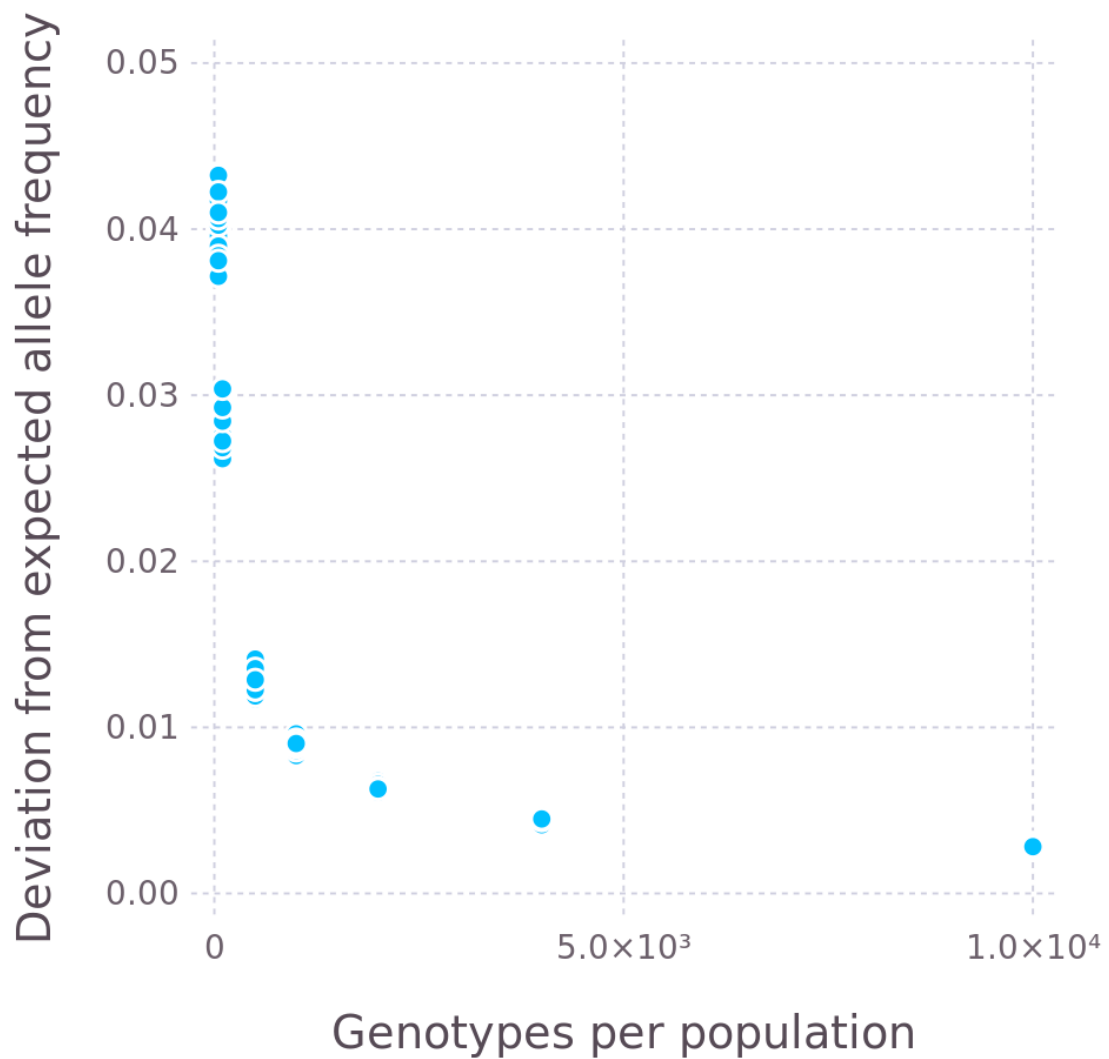
Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

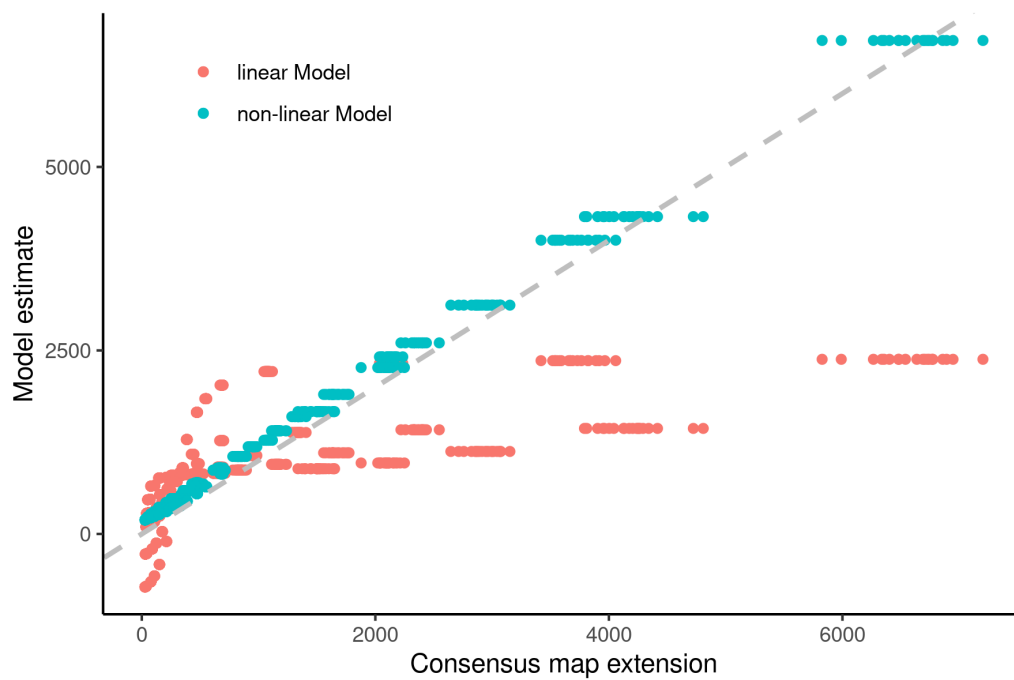
At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

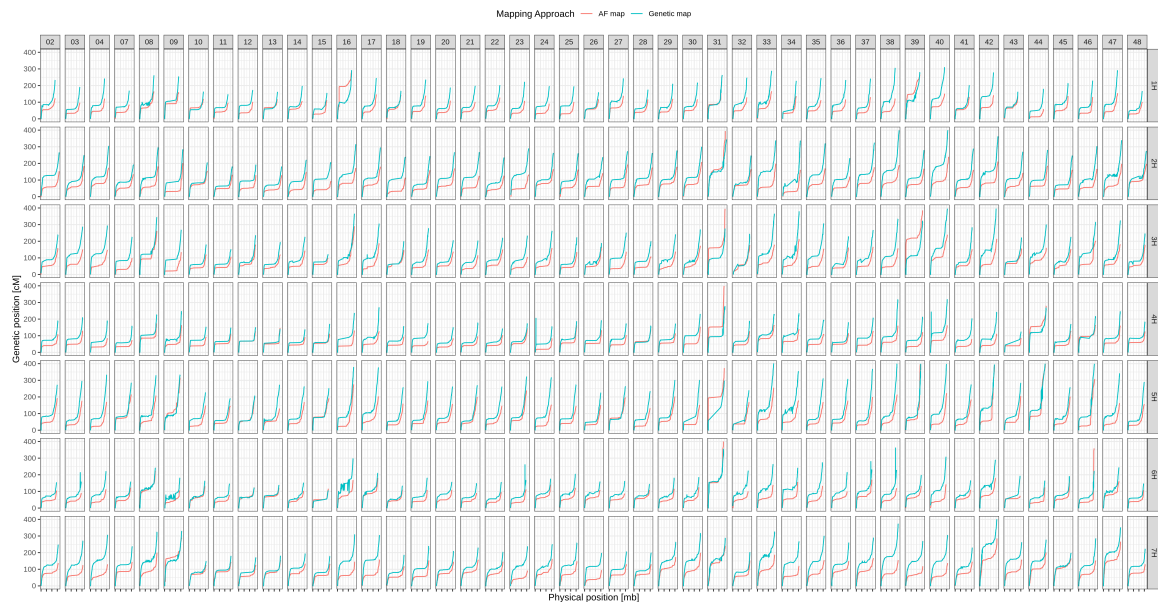




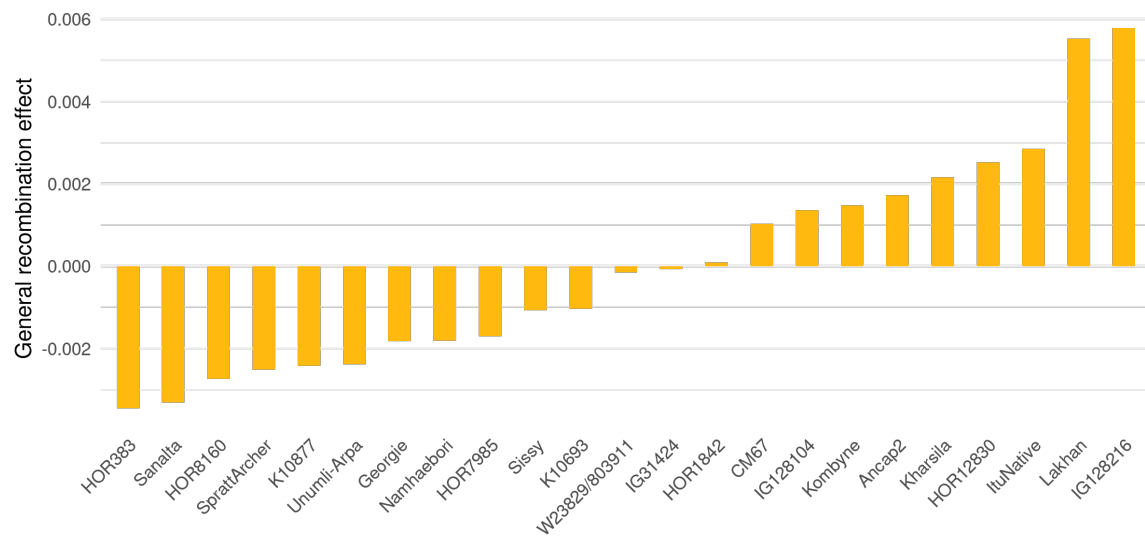
Suppl. Figure 1: Deviation between observed and expected allele frequency (y-axis) for different numbers of genotypes per population (x-axis). The expected allele frequency value in an F2 population of infinite size is 0.5 and was set as the expected allele frequency. The observed allele frequency results from simulating a population with a given genotype count by AlphaSim. Each dot presents one simulated population. A total of 1260 populations were simulated.



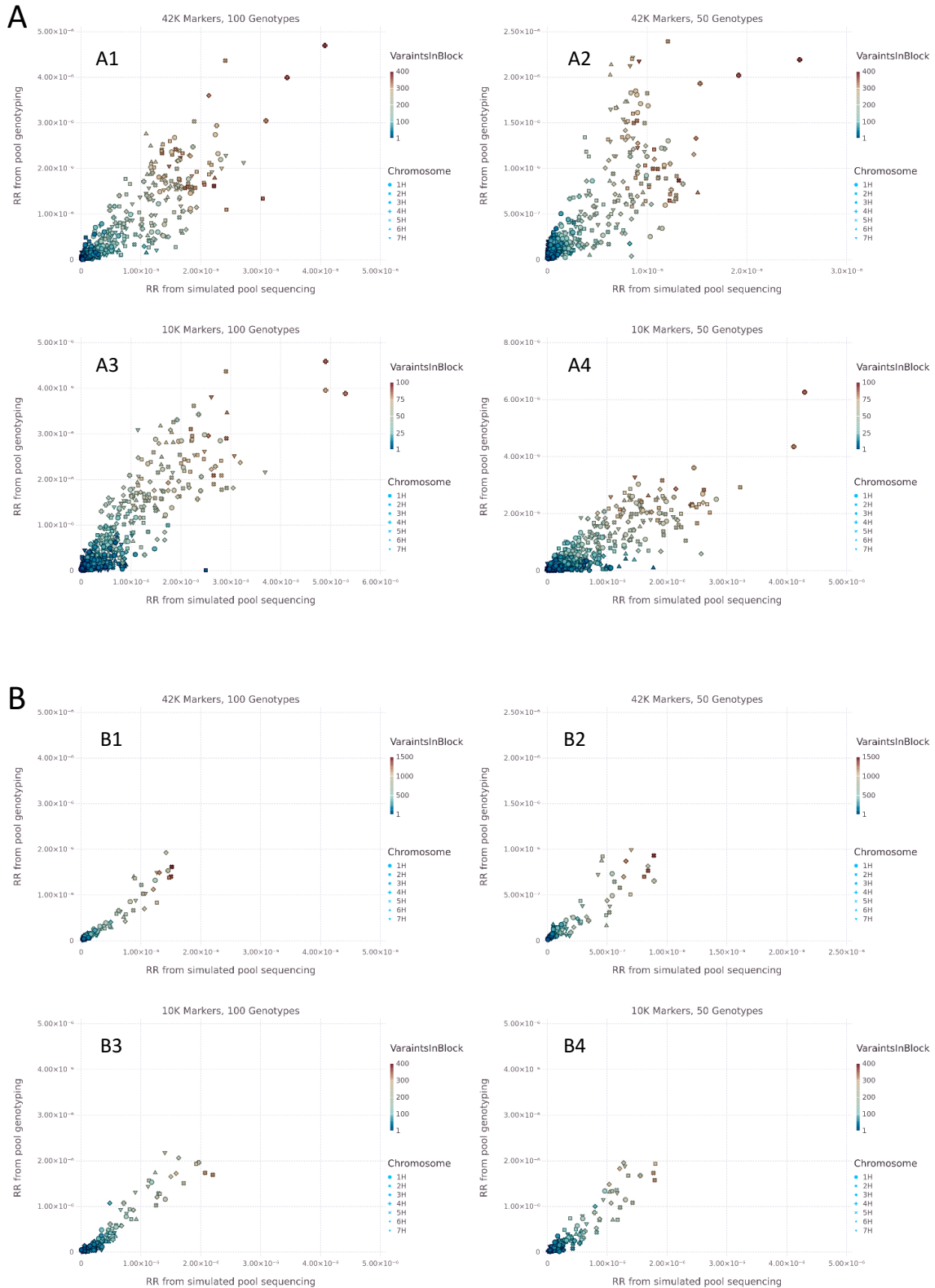
Suppl. Figure 2: Linear model (magenta) and non-linear least square (turquoise) models to predict the impact of a population's size and genotyping depth on the map extension (length). The model estimate on the y-axis is based on the pool genetic map estimation. Each point illustrates an individual population. The dashed line indicates the ideal fit. (step 8.1 & 8.2 in figure 1)



Suppl. Figure 3: Marey map of genetic position (y-axis) against the physical position (x-axis) for all 45 experimental populations. The nPGM (coral) is compared against the HGM (blue). Chromosomes and populations are faceted.

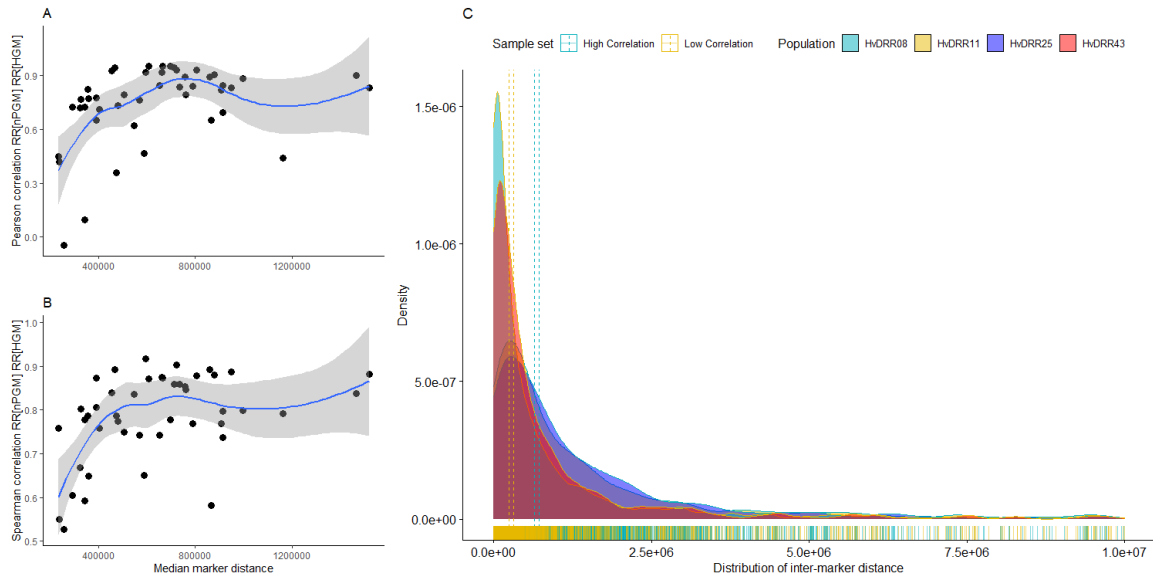


Suppl. Figure 4: The genome-wide general recombination effect for each parental inbred line, computed using a GBLUP model, based on the nPGM genome-wide RR observations.



Suppl. Figure 5: The correlation plot of the RR from pooled genotyping (y-axis) compared to the RR from simulated pool sequencing (x-axis) at 100 reads coverage in 10 MB (**A**) and 50 MB (**B**) genomic windows. Four samples, differing in marker or genotype count, are indicated by the numbers 1 to 4

for both A and B. The number of variants in the genomic windows is indicated by color, while the chromosomes are differentiated by shape.



Suppl. Figure 6: The effect of the median marker distance on the RR_{nPGM} to RR_{HGM} correlation coefficients across all HvDRR populations. **A**– the effect of median marker distance (bp) on the Pearson correlation. **B** – the effect of the median marker distance on the Spearman correlation. **C** – the genome-wide distribution of inter-marker distance (bp) for four HvDRR populations, characterized by a low (yellow, HvDRR08, HvDRR43) and a high (turquoise, HvDRR11, HvDRR25) RR_{nPGM} to RR_{HGM} Pearson correlation.

7 List of publications

1. **Casale, F. A.**, van Inghelandt, D., Weisweiler, M., Li J., and Stich, B. (2022). Genomic prediction of the recombination rate variation in barley – A route to highly recombinogenic genotypes. *Plant Biotechnology Journal*, 20: 676–690.
2. Schneider, M., **Casale, F. A.**, and Stich, B. (2022). Accurate recombination estimation from pooled genotyping and sequencing: a case study on barley. *BMC Genomics*, 23:468.
3. Shrestha, A., Cosenza, F., van Inghelandt, D., Wu, P.-Y., Li, J., **Casale, F. A.**, Weisweiler, M., and Stich, B. (2022). The double round-robin population unravels the genetic architecture of grain size in barley. *Journal of Experimental Botany*, 73: 7344–7361.
4. Cosenza, F., Shrestha, A., van Inghelandt, D., **Casale, F. A.**, Wu, P.-Y., Weisweiler, M., Li, J., Wespel, F., and Stich, B. (2024). Genetic mapping reveals new loci and alleles for flowering time and plant height using the double round-robin population of barley. *Journal of Experimental Botany*, 75: 2385–2402.
5. **Casale, F. A.**, Arlt, C., Köhl, M., Li, J., Engelhorn, J., Hartwig, T., and Stich, B. (2024). The role of methylation and structural variants in shaping the recombination landscape of barley. Submitted for publication.

8 Acknowledgements

I feel thankful to Prof. Dr. Benjamin Stich for his guidance, teaching, support, and patience during this journey.

Sincere thanks to Prof. Dr. Korbinian Schneeberger for agreeing to review this work.

Also, I am grateful for the advice and support of Dr. Delphine van Inghelandt during many stages of the presented experiments.

Special thanks to Dr. Po-Ya Wu for the many times she gave me a hand, including this Latex compilation, and to my dear friend Francesco Cosenza for making my days a lot funnier.

Thanks to my PhD and post-doc colleagues of the "early" days: Ricardo "Ricardinho" Guerreiro, Dr. Marius Weisweiler, Dr. David Ries, Dr. Amaury de Montaigu, Dr. Ruth Freire, Maria Schmidt, and Remy Bastiaan Tjeng. Thanks to my PhD and post-doc colleagues of the "late" days: Francesco Cosenza, Dr. Asis Shrestha, Dr. Michael Schneider, Dr. Po-Ya Wu, Yanrong Gao, Nadia Baig, Christopher Arlt, Marius Kühl, and Dr. Suresh Bonthala. Thanks to all who provided field-related technical support: Florian Esser, Marianne Haperscheid, Isabelle Scheibert, Agata Stoltman, George Alskief, Stephanie Krey, and Konstantin Shek. Special thanks to Ines Sigge for her help in resolving administrative stuff. Also, thanks to all unmentioned members and former members of the Institute for Quantitative Genetics and Genomics of Plants for creating a great and pleasant work atmosphere.

The financial support from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy is gratefully acknowledged. I feel thankful to the Heinrich-Heine-Universität Düsseldorf for its great staff and beautiful facilities.

Thanks to my colleagues in Computomics for the support in finishing this thesis.

Thanks to my dear friends in Europe, USA, and Argentina. Gracias por tolerar la palabra "doctorado" en cada conversación.

Thanks to my dear family: Mami, Papi and Flor. Gracias eternas por aguantarse toda esta lejanía y venirme a visitar. Abu, gracias por llamarme siempre! Tíos, primitos, y suegritos gracias por siempre acordarse de mí y venirme a visitar tan lejos, tantas veces. Los quiero mucho. This is just another Dr. Casale on the list.

Finally, I want to thank my little own family, Maru and Delfi. Gracias por aguantarme trabajando días eternos que fueron años. Piojo, te voy a estar eternamente agradecido.

This work is dedicated to my wife, María Eugenia Bevilacqua.