Genetic factors in *Arabidopsis thaliana* meiotic recombination: Mapping new crossover modifiers and characterizing MutL complexes

M.Sc. Nadia Kbiri

Under the supervision of: Dr. hab. Piotr A. Ziółkowski, prof. UAM

Adam Mickiewicz University, Poznań
Department Of Biology
Laboratory Of Genome Biology

This dissertation is submitted for the degree of:
Philosophy Doctor in Biology

July 2023, Poznań
Poland

Czynniki genetyczne w rekombinacji mejotycznej u Arabidopsis thaliana: mapowanie nowych modyfikatorów crossing-over i charakterystyka kompleksów MutL

M.Sc. Nadia Kbiri

Under the supervision of: Dr. hab. Piotr A. Ziółkowski, prof. UAM

Adam Mickiewicz University, Poznań
Department Of Biology
Laboratory Of Genome Biology

This dissertation is submitted for the degree of:
Philosophy Doctor in Biology

July 2023, Poznań
Poland

Table of content:

Acknowledgments

To my PI, Dr Piotr Ziółkowski,

Thank you for all the guidance and support you have provided me during my time as your student. Your dedication to teaching, unique humor, amazing forever positive attitude, and passion for meiosis and science has profoundly impacted my academic journey. Thank you for your willingness to answer my questions and provide constructive feedback. I am grateful for the knowledge and skills that you have imparted to me.

To my friends and colleagues,

Shoutout to all of you amazing past and present members of the Genome Biology Lab! Special thank you to Weronika my confidante, Longfei my lab brother, Alex my friend and Espe my unicorn. Thank you to Maja, Julia, and Wojtek, my fellow brave PhD soldiers. Thank you to Franz for your support and kindness. Very special thank you to Mrs. Hania, our awesome Lab manager. Thank you to Piotrek and Karolina for being valuable colleagues. Thank you to all my students, Olga, Matti, Natalie, and Marcelina. I learned so much working with you.

Your hard work and collaboration are what make this team so successful and fun to work with. Thank you all for always being there to lend a helping hand and for bringing such positive energy to our work together. I have made what I hope to be lifelong friendships.

To my jury,

Thank you for taking the time to read my doctoral dissertation. I hope it will be to your liking. I look forward to your insight and to having a chance to discuss it with you.

To the administration of UAM faculty of biology and natural sciences doctoral school,

Your administrative expertise and professionalism helped me navigate all various formalities. You have been of great help in addressing my queries and concerns. Thank you very much.

To my family,

Thank you to my amazing parents, brothers, and sister for always being so loving, supportive, and understanding. Thank you to my family back in France for your help and support, during my undergraduate studies, which eventually got me to where I am today. Thank you to my lifelong friends, my family I got to choose, you just rock. Thank you, all of you, for making me the person I am today.

Dedicated to,

The paths never taken

Funding

This work was supported by:

1. The Polish National Science Centre grants: Opus 11 (2016/21/B/NZ2/01757) and Sonata Bis 6 (2016/22/E/NZ2/00455) to PAZ. Preludium 18 (2019/35/N/NZ2/02933) to NK.
2. Foundation for Polish Science grant: Team POIR.04.04.00-00-5C0F/17 to PAZ.
3. The European Union: Passport to the future - Interdisciplinary doctoral studies at the Faculty of Biology, Adam Mickiewicz University in Poznan (POWR.03.02.00-00-I006/17) and POWER17 mini-grant: MG/POWER17/2021/2 to NK.

Scientific collaborations

Scientific experiments were performed at the Laboratory of Genome Biology, Institute of Molecular Biology and Biotechnology, Faculty of Biology, Adam Mickiewicz University, Poznan.

Publications

MSH 2 shapes the meiotic crossover landscape in relation to interhomolog polymorphism in Arabidopsis. Blackwell and Dluzewska et al. The EMBO journal — Sep. 2020, doi.org/10.15252/embj.2020104858.

Quantifying Meiotic Crossover Recombination in Arabidopsis Lines Expressing Fluorescent Reporters in Seeds Using SeedScoring Pipeline for CellProfiler. Kbiri et al. Plant Gametogenesis, Methods in Molecular Biology book series — Apr. 2022, volume 2484, doi.org/10.1007/978-1-0716-2253-7_10.

The effect of DNA polymorphisms and natural variation on crossover hotspot activity in Arabidopsis hybrids. Szymańska-Lejman et al. Nature communications — Jan. 2023, doi.org/10.1038/s41467-022-35722-3.

Abstract

Meiotic crossover recombination events are indispensable for proper chromosome segregation and genetic information mixing. A better understanding of crossover designation and distribution processes is of high interest for breeding and crop improvement. In this work, I use forward and reverse genetics to identify and characterize meiotic crossover recombination factors.

In the first chapter, I briefly introduce the state of the art on meiotic cell division, pro and anti-crossover pathways, and other factors and phenomena that affect crossover frequency and distribution.

In the second chapter, I explore the differences in crossover distribution and frequencies in 5 Arabidopsis bi-parental populations. These populations were obtained from crosses between the reference accession Col-0 and 5 accessions that originate from 5 different climates. My results show that crossover distribution in all the tested accessions follows the same trends with subtelomeric and pericentromeric regions receiving most of the crossover events, but some differences between accessions can still be observed. I also use these populations to map for recombination quantitative loci and was able to identify QTLs in two of the tested populations.

In the third and final chapter, I characterize the effect of MutL genes expression level on Arabidopsis crossover rates in specific genomic intervals. For this, I used commercially available mutants, in-house CRISPR-cas9-mediated deletion mutants, and two different levels of overexpression. My results show that MLH1 and MLH3, but not PMS1, affect recombination frequency in the tested intervals. Cytogenetic characterization shows that MLH1 is indispensable for crossover formation and assurance. Interestingly MLH1 and MLH3 display a dosage stabilization behavior, where both loss of function and excessive overexpression are detrimental for Arabidopsis fertility. Moreover, my results suggest that MutLγ is not exclusive to

class I crossover resolution and that Class I designated intermediates may also be resolved by other endonucleases.

Streszczenie

Rekombinacja crossing-over jest niezbędna do prawidłowej segregacji chromosomów i mieszania informacji genetycznej podczas mejozy. Pełniejsze zrozumienie procesów determinujących zachodzenie zdarzeń crossing-over i ich dystrybucji ma duże znaczenie w hodowli roślin uprawnych. W tej pracy wykorzystuję zarówno podejścia „forward genetics", jak i „reverse genetics", aby zidentyfikować i scharakteryzować mejotyczne czynniki rekombinacji crossing-over. W pierwszym rozdziale pokrótce przedstawiam aktualny stan wiedzy na temat podziału komórek mejotycznych, szlaków pro- i antyrekombinacyjnych, a także innych czynników i zjawisk, które wpływają na częstotliwość i rozkład crossing-over u roślin.

W drugim rozdziale badam różnice w rozkładzie i częstości występowania crossing-over w pięciu populacjach Arabidopsis. Populacje te uzyskano z krzyżówek między linią referencyją Col-0 a pięcioma ekotypami pochodzącymi z pięciu różnych stref klimatycznych. Moje wyniki pokazują, że rozkład crossing-over we wszystkich testowanych populacjach przebiega zgodnie z tymi samymi trendami, przy czym regiony przytelomerowe i okołocentromerowe otrzymują większość zdarzeń crossing-over; nadal można jednak zaobserwować pewne różnice między badanymi populacjami. Krzyżówki te zostały również przeze mnie użyte do mapowania rekombinacyjnych loci cech ilościowych, co pozwoliło mi na zidentyfikowanie QTL w dwóch testowanych populacjach.

W trzecim i ostatnim rozdziale charakteryzuję wpływ poziomu ekspresji genów MutL na częstość crossing-over u Arabidopsis w określonych interwałach genomowych. W tym celu użyłam dostępnych komercyjnie mutantów, uzyskanych przeze mnie za pomocą CRISPR-Cas9 mutantów delecyjnych i dwóch różnych poziomów nadekspresji. Moje wyniki pokazują, że MLH1 i MLH3, ale nie PMS1, wpływają na częstość rekombinacji w testowanych interwałach. Charakterystyka cytogenetyczna wykazała, że MLH1 jest niezbędny do tworzenia i zapewniania

crossing-over. Co ciekawe, MLH1 i MLH3 wykazują silną tendencję do stabilizacji dawki genu, gdzie zarówno utrata funkcji, jak i nadmierna ekspresja są szkodliwe dla płodności roślin. Co więcej, moje wyniki sugerują, że MutLγ nie jest ograniczony do rozdzielania zdarzeń crossing-over klasy I, i że produkt pośredni rekombinacji oznaczony jako klasa I może być również rozdzielany przez inne endonukleazy.

Chapter 1:

General introduction

Chapter 1:

General introduction

I. List of figures:

II. List of tables:

III. Abbreviations:

2n	Diploid
CO	Crossover
D-loop	Displacement loop
dHJ	Double Holliday junction
DMC1	Disrupted meiotic coding DNA
DNA	Desoxyribonucleic acid
DSB	Double strand break
EME1	Essential meiotic endonuclease 1
Exo1	Exonuclease 1
FANCM	Fanconi anemia complementation M
FIGL1	Fidgetin like 1
G1	Growth phase 1
HEI10	Homolog of human enhancer of cell invasion 10
JM	Joint molecule
Mer3	ATP-dependent DNA helicase
MiMC	Microspore mother cell
MLH	Mutator S like homolog
MMC	Megaspore mother cell
MMR	Mismatch repair
Mms4	Methyl methanesulfonate sensitivity 4
Mre11	Meiotic recombination 11
MSH	Mutator S homolog
MUS81	MMS and UV sensitive 81
MutL	Mutator S like
MutS	Mutator S
Nbs1	Nijmegen breakage syndrome
NCO	Noncrossover
NHEJ	Non-hologous end joinning
PCNA	Proliferating cell nuclear antigen
PMS	Post meiotic segregation
PTD	Parting dancers
RAD51	Gamma-radiation hypersensitive 51
RECQ4	RecQ4 helicase
RING	Really interesting new gene
RPA	Replication Protein A
S Phase	Synthesis phase
SC	Synaptonemal complex
SDSA	Synthesis dependent strand annealling
SHOC	Shortage In Chiasmata 1
Spo16	Sporulation 16
ssDNA	Single stranded DNA

SUMO	Small Ubiquitin-like Modifier
Zip1	Zing transporter precursor
ZIP4	TPR repeat-containing protein ZIP4
ZMM	Zip1, Zip2, Zip3, Zip4, Msh4-Msh5, Mer3, and Spo16
ZYP1	Zip homolog 1
ZZS	Zip2, Zip4, Spo16

Meiosis is a reductive cell division characteristic of sexually reproducing organisms. It produces haploid cells that mature into gametes, eggs for females, and sperm or pollen for males. When fused through the process of fertilization, gametes produce a zygote restoring the ploidy level of the parent organism (Figure 1). Meiosis is a very complex process during which the genetic material is first replicated (interphase) then homologous chromosomes are segregated (meiosis I) and finally sister chromatids are separated (meiosis II). Unlike mitosis, which produces two cells identical to the mother cell, spores and so gametes are genetically different from each other and from their progenitor. This is made possible through the random segregation of chromosomes and homologous recombination between maternal and paternal chromosomes. Homologous recombination is one of the most deterministic and defining meiotic phenomena. It takes place during prophase I and is triggered by programmed double-strand breaks (Hunter, 2015; Lam and Keeney, 2015a; Mercier et al., 2015; Wang and Copenhaver, 2018).

Meiotic crossover recombination provides many advantages to sexually reproducing organisms. i. Homologous chromosomes are physically linked through pairing and reciprocal DNA exchanges during meiosis I. This ensures the proper alignment and segregation of homologs. (Mercier et al., 2015; Wang and Copenhaver, 2018; Lloyd, 2022). ii. Crossovers taking place in polymorphic regions can help preventing inbred depression and introduce new alleles and combinations of alleles after each meiotic division. This is true for both selfing and outcrossing scenarios. iii. For organisms that are outcrossing compatible, subsequent crossover events can open the possibility of breaking inter-species barriers and acquiring novel advantageous traits (Feldman and Levy, 2005; Feldman and Levy, 2012; Hollister, 2015; Pelé et al., 2018; Qiu et al., 2020). The evolutionary power of meiosis translates into the observable success of sexually reproducing organisms. Shuffling and accumulation of advantageous traits can boost adaptive abilities. As seen for

angiosperms which represent 89,4% of the referenced vegetal population (Crepet and Niklas, 2009) and the vast majority of the cultivated species.

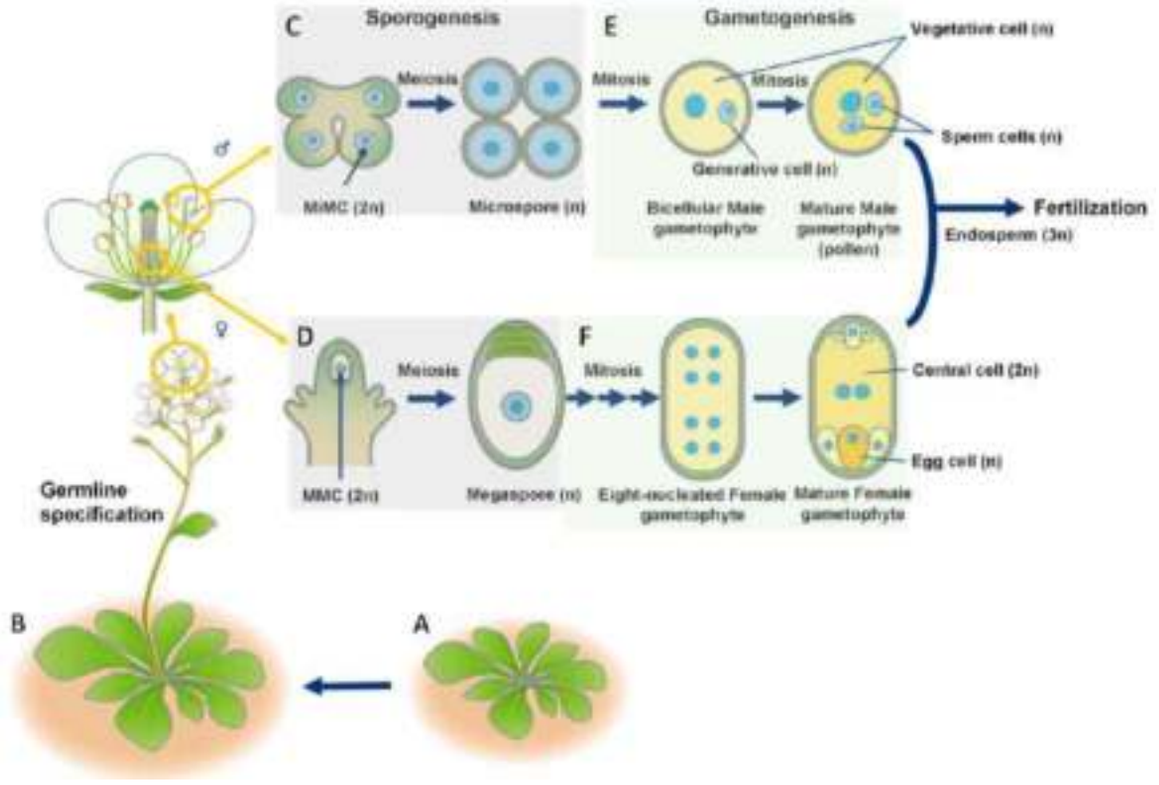


Figure 1. Representation of sporogenesis and gametogenesis in *Arabidopsis thaliana* leading to fertilization. (A-B) The transition from the vegetative to the reproductive stage of *Arabidopsis thaliana* initiates the production of gametes. (C-D) Sporogenesis, male (C) and female (D) representing reproductive organs. Anthers with Microspore Mother Cells (MiMC, 2n) and stigma with Megaspore Mother Cells (MMC, 2n) lead to microspores and megaspores after a meiotic division (1n). (E-F) Gametogenesis, male (E) and female (F) representing mitotic divisions, two for male and three for female, leading to the formation of gametophytes. The fertilization of the female gamete by the pollen grain restores the ploidy level (2n). The ploidy level on the different cells and tissues is represented (xn). Adapted from Ono and Kinoshita, 2021.

1 Meiotic cell division

Germ cells are meiocytes progenitors. In some organisms, like mammals, they are determined as early as embryonic development. Whereas in flowering plants, they are determined later when the transition from vegetative to reproductive stage happens. Meiosis is a lengthy process that can be divided into 3 major stages: Interphase, Meiosis I and Meiosis II (Zamariola et al., 2014; Mercier et al., 2015). It

lasts about 36 hours for *Arabidopsis thaliana* (Armstrong et al., 2003; Armstrong and Jones, 2003).

1.1 Interphase

Interphase is a sequence of 3 stages, growth phase 1 (G1), synthesis phase (S phase) and growth phase 2 (G2). During interphase, germ cells prepare their genetic material for the subsequent divisions. In G1, meiocytes grow in size and produce the necessary transcripts and proteins needed for the S phase. Over the S phase, cells replicate the entirety of their genome by producing neo-synthesized chromatids that stay attached to their sisters at the centromeres using cohesins. G2 phase is used to produce the necessary transcripts and proteins for subsequent cell division (Zamariola et al., 2014).

1.2 Meiosis I

Meiosis I is the first meiotic division through which homologous chromosomes are segregated. It consists of four phases, prophase I, metaphase I, anaphase I, and telophase I (Figure 2B, yellow sector). During prophase I, which is the longest phase of the entirety of the division, lasting around 30 hours in Arabidopsis, chromosomes condense, and homologs pair in preparation for their segregation. Prophase I itself is divided into five sequential stages, leptotene, zygotene, pachytene, diplotene, and diakinesis (Armstrong et al., 2003; Mercier et al., 2015). From leptotene to pachytene, programmed DNA double-strand breaks (DSBs) and meiotic crossover events take place. Crossover events are important for the proper segregation of chromosomes, and in most eukaryotes at least one is required between each pair of homologs (Jones and Franklin, 2006). By late prophase I, chromosomes are fully condensed and paired through the synaptonemal complex (hereafter SC, Hunter and Kleckner, 2001; Hunter, 2003; Hunter, 2015). The migration of the paired chromosomes towards the metaphasic plate marks the end of prophase I and the beginning of metaphase I. The migration and the positioning

of the homologs at the metaphasic plate is orchestrated by the karyokinetic spindle and cytoskeleton. These same structures are also responsible for homologous chromosomes segregation during anaphase I (Bhalla and Dernburg, 2008; Koszul and Kleckner, 2009; Woglar and Jantsch, 2014; Zamariola et al., 2014). The end of anaphase I and the beginning of telophase I are marked by the positioning of the chromosomes at the poles and nuclei separation (Mercier et al., 2015).

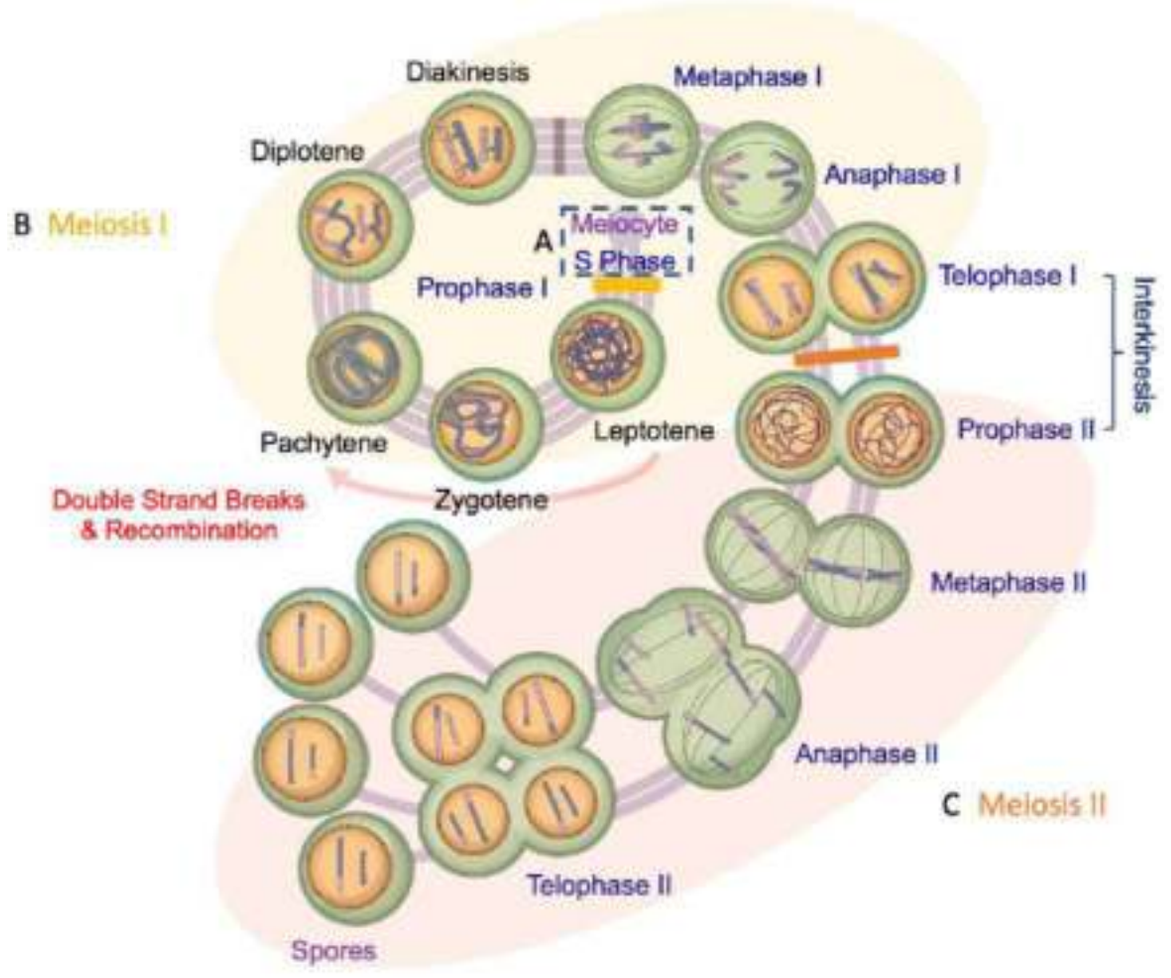


Figure 2. Overview representation of meiosis. (A) Pre-meiosis. In preparation for meiotic division, each meiocyte differentiates, grows, and replicates its genome during the S phase, dashed line blue rectangle. The yellow full rectangle represents the transition from pre-meiosis to meiosis I. **(B) Meiosis I**, yellow sector. Prophase I is the longest-lasting phase of meiotic division, it consists of: **Leptotene**, axis structures form onto chromosomes that start compacting, and homologous pairing. Simultaneously chromosomes are subjected to programmed double-strand breaks and initiate recombination. **Zygotene**, the synaptonemal complex is established and homologous chromosomes start synapsing. **Pachytene**, synapsis is completed, and recombination events are determined. Double strand breaks and recombination occur from leptotene to pachytene, represented with the red arrow. **Diplotene**, the synaptonemal complex is disassembled and homologs are

connected by crossovers. **Diakinesis**, homologous chromosomes are fully compacted and start migrating toward the metaphasic plate. **Metaphase I**, chromosomes are maintained at the metaphasic plate by the spindle. **Anaphase I**, crossovers are resolved, and chromosomes migrate in opposite directions. **Interkinesis**, stage comprising telophase I and prophase II. **Telophase I**, segregated chromosomes are separated in space. The orange full rectangle represents the transition from meiosis I to meiosis II. **(B) Meiosis II**, orange sector. **Prophase II**, chromosomes decondense shortly as they engage in Meiosis II. **Metaphase II**, chromosomes align at the metaphasic plate once again. **Anaphase II**, cohesion is released, and sister chromatids separate and migrate to the poles. **Telophase II**, a tetrad is formed, and the four nuclei are separated. Following cytokinesis, four haploid spores are released. Meiotic phases are represented in blue, sub-phases in black and cell types in purple. The lines connecting the cells represent the evolution of the level of ploidy. Adapted from Mercier et al., 2015.

1.3 Meiosis II

During Meiosis II, sister chromatids are separated. This division culminates in the production of four haploid cells (Figure 2C, orange sector, Mercier et al., 2015). It also consists of four stages, prophase II, metaphase II, anaphase II and telophase II. Telophase I and prophase II are very rapid and are combined into interkinesis. Chromosomes partially decondense before recondensing and positioning at the two metaphasic plates, marking metaphase II. Cohesion between sister chromatids is released initiating their separation and anaphase II. Subsequently, at telophase II, sister chromatids are fully separated, and a tetrad is obtained. Meiosis II ends with cytokinesis and the release of four microspores (Liu and Qu, 2008; Mercier et al., 2015; Ono and Kinoshita, 2021).

Female and male sporogenesis processes are largely similar for *Arabidopsis thaliana*. Alternatively, gametogenesis is rather different, and the two different processes are important to keep in mind for all experiments that are not directly conducted on meiocytes. Indeed, spores undergo mitotic events to produce gametes, three for female and two for male. This is important for factors that are not meiosis-specific. Many factors involved in meiosis are also involved in DNA damage repair, mismatch repair, compaction, stability, chromosome segregation *etc*. All these processes can affect mitosis success and so affect the outcome

observed at post-sporogenesis stages such as gamete, fruit or seed formation. These are products of both maternal and paternal meioses and many mitoses.

2 Meiotic recombination

Meiotic recombination characterizes the genetic material exchanges that take place during early prophase I of meiotic divisions. These exchanges can be reciprocal or non-reciprocal. When two DNA molecules exchange portions of their sequence they yield a crossover event (reciprocal exchange). Meiotic crossover recombination events are specifically defined by this exchange taking place between two homologous chromosomes. When one DNA molecule uses the second one as a template for repair, they yield a noncrossover event (non-reciprocal exchange). Meiotic recombination is very tightly regulated. It is orchestrated by multiple pathways that play antagonistic roles to maintain the number of crossover events at a physiological level (Mercier et al., 2015; Mézard et al., 2015a; Dluzewska et al., 2018; Ziolkowski, 2022).

2.1 Meiotic crossover interference

Crossover recombination events are subjected to different regulations that translate into the phenomena called: crossover assurance, crossover interference, and crossover suppression. Crossover assurance consists of the warranty that each pair of homologs receives at least one crossover event. This is important for proper homolog segregation. Crossover interference consists of distributing crossover events with inter-event distances that are larger than statistically random distribution. This is an indicator of molecular interactions between crossovers to position them within possibly more advantageous regions. Finally, crossover suppression ensures the exclusion of crossovers from centromeres, repetitive sequences, and generally heterochromatin regions. This is believed to shield the organism from genetic instability and possible activation of dangerous

transposable elements (Shinohara et al., 2008; Rosu et al., 2011; Li et al., 2021; Pazhayam et al., 2021; Lloyd, 2022). In a regard to concision and interests, only crossover interference will be detailed.

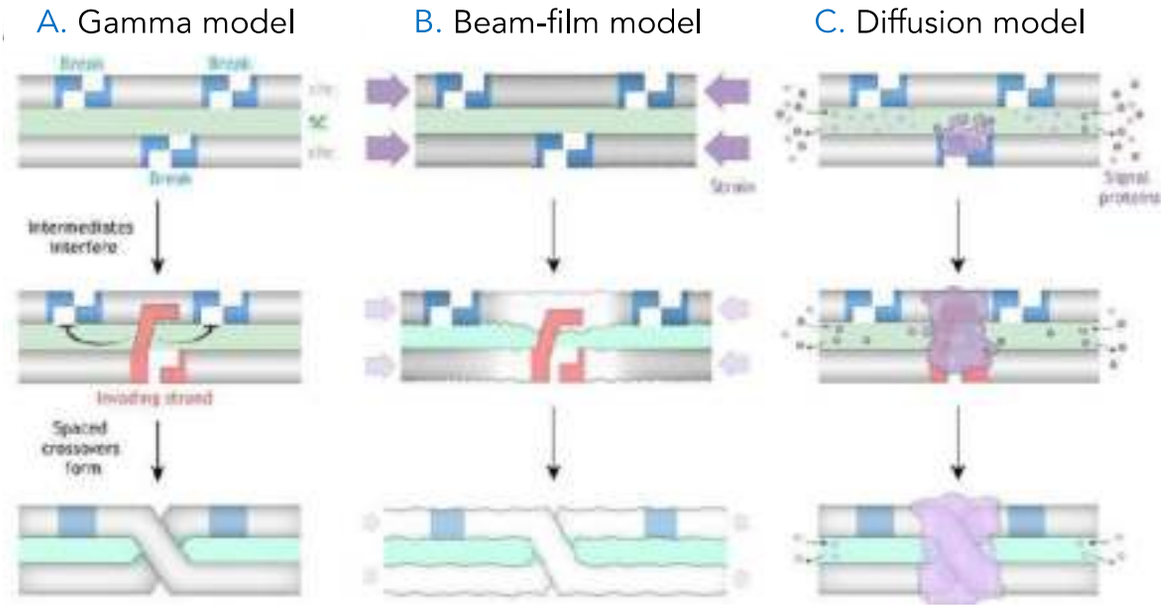


Figure 3. Schematic representation of the mechanistic action of the current models explaining meiotic interference. (A) The statistical Gamma model, where the occurrence of a crossover inhibits the repair of close DSBs into crossovers. (B) The beam-film model, where interference is exercised through mechanical straining of DNA molecules (purple arrows). (C) The diffusion model, where the coarsening of a group of molecules to distinct DSBs and depleting them from other DSBs ensure physical spacing. Adapted from von Diezmann and Rog, 2021.

Interference, when applied to meiotic crossover events, is observed through how these events are more spaced physically than would be expected from random distribution (Jones and Franklin, 2006; Wang et al., 2015; Otto and Payseur, 2019; von Diezmann and Rog, 2021). It is closely related to the synaptonemal complex, where interference is lost with the disruption of this protein structure, most notably through the loss of ZYP1 (Capilla-Pérez et al., 2021; Durand et al., 2022). Although how interference mechanistically operates is still elusive, several models of its *modus operandi* were proposed: the gamma model, the beam-film model, and the diffusion model. The gamma model relies on the statistically normal distribution of crossovers along chromosomes (McPeek and Speed, 1995; Broman and Weber, 2000; Housworth and Stahl, 2003). In this model, the physical occurrence of a

crossover event inhibits the formation of additional crossovers in the close vicinity. The beam-film model is based on the physical strain exercised on the DNA molecules through strand invasion, and joint molecule structures (Kleckner et al., 2004; Zhang et al., 2014). The diffusion model is based on licensing of crossover factors to designated crossover sites and their depletion from other recombination intermediates. Current data in Arabidopsis offers HEI10 as the main player through its coarsening (von Diezmann and Rog, 2021; Morgan et al., 2021; Durand et al., 2022; Lloyd, 2022). These models are complementary and not mutually exclusive. However, the existence of the interference-insensitive crossovers, so-called class II crossovers (see below), discredits some aspects of the gamma and beam-film models giving more weight to the diffusion model. When class II crossovers are uninhibited, through the loss of anti-crossover factors such as FANCM, RECQ4 or FIGL1 (reviewed below in 2.3 Noncrossover and anti-crossover factors), crossover recombination events take place within very close vicinity from each other (Crismani et al., 2012; Girard et al., 2015; Blary et al., 2018; Fernandes et al., 2018; Mieulet et al., 2018; Serra et al., 2018). Purely physical constraints would have a more general unbiased effect independently of the molecules involved in the crossover formation relying solely on steric hindrance.

2.2 Crossover recombination

Meiotic crossover recombination events are produced through several pathways. These pathways are classified as interfering class I crossovers and non-interfering class II crossovers. At the early stages of prophase I, sister chromatids are connected through the formation of the chromosome axis (Zickler and Kleckner, 1999; Hunter, 2015; Zickler and Kleckner, 2015). Simultaneously chromosomes are subjected to programmed DSBs (~ 200) mediated by the Topo VIB-like transesterase SPO11 (Hartung, 2000; Keeney, 2008; Serrentino and Borde, 2012; Lam and Keeney, 2015b; Lam and Keeney, 2015a; Robert et al., 2016). The DSBs are then resected by MRE11/RAD50/NBS1-EXO1 complex to generate 3’ ssDNA

overhangs on both sides of the DSB (Li, 2008; Fernandes et al., 2017). The overhangs are then bound by DMC1 and RAD51 recombinases forming a nucleoprotein filament that can be involved in strand invasion (Hunter, 2015; Lambing et al., 2017). A proportion of the nucleofilaments that were successful at invading the homologous chromosome will form joint molecules (JMs). The formed JMs can either be rejected and dissolved or resolved as recombination events. Recombination events can yield a reciprocal exchange of genetic material, which are called crossovers (hereafter COs). They can also yield a one-way homologous recombination event in which case they are called non-crossovers (hereafter NCOs). In the case where JMs are maintained, they can mature into Holiday junctions (hereafter HJs). HJs are stabilized by RPA and PCNA and progress into double Holiday junctions (hereafter dHJs). Class I or class II recombination machinery is then recruited to dHJs. Class I machinery consists of the ZMM factors and yields exclusively COs (Lynn et al., 2007; Hunter, 2015; Ziolkowski, 2022). Class I COs represent 85% to 95% of the CO events in Arabidopsis. Class II COs are made through several pathways, but the majority is attributed to MUS81-EME1/Mms4 complex. Class II is responsible for 5% to 15% of the total COs in Arabidopsis (Hunter, 2004; Hunter, 2007; Egel and Lankenau, 2008; Hunter, 2015; Mézard et al., 2015b; Zickler and Kleckner, 2015; Dluzewska et al., 2018; Wang and Copenhaver, 2018).

2.2.1 ZMM class I crossovers

Class I crossovers are produced throw the ZMM pathway. ZMM stands for yeast proteins Zip1, Zip2, Zip3, Zip4, Msh4-Msh5, Mer3, and Spo16. It is very widely conserved through Eukaryotes and sexually reproducing organisms that generate crossovers through the class I pathway (Hunter, 2007; Lynn et al., 2007; Hunter, 2015; Mercier et al., 2015; Pyatnitskaya et al., 2019; Ziolkowski, 2022). For specificity, I will hereafter use the Arabidopsis (plants) nomenclature as presented in Table 1:

ZYP1 a and b, SHOC1, HEI10, ZIP4, MSH4-MSH5, MER3, and PTD (in sequence to the names listed previously).

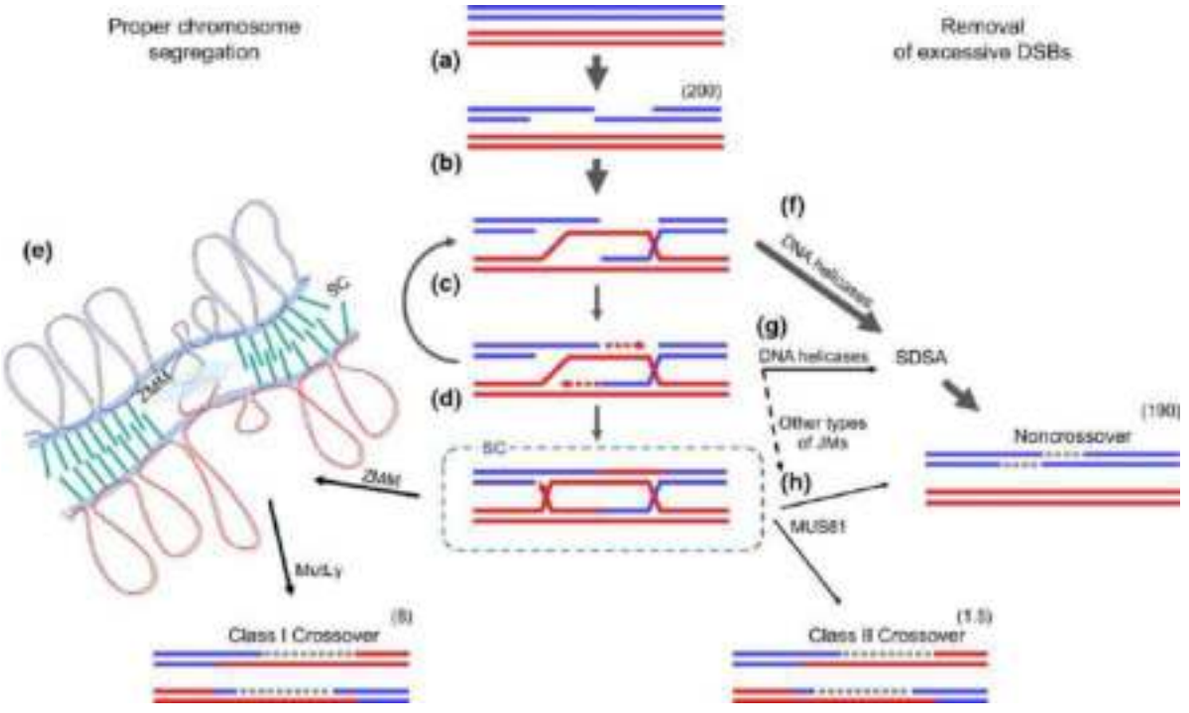


Figure 4. Representation of DSB repair through inter-homolog recombination. (a) Meiotic recombination is initiated with programmed DSB formation and 5′-3′ DNA resection. (b) 3′ single-stranded DNA invades the homologous chromatid and forms a D-loop. (c, d) This can lead to DNA synthesis (dashed red arrows) and second-end capture, which results in dHJ formation. (e) dHJs, when protected by ZMM proteins, will be converted to Class I crossovers by the MutLγ resolvase. This normally takes place in the environment of the SC, which is involved in Class I crossover regulation. (f, g) DSBs that were not processed as crossovers are repaired by pathways leading to synthesis-dependent strand annealing (SDSA), which results in NCOs. (h) A small proportion of JMs (including dHJs) that were not dissolved by helicases can eventually be resolved by MUS81 to produce either a crossover or an NCO. The numbers in brackets indicate approximate estimates of the frequency of each event per Arabidopsis meiosis. The arrow between (d) and (c) indicates recurrent rounds of invasion, extension, and displacement resulting in complex structures; for simplicity, multiple conversion tracts are not shown on (d), and recombination outcomes. From Ziolkowski, 2022.

Arabidopsis has two homologs of Zip1, ZYP1a and ZYP1b. They were identified by homology then characterized and proven to constitute the central element of the synaptonemal complex (SC) (Bleuyard and White, 2004; Higgins et al., 2005). The SC's role is to ensure homologous chromosomes synapsis. It is also involved in the interfering nature of class I crossovers as its loss results in the loss of interference

(Capilla-Pérez et al., 2021; France et al., 2021; Durand et al., 2022). However, strictly speaking, plant ZYP1 may not be considered a ZMM factor. Indeed, the ZMM pathway is still functional *zyp1* null plants, and an increase in crossover number is also observed.

SHOC1, ZIP4 and PTD form a complex that is involved, with MER3 in stabilizing branched DNA. They stimulate displacement loops (hereafter D-loops) into forming joint molecules and serve as a recruitment scaffold for downstream ZMM factors. Their loss of function mutants show a significant decrease in crossover rate and a loss of interference (Osman et al., 2011; Hunter, 2015; Wang and Copenhaver, 2018).

Table 1. List of the yeast ZMM factors, the Arabidopsis homologs, their activities, and functions during crossover formation.

ZMM factor	*A.thaliana* homolog	Activity	Function
Zip1	**ZYP1a and ZYPb**	Coiled-coil protein	Central element of the synaptonemal complex
Zip2	**SHOC1**	Putative XPF endonuclease	Part of the Zip2-Zip4-Spo16 complex (ZZS). Binds branched DNA *in vitro*
Zip3	**HEI10**	RING finger protein	Predicted to be a SUMO and /or ubiquitin E3 ligase. Plays a pivotal role in CO designation.
Zip4	**ZIP4**	TPR motif protein	Part of the ZZS complex. Scaffold protein with multiple protein-protein interactions with ZMM and axis proteins
Msh4	**MSH4**	Mismatch repair family proteins, DNA structure recognition	Form the MutSγ heterodimer MSH4/MSH5. Binds and stabilizes D-loops and joint molecules.
Msh5	**MSH5**		
Mer3	**MER3**	DNA helicase	Involved in DNA heteroduplex stabilization and stimulation of branch migration
Spo16	**PTD**	ERCC1-like protein	Part of the Zip2-Zip4-Spo16 complex (ZZS). Binds branched DNA *in vitro*

HEI10 was identified by homology to be a stand-in for Zip3. It is a RING-finger protein with a predicted SUMO and/or ubiquitin E3 ligase activity. Although its specific activity and targets are still unknown, its pivotal role is very well characterized. Loss of function *hei10* mutant displays a drastic decrease in

recombination rate. It has an additional dosage effect where its expression level correlates positively with the class I recombination events number. Moreover, it also has been shown to have a diffusion / coarsening behavior that involves it in meiotic crossover interference (Ward et al., 2007; Chelysheva et al., 2012; Wang et al., 2012; Ziolkowski et al., 2017; Serra et al., 2018; Morgan et al., 2021; Durand et al., 2022).

Finally, MSH4 and MSH5, which form the MutSγ heterodimer, are meiosis specific mismatch repair (MMR) proteins (Eisen, 1998; Sachadyn, 2010; Han et al., 2022). They form a ring-shaped structure that scans DNA and recognizes recombination intermediates such as displacement loops (D-loops), and joint molecules (Snowden et al., 2004; Lahiri et al., 2018). Its role is to stabilize joint molecules and recruit downstream machinery to resolve them into crossovers. MutSγ is itself regulated through phosphorylation of its degron, which extends its half-life, and sumolaytion of MSH4, which stabilizes it and extends the duration of its presence onto DNA (He et al., 2020; He et al., 2021). Loss of function mutants for MSH4 and MSH5 also display a strong decrease in crossover rate (Higgins et al., 2004; Franklin et al., 2006; Lu et al., 2008; Milano et al., 2019; Desjardins et al., 2020).

Mechanistically, ZMM factors ensure homologous chromosomes synapsis. They stabilize and promote branched DNA to progress into D-loops, Holliday Junctions (HJ) then double Holliday junctions (dHJ) leading to their resolution as a crossover. They are assisted by other factors such as DMC1, RAD51, RPA, PCNA, *etc*. ZMM factors do not resolve dHJ themselves. This step is covered by the MutLγ endonuclease (Hunter, 2015; Mercier et al., 2015; Wang and Copenhaver, 2018).

2.2.2 Class II crossovers

Class II crossovers are mainly described by their non-interfering nature. They are responsible for 10% to 15% of the total crossover events in Arabidopsis (Figure 4). Class II crossovers are mostly attributed to the MUS81 pathway (Hunter, 2007;

Hunter, 2015; Mercier et al., 2015). MUS81 is a structure-specific endonuclease, which in complex with MMS4 and in a RAD52-dependent manner resolves single-end invasions and dHJs into crossovers or non-crossovers (Hunter, 2007; Hunter, 2015). *mus81* loss of function mutants display a 10% decrease of total crossover rate. For Arabidopsis, the additional loss of ZMM-dependent crossovers induces an additional 85% to 90% loss. The remaining about 5% residual crossovers suggest that non-interfering crossovers can also be produced through other, still unknown, pathways (Gerton and Hawley, 2005; Berchowitz et al., 2007; Higgins et al., 2008; Geuting et al., 2009; Macaisne et al., 2011).

2.3 Noncrossover and anti-crossover factors

During early prophase I of a meiotic division, a few hundreds of programmed DSBs are formed. Only a very small portion of these is repaired as crossovers, while the remaining majority is repaired as noncrossovers (Keeney, 2001; Lam and Keeney, 2015b; Mercier et al., 2015). Noncrossover events are non-reciprocal DNA exchanges that are formed through several pathways.

Two major Arabidopsis noncrossover pathways were identified through forward genetic screens, FANCM and RecQ4. *FANCM* is a single-copy gene whereas *RecQ4* exists in two copies in *A. thaliana*, *RecQ4a* and *RecQ4b* (Hartung et al., 2000). Their meiotic roles were discovered through their ability to suppress the *zip4* phenotype. ZIP4 is a core component of the ZMM factors, which are responsible for 85% to 90% of crossovers in Arabidopsis. The *zip4* mutant characteristically displays a dysfunctional meiotic behavior, imbalanced gametes due to low bivalent count and crossover rate, and dramatically decreased fertility (Crismani et al., 2012; Séguéla-Arnaud et al., 2015; Kumar et al., 2019). FANCM and RecQ4 are helicases that are known for their role in suppressing class II crossovers. Loss of function mutants for these helicases translates in a significant increase in class II crossovers, which results in suppression of the *zip4* sterility phenotype. Indeed, cytology shows that class I

crossovers number is unaffected, but the genetic maps are significantly longer (Mannuss et al., 2010; Girard et al., 2015; Séguéla-Arnaud et al., 2015; Blary et al., 2018; Serra et al., 2018; Desjardins et al., 2022). FANCM and RecQ4 induce the dissolution of joint molecules, triggering the synthesis-dependent strand annealing (SDSA) pathway for the repair of the DSBs (Figure 4). Mechanistically, a one-sided recombination event takes places with one homolog using the other as template. Molecularly, the resected DNA sequence is replaced by the DNA fragment synthesized based on a homologous sequence from the homologous chromosome often yielding a conversion tract. Conversion tracts can be mapped when the resynthesized sequence contains distinctive polymorphisms (Chelysheva et al., 2008; Hartung et al., 2008; De Muyt et al., 2012; Pradillo et al., 2014).

Other factors tend to act at the very early stages of DSB processing, frequently affecting the success of strand invasion. This directs these DSBs to be repaired through pathways that do not involve inter-homolog recombination, such as sister chromatid repair. In Arabidopsis, this is distinctively observed with the AAA-ATPase FIGL1. FIGL1 was shown to antagonize chromosome synapsis and influence the turnover of DMC1. By antagonizing homolog synapsis FIGL1 is able to favor sister chromatid instead of inter-homolog recombination. In the absence of FIGL1, the number of class I crossovers is unaffected, but the total crossover count is slightly increased at the genome-wide scale, signifying an increase in the number of class II crossovers.

Additional, less specific and more nuanced, crossover factors are the mismatch repair (MMR) complexes from the MutS family. All these complexes, except the MutSγ complex (MSH4/MSH5), are formed by a heterodimer of MSH2 and another MSH protein (more details in Chapter 3). MSH2-dependent complexes form a sliding clamp that detects post-replicative mismatches and recruits downstream machinery to excise and repair them (Larrea et al., 2010; Jiricny, 2013; Fishel, 2015;

Groothuizen et al., 2015; Han et al., 2022). The canonical role of MMR during meiosis is to prevent recombination between low homology sequences (Figure 5).

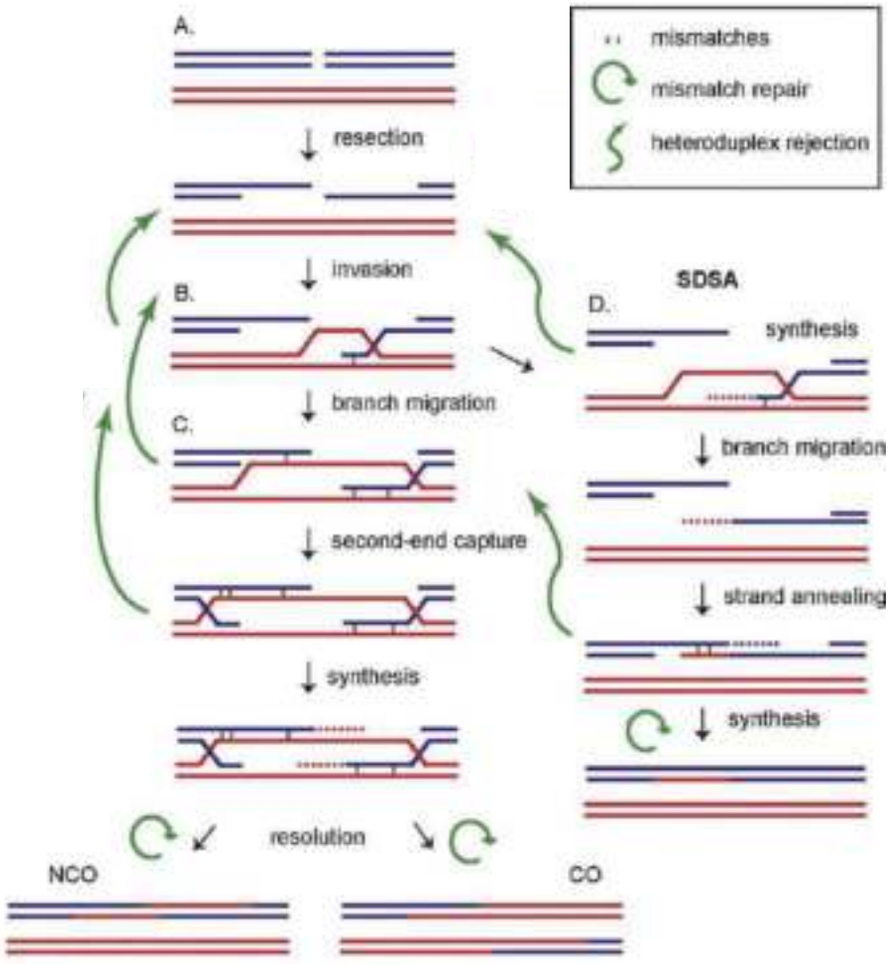


Figure 5. Staling of homologous recombination (HR) by the MMR system. (A) Programmed DSBs are formed and resected. (B) DMC1/RAD51 bind ssDNA forming a nucleoprotein allowing for strand invasion and D-loop formation. (C) Progression of the class I and/or II pathways. Branch migration, heteroduplex extension, second end capture and a secondary D-loop structure take place. DNA is resynthesised on the homologous template and a double Holliday junction is formed. The resolution of these joint molecules leads to the formation of crossover or non-crossover products depending on the cleavage orientation of structure-specific endonucleases. (D) In the SDSA pathway the second end is unlikely to be captured and the extended D-loop is dismantled. The strand invasion and D-loop were maintained long enough for DNA to be resynthesized using the homolog as template. This pathway never results in CO events. Mismatches are potentially formed in regions were donor and recipient DNA molecules anneal (red and blue strands, respectively) either upon strand invasion or after resolution of extended D-loops or Holliday junctions, and are indicated withgreen blocks. Green curly arrows indicate possible heteroduplex rejection events carried out by MMR proteins. Repair of mismatches in maturing heteroduplex regions, resulting in gene conversion, is indicated with circular arrows. Adapted from Tham et al., 2016.

The anti-crossover effect would take place during strand invasion, where mismatches would be detected causing the invasion to be rejected (Kolas et al., 2005; Tham et al., 2016). Yet, it is important to note that in Arabidopsis the MMR system plays a much more complex role where it is also involved in favoring crossovers events in polymorphic regions (Blackwell et al., 2020; Szymanska-Lejman et al., 2023).

3 Research interests and objectives

In the Ziółkowski lab, we are interested in factors influencing meiotic crossover recombination level and pattern, recombination modifiers, environmental stresses, and ploidy. We aspire to unravel the genetic basis and molecular mechanisms governing crossover profile variability. This is to better understand how plants adapt at the genetic level through allelic shuffling and mixing. In the long run, we aim to facilitate the increase in genetic diversity and the introgression of valuable loci in crops.

3.1 Aims and research hypotheses

My doctoral research investigates meiotic crossover recombination level and distribution variability. In the following thesis, I discuss my contribution to the effort towards a better understanding of these variables. In my endeavor, I used two different approaches, forward and reverse genetics. In this chapter, I introduced the current understanding of meiosis and meiotic recombination in plants. In the second chapter, I will discuss how I used five different *Arabidopsis thaliana* accessions to generate segregating populations and investigate the effect of polymorphisms on meiotic recombination. I also used these segregating populations to map recombination quantitative loci. My aim is to investigate the effect of genetic divergence between different Arabidopsis accessions on crossover profiles and eventually identify the underlying genetic factors. In the third chapter,

I will present a genetic characterization of the MutL family genes, *MLH1*, *PMS1* and *MLH3*. The purpose of this characterization is to investigate how different expression levels of the genes coding for MLH1/PMS1 (MutLα) and MLH1/MLH3 (MutLγ) endonucleases affect meiotic crossover recombination. My work focuses mainly on MutLγ, the main class I resolvase. The different results will be concluded on and discussed at the end of each section.

3.2 Biological relevance

Meiotic recombination is at the heart of genetic variability and the ability of sessile organisms such as plants to adapt to their environment. The shuffling and propagation of advantageous traits contributed greatly to the success of the sexually reproducing plants. Domestication induced very high levels of inbreeding and the counterselection of natural variants. This made crops vulnerable to pests and changing climate conditions. Understanding how meiotic recombination operates and is regulated can be implemented into breeding strategies to favor the introgression of interesting traits. It could also bend genetic barriers allowing for directed horizontal genetic flows.

4 Bibliography

Armstrong SJ, Franklin FCH, Jones GH (2003) A meiotic time-course for Arabidopsis thaliana. Sex Plant Reprod **16**: 141–149

Armstrong SJ, Jones GH (2003) Meiotic cytology and chromosome behaviour in wild-type Arabidopsis thaliana. J Exp Bot **54**: 1–10

Berchowitz LE, Francis KE, Bey AL, Copenhaver GP (2007) The Role of AtMUS81 in Interference-Insensitive Crossovers in A. thaliana. PLoS Genet **3**: e132

Bhalla N, Dernburg AF (2008) Prelude to a division. Annu Rev Cell Dev Biol **24**: 397–424

Blackwell AR, Dluzewska J, Szymanska-Lejman M, Desjardins S, Tock AJ, Kbiri N, Lambing C, Lawrence EJ, Bieluszewski T, Rowan B, et al (2020) MSH 2 shapes the meiotic crossover landscape in relation to interhomolog polymorphism in Arabidopsis. EMBO J. doi: 10.15252/embj.2020104858

Blary A, Gonzalo A, Eber F, Bérard A, Bergès H, Bessoltane N, Charif D, Charpentier C, Cromer L, Fourment J, et al (2018) FANCM Limits Meiotic Crossovers in Brassica Crops. Front Plant Sci **9**: 1–13

Bleuyard J-Y, White CI (2004) The Arabidopsis homologue of Xrcc3 plays an essential role in meiosis. EMBO J **23**: 439–449

Broman KW, Weber JL (2000) Characterization of human crossover interference. Am J Hum Genet **66**: 1911–1926

Capilla-Pérez L, Durand S, Hurel A, Lian Q, Chambon A, Taochy C, Solier V, Grelon M, Mercier R (2021) The synaptonemal complex imposes crossover interference and heterochiasmy in Arabidopsis. Proceedings of the National Academy of Sciences **118**: e2023613118

Chelysheva L, Vezon D, Belcram K, Gendrot G, Grelon M (2008) The Arabidopsis BLAP75/Rmi1 Homologue Plays Crucial Roles in Meiotic Double-Strand Break Repair. PLoS Genet **4**: e1000309

Chelysheva L, Vezon D, Chambon A, Gendrot G, Pereira L, Lemhemdi A, Vrielynck N, Le Guin S, Novatchkova M, Grelon M (2012) The Arabidopsis HEI10 is a new ZMM protein related to Zip3. PLoS Genet. doi: 10.1371/journal.pgen.1002799

Crepet WL, Niklas KJ (2009) Darwin's second "abominable mystery": Why are there so many angiosperm species? Am J Bot **96**: 366–381

Crismani W, Girard C, Froger N, Pradillo M, Santos JL, Chelysheva L, Copenhaver GP, Horlow C, Mercier R (2012) FANCM limits meiotic crossovers. Science (1979) **336**: 1588–1590

De Muyt A, Jessop L, Kolar E, Sourirajan A, Chen J, Dayani Y, Lichten M (2012) BLM Helicase Ortholog Sgs1 Is a Central Regulator of Meiotic Recombination Intermediate Metabolism. Mol Cell **46**: 43–53

Desjardins SD, Ogle DE, Ayoub MA, Heckmann S, Henderson IR, Edwards KJ, Higgins JD (2020) MutS homologue 4 and MutS homologue 5 Maintain the Obligate Crossover in Wheat Despite Stepwise Gene Loss following Polyploidization. Plant Physiol **183**: 1545–1558

Desjardins SD, Simmonds J, Guterman I, Kanyuka K, Burridge AJ, Tock AJ, Sanchez-Moran E, Franklin FCH, Henderson IR, Edwards KJ, et al (2022) FANCM promotes class I interfering crossovers and suppresses class II non-interfering crossovers in wheat meiosis. Nat Commun **13**: 3644

von Diezmann L, Rog O (2021) Let's get physical – mechanisms of crossover interference. J Cell Sci. doi: 10.1242/jcs.255745

Dluzewska J, Szymanska M, Ziolkowski PA (2018) Where to Cross Over ? Defining Crossover Sites in Plants. **9**: 1–20

Durand S, Lian Q, Jing J, Ernst M, Grelon M, Zwicker D, Mercier R (2022) Joint control of meiotic crossover patterning by the synaptonemal complex and HEI10 dosage. Nature Communications 2022 13:1 **13**: 1–13

Egel R, Lankenau D-H, eds (2008) Recombination and Meiosis. Recombination and Meiosis. doi: 10.1007/978-3-540-75373-5

Eisen J (1998) A phylogenomic study of the MutS family of proteins. Nucleic Acids Res **26**: 4291–4300

Feldman M, Levy AA (2005) Allopolyploidy - A shaping force in the evolution of wheat genomes. Cytogenet Genome Res **109**: 250–258

Feldman M, Levy AA (2012) Genome Evolution Due to Allopolyploidization in Wheat. Genetics **192**: 763–774

Fernandes JB, Séguéla-Arnaud M, Larchevêque C, Lloyd AH, Mercier R (2018) Unleashing meiotic crossovers in hybrid plants. Proceedings of the National Academy of Sciences **115**: 2431–2436

Fernandes JB, Seguéla-Arnaud M, Larchevêque C, Lloyd AH, Mercier R (2017) Unleashing meiotic crossovers in hybrid plants. Proceedings of the National Academy of Sciences 201713078

Fishel R (2015) Mismatch repair. Journal of Biological Chemistry **290**: 26395–26403

France MG, Enderle J, Röhrig S, Puchta H, Franklin FCH, Higgins JD (2021) ZYP1 is required for obligate cross-over formation and cross-over interference in Arabidopsis. Proc Natl Acad Sci U S A **118**: e2021671118

Franklin FCH, Higgins JD, Sanchez-Moran E, Armstrong SJ, Osman KE, Jackson N, Jones GH (2006) Control of meiotic recombination in Arabidopsis: role of the MutL and MutS homologues. Biochem Soc Trans **34**: 542–544

Gerton JL, Hawley RS (2005) Homologous chromosome interactions in meiosis: Diversity amidst conservation. Nat Rev Genet **6**: 477–487

Geuting V, Kobbe D, Hartung F, Diirr J, Focke M, Puchta H (2009) Two Distinct MUS81-EME1 Complexes from Arabidopsis Process Holliday Junctions. Plant Physiol **150**: 1062–1071

Girard C, Chelysheva L, Choinard S, Froger N, Macaisne N, Lehmemdi A, Mazel J, Crismani W, Mercier R (2015) AAA-ATPase FIDGETIN-LIKE 1 and Helicase FANCM Antagonize Meiotic Crossovers by Distinct Mechanisms. PLoS Genet **11**: e1005369

Groothuizen FS, Winkler I, Cristóvão M, Fish A, Winterwerp HH, Reumer A, Marx AD, Hermans N, Nicholls RA, Murshudov GN, et al (2015) MutS/MutL crystal structure reveals that the MutS sliding clamp loads MutL onto DNA. Elife. doi: 10.7554/elife.06744

Han X-P, Yang X-W, Liu J (2022) MutS and MutL sliding clamps in DNA mismatch repair. Genome Instab Dis **1**: 3

Hartung F (2000) Molecular characterisation of two paralogous SPO11 homologues in Arabidopsis thaliana. Nucleic Acids Res **28**: 1548–1554

Hartung F, Plchová H, Puchta H (2000) Molecular characterisation of RecQ homologues in Arabidopsis thaliana.

Hartung F, Suer S, Knoll A, Wurz-Wildersinn R, Puchta H (2008) Topoisomerase 3α and RMI1 Suppress Somatic Crossovers and Are Essential for Resolution of Meiotic Recombination Intermediates in Arabidopsis thaliana. PLoS Genet **4**: e1000285

He W, Rao HBDP, Tang S, Bhagwat N, Kulkarni DS, Ma Y, Chang MAW, Hall C, Bragg JW, Manasca HS, et al (2020) Regulated Proteolysis of MutSγ Controls Meiotic Crossing Over. Mol Cell **78**: 168-183.e5

He W, Verhees GF, Bhagwat N, Yang Y, Kulkarni DS, Lombardo Z, Lahiri S, Roy P, Zhuo J, Dang B, et al (2021) SUMO fosters assembly and functionality of the MutSγ complex to facilitate meiotic crossing over. Dev Cell **56**: 2073-2088.e3

Higgins JD, Armstrong SJ, Franklin FCH, Jones GH (2004) The Arabidopsis MutS homolog AtMSH4 functions at an early step in recombination: Evidence for two classes of recombination in Arabidopsis. Genes Dev **18**: 2557–2570

Higgins JD, Buckling EF, Franklin FCH, Jones GH (2008) Expression and functional analysis of AtMUS81 in Arabidopsis meiosis reveals a role in the second pathway of crossing-over. The Plant Journal **54**: 152–162

Higgins JD, Sanchez-Moran E, Armstrong SJ, Jones GH, Franklin FCH (2005) The Arabidopsis synaptonemal complex protein ZYP1 is required for chromosome synapsis and normal fidelity of crossing over. Genes Dev **19**: 2488–2500

Hollister JD (2015) Polyploidy: adaptation to the genomic environment. New Phytologist **205**: 1034–1039

Housworth EA, Stahl FW (2003) Crossover interference in humans. Am J Hum Genet **73**: 188–197

Hunter N (2015) Meiotic recombination: The essence of heredity. Cold Spring Harb Perspect Biol **7**: 1–35

Hunter N (2003) Synaptonemal Complexities and Commonalities. Mol Cell **12**: 533–535

Hunter N (2007) Meiotic recombination. Molecular Genetics of Recombination. Springer, Berlin, Heidelberg, pp 381–442

Hunter N (2004) Meiosis. Encyclopedia of Biological Chemistry. Elsevier, pp 610–616

Hunter N, Kleckner N (2001) The Single-End Invasion: An Asymmetric Intermediate at the Double-Strand Break to Double-Holliday Junction Transition of Meiotic Recombination.

Jiricny J (2013) Postreplicative Mismatch Repair. Cold Spring Harb Perspect Biol **5**: a012633–a012633

Jones GH, Franklin FCH (2006) Meiotic Crossing-over: Obligation and Interference. Cell **126**: 246–248

Keeney S (2008) Spo11 and the Formation of DNA Double-Strand Breaks in Meiosis. Recombination and Meiosis. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 81–123

Keeney S (2001) Mechanism and control of meiotic recombination initiation. pp 1–53

Kleckner N, Zickler D, Jones GH, Dekker J, Padmore R, Henle J, Hutchinson J (2004) A mechanical basis for chromosome function. Proc Natl Acad Sci U S A **101**: 12592–12597

Kolas NK, Svetlanov A, Lenzi ML, Macaluso FP, Lipkin SM, Liskay RM, Greally J, Edelmann W, Cohen PE (2005) Localization of MMR proteins on meiotic

chromosomes in mice indicates distinct functions during prophase I. Journal of Cell Biology **171**: 447–458

Koszul R, Kleckner N (2009) Dynamic chromosome movements during meiosis: a way to eliminate unwanted connections? Trends Cell Biol **19**: 716–724

Kumar R, Duhamel M, Coutant E, Ben-Nahia E, Mercier R (2019) Antagonism between BRCA2 and FIGL1 regulates homologous recombination. Nucleic Acids Res **47**: 5170–5180

Lahiri S, Li Y, Hingorani MM, Mukerji I (2018) MutSγ-Induced DNA Conformational Changes Provide Insights into Its Role in Meiotic Recombination. Biophys J **115**: 2087–2101

Lam I, Keeney S (2015a) Mechanism and regulation of meiotic recombination initiation. Cold Spring Harb Perspect Biol. doi: 10.1101/cshperspect.a016634

Lam I, Keeney S (2015b) Mechanism and Regulation of Meiotic Recombination Initiation. Cold Spring Harb Perspect Biol **7**: a016634

Lambing C, Franklin FCH, Wang C-JR (2017) Understanding and Manipulating Meiotic Recombination in Plants. Plant Physiol **173**: 1530–1542

Larrea AA, Lujan SA, Kunkel TA (2010) SnapShot: DNA Mismatch Repair. Cell **141**: 730-730.e1

Li GM (2008) Mechanisms and functions of DNA mismatch repair. Cell Res **18**: 85–98

Li X, Zhang J, Huang J, Xu J, Chen Z, Copenhaver GP, Wang Y (2021) Regulation of interference-sensitive crossover distribution ensures crossover assurance in Arabidopsis. Proc Natl Acad Sci U S A. doi: 10.1073/PNAS.2107543118/-/DCSUPPLEMENTAL

Liu J, Qu L-J (2008) Meiotic and Mitotic Cell Cycle Mutants Involved in Gametophyte Development in Arabidopsis. Mol Plant **1**: 564–574

Lloyd A (2022) Crossover patterning in plants. Plant Reproduction 2022 **1**: 1–18

Lu X, Liu X, An L, Zhang W, Sun J, Pei H, Meng H, Fan Y, Zhang C (2008) The Arabidopsis MutS homolog AtMSH5 is required for normal meiosis. Cell Res. doi: 10.1038/cr.2008.44

Lynn A, Soucek R, Börner GV (2007) ZMM proteins during meiosis: Crossover artists at work. Chromosome Research **15**: 591–605

Macaisne N, Vignard J, Mercier R (2011) SHOC1 and PTD form an XPF–ERCC1-like complex that is required for formation of class I crossovers. J Cell Sci **124**: 2687–2691

Mannuss A, Dukowic-Schulze S, Suer S, Hartung F, Pacher M, Puchta H (2010) RAD5A, RECQ4A, and MUS81 Have Specific Functions in Homologous Recombination and Define Different Pathways of DNA Repair in Arabidopsis thaliana. Plant Cell **22**: 3318–3330

McPeek MS, Speed TP (1995) Modeling interference in genetic recombination. Genetics **139**: 1031–1044

Mercier R, Mézard C, Jenczewski E, Macaisne N, Grelon M (2015) The Molecular Biology of Meiosis in Plants. Annu Rev Plant Biol **66**: 297–327

Mézard C, Tagliaro Jahns M, Grelon M (2015a) Where to cross? New insights into the location of meiotic crossovers. Trends in Genetics **31**: 393–401

Mézard C, Tagliaro Jahns M, Grelon M (2015b) Where to cross? New insights into the location of meiotic crossovers. Trends in Genetics **31**: 393–401

Mieulet D, Aubert G, Bres C, Klein A, Droc G, Vieille E, Rond-Coissieux C, Sanchez M, Dalmais M, Mauxion J-P, et al (2018) Unleashing meiotic crossovers in crops. Nat Plants **4**: 1010–1016

Milano CR, Kim Holloway J, Zhang Y, Jin B, Smith C, Bergman A, Edelmann W, Cohen PE (2019) Mutation of the ATPase domain of mutS homolog-5 (MSH5) reveals a requirement for a functional MutSG complex for all crossovers in mammalian meiosis. G3: Genes, Genomes, Genetics **9**: 1839–1850

Morgan C, Fozard JA, Hartley M, Henderson IR, Bomblies K, Howard M (2021) Diffusion-mediated HEI10 coarsening can explain meiotic crossover positioning in Arabidopsis. Nature Communications 2021 12:1 **12**: 1–11

Ono A, Kinoshita T (2021) Epigenetics and plant reproduction: Multiple steps for responsibly handling succession. Curr Opin Plant Biol. doi: 10.1016/j.pbi.2021.102032

Osman K, Higgins JD, Sanchez-Moran E, Armstrong SJ, Franklin FCH (2011) Pathways to meiotic recombination in Arabidopsis thaliana. New Phytologist **190**: 523–544

Otto SP, Payseur BA (2019) Crossover Interference: Shedding Light on the Evolution of Recombination. Annu Rev Genet **53**: 19–44

Pazhayam NM, Turcotte CA, Sekelsky J (2021) Meiotic Crossover Patterning. Front Cell Dev Biol **9**: 1940

Pelé A, Rousseau-Gueutin M, Chèvre AM (2018) Speciation success of polyploid plants closely relates to the regulation of meiotic recombination. Front Plant Sci **9**: 1–9

Pradillo M, Varas J, Oliver C, Santos JL (2014) On the role of AtDMC1, AtRAD51 and its paralogs during Arabidopsis meiosis. Front Plant Sci. doi: 10.3389/fpls.2014.00023

Pyatnitskaya A, Borde V, De Muyt A (2019) Crossing and zipping: molecular duties of the ZMM proteins in meiosis. Chromosoma **128**: 181–198

Qiu T, Liu Z, Liu B (2020) The effects of hybridization and genome doubling in plant evolution via allopolyploidy. Mol Biol Rep **47**: 5549–5558

Robert T, Vrielynck N, Mézard C, de Massy B, Grelon M (2016) A new light on the meiotic DSB catalytic complex. Semin Cell Dev Biol **54**: 165–176

Rosu S, Libuda DE, Villeneuve AM (2011) Robust crossover assurance and regulated interhomolog access maintain meiotic crossover number. Science (1979) **334**: 1286–1289

Sachadyn P (2010) Conservation and diversity of MutS proteins. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis **694**: 20–30

Séguéla-Arnaud M, Crismani W, Larchevêque C, Mazel J, Froger N, Choinard S, Lemhemdi A, Macaisne N, van Leene J, Gevaert K, et al (2015) Multiple mechanisms limit meiotic crossovers: TOP3α and two BLM homologs antagonize crossovers in parallel to FANCM. Proceedings of the National Academy of Sciences **112**: 4713–4718

Serra H, Lambing C, Griffin CH, Topp SD, Nageswaran DC, Underwood CJ, Ziolkowski PA, Séguéla-Arnaud M, Fernandes JB, Mercier R, et al (2018) Massive crossover elevation via combination of HEI10 and recq4a recq4b during Arabidopsis meiosis. Proceedings of the National Academy of Sciences **115**: 2437–2442

Serrentino M-E, Borde V (2012) The spatial regulation of meiotic recombination hotspots: Are all DSB hotspots crossover hotspots? Exp Cell Res **318**: 1347–1352

Shinohara M, Oh SD, Hunter N, Shinohara A (2008) Crossover assurance and crossover interference are distinctly regulated by the ZMM proteins during yeast meiosis. Nat Genet **40**: 299–309

Snowden T, Acharya S, Butz C, Berardini M, Fishel R (2004) hMSH4-hMSH5 recognizes holliday junctions and forms a meiosis-specific sliding clamp that embraces homologous chromosomes. Mol Cell **15**: 437–451

Szymanska-Lejman M, Dziegielewski W, Dluzewska J, Kbiri N, Bieluszewska A, Poethig RS, Ziolkowski PA (2023) The effect of DNA polymorphisms and natural variation on crossover hotspot activity in Arabidopsis hybrids. Nat Commun **14**: 33

Tham KC, Kanaar R, Lebbink JHG (2016) Mismatch repair and homeologous recombination. DNA Repair (Amst) **38**: 75–83

Wang K, Wang M, Tang D, Shen Y, Miao C, Hu Q, Lu T, Cheng Z (2012) The role of rice HEI10 in the formation of meiotic crossovers. PLoS Genet. doi: 10.1371/journal.pgen.1002809

Wang S, Zickler D, Kleckner N, Zhang L (2015) Meiotic crossover patterns: Obligatory crossover, interference and homeostasis in a single process. Cell Cycle **14**: 305–314

Wang Y, Copenhaver GP (2018) Meiotic Recombination: Mixing It Up in Plants. Annu Rev Plant Biol **69**: 577–609

Ward JO, Reinholdt LG, Motley WW, Niswander LM, Deacon DC, Griffin LB, Langlais KK, Backus VL, Schimenti KJ, O'Brien MJ, et al (2007) Mutation in mouse Hei10, an E3 ubiquitin ligase, disrupts meiotic crossing over. PLoS Genet. doi: 10.1371/journal.pgen.0030139

Woglar A, Jantsch V (2014) Chromosome movement in meiosis I prophase of Caenorhabditis elegans. Chromosoma **123**: 15–24

Zamariola L, Tiang CL, de Storme N, Pawlowski W, Geelen D (2014) Chromosome segregation in plant meiosis. Front Plant Sci **5**: 279

Zhang L, Liang Z, Hutchinson J, Kleckner N (2014) Crossover Patterning by the Beam-Film Model: Analysis and Implications. PLoS Genet **10**: e1004042

Zickler D, Kleckner N (1999) Meiotic Chromosomes: Integrating Structure and Function. Annu Rev Genet **33**: 603–754

Zickler D, Kleckner N (2015) Recombination, Pairing, and Synapsis of Homologs during Meiosis. Cold Spring Harb Perspect Biol **7**: a016626

Ziolkowski PA (2022) Why do plants need the ZMM crossover pathway? A snapshot of meiotic recombination from the perspective of interhomolog polymorphism. Plant Reprod **1**: 3

Ziolkowski PA, Underwood CJ, Lambing C, Martinez-Garcia M, Lawrence EJ, Ziolkowska L, Griffin C, Choi K, Franklin FCH, Martienssen RA, et al (2017) Natural variation and dosage of the HEI10 meiotic E3 ligase control Arabidopsis crossover recombination. Genes Dev **31**: 306–317

Chapter 2:

Study of *Arabidopsis thaliana* natural crossover variability and identification of natural meiotic recombination modifiers

This work was funded by NCN grants Sonata Bis 6 (2016/22/E/NZ2/00455) to PAZ. Preludium 18 (2019/35/N/NZ2/02933) to NK, FNP grant Team POIR.04.04.00-00-5C0F/17 to PAZ, and POWER17 mini-grant: MG/POWER17/2021/2 to NK.

Chapter 2:

Study of *Arabidopsis thaliana* natural crossover variability and identification of natural meiotic recombination modifiers

I. List of abbreviations

"a" x "b"	"a" crossed to "b"
Acc	Accession
bp	Base pair
Cdm	Caldas de Miravete
Co	Coimbra
Col	Columbia
CTAB	Cetyltrimethylammonium bromide
dCAPS	Derived Cleaved Amplified Polymorphic Sequences
DMF	Dimethylformamide
DNA	Deoxyribonucleic acid
DSB	Double strand break
dsDNA	double stranded DNA
dsRED	Discosoma red
EDTA	Ethylenediaminetetraacetic acid
eGFP	Enchanced Green fluorescent protein
F"n"	Filial generation number "n"
FTL	Fluorescent tagged line
GBS	Genotyping by sequencing
kb	Kilobase
L*er*	Lansberg *erecta*
LOD	Logarithm of the odds
Mbp	Mega base pair
MMR	Mismatch repair
NASC	Nottingham Arabidopsis Stock Centre
Neo	Neo-Shahdara
Oy	Oystese
PCR	Polymerase chain reaction
Per	Perm
QTL	Quantitative trait locus
RNA	Ribonucleic acid
SNP	Single nucleotide polyporphism
TIGER	Trained Individual GenomE Reconstruction

II. List of figures

III. List of tables

IV. List of equations

V. List of Supplemental figures

VI. List of supplemental tables

1 Introduction

Meiotic recombination, the process by which reciprocal, crossovers, and non-reciprocal, non-crossover, exchanges of genetic material are made, is believed to be an evolution driver but also to be subjected to natural selection (Bomblies et al., 2015; Henderson and Bomblies, 2021). The genetic material exchanges take place during the first prophase of a meiotic division and are very deterministic in the successful production of spores. During anaphase I, chromosomes segregate randomly which constitutes the first level of mixing between maternal and paternal genetic material. Simultaneously, in order to ensure proper segregation, these same chromosomes undertake at least one crossover event per pair of homologs (Jones and Franklin, 2006; Mercier et al., 2015; Wang et al., 2015). These reciprocal genetic exchanges can create new combinations of alleles and sometimes new alleles. This constitutes the second level of mixing. It can ultimately create advantageous allelic combinations allowing, with successive generations, for heritable changes to propagate in a given population (Figure 1), (Egel and Lankenau, 2008; Zamariola et al., 2014; Hunter, 2015; Zickler and Kleckner, 2015; Wang and Copenhaver, 2018). Additionally, meiotic factors are subjected to natural selection under the pressure of the growth environment. In fact, meiotic processes, including recombination, are sensitive to environmental conditions such as temperature and biotic stress (Hamilton et al., 1990; Fischer and Schmid-Hempel, 2005; Salathé et al., 2008; Lasky et al., 2014; de Storme and Geelen, 2014; Bomblies et al., 2015; Morgan et al., 2017; Lloyd et al., 2018; Henderson and Bomblies, 2021).

In this chapter, I use five biparental populations, produced using the cross of five *Arabidopsis thaliana* accessions from different climates to the reference Col accession (Figure 4 and Figure 5). To investigate the crossover distribution at the whole genome scale and the correlation between crossover rate and single nucleotide polymorphism (hereafter SNP) density. Moreover, I use these five populations to map for recombination quantitative trait loci (hereafter QTLs). The

chosen quantitative trait is the crossover recombination frequency within a 5.1 Mbp subtelomeric region of chromosome 3, the 420 interval.

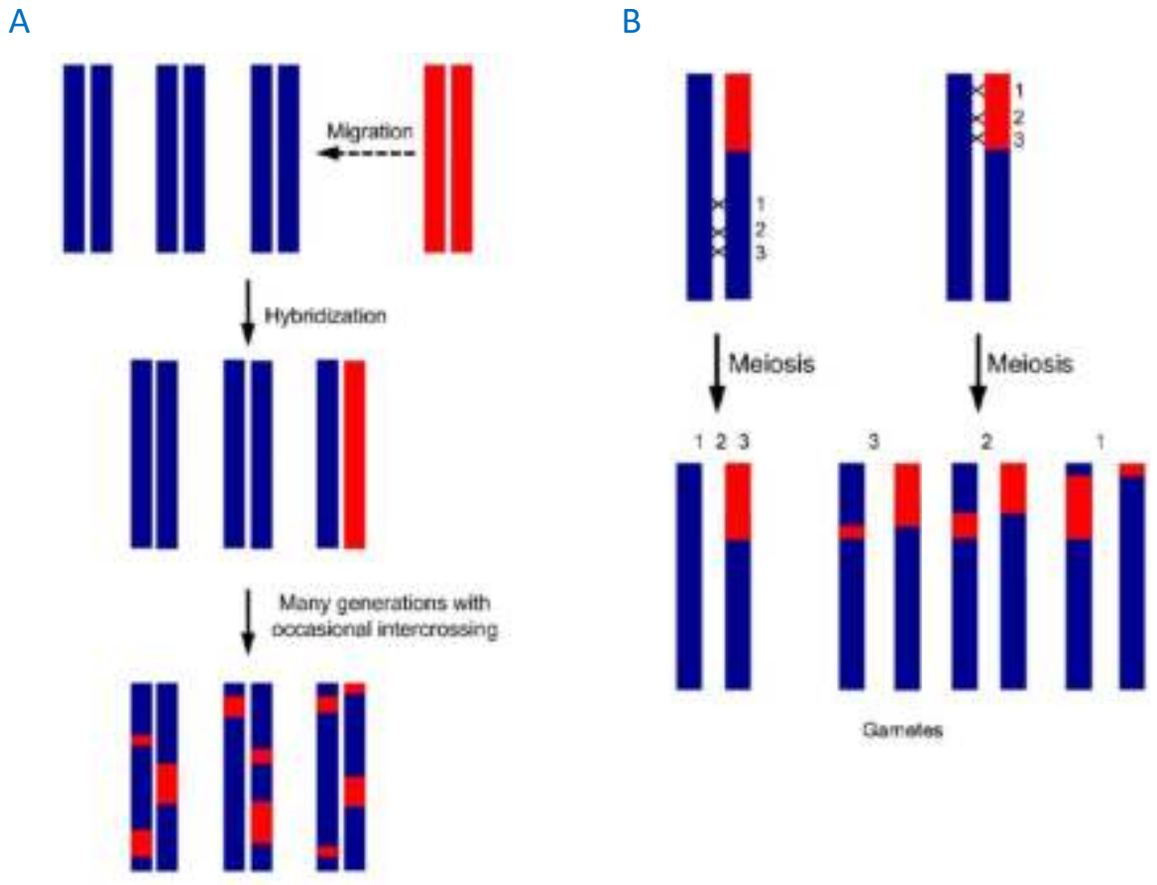


Figure 1. Introduction of heterozygosity through outcrossing and evolutionary advantage of favoring recombination events within heterozygous regions. Single chromosome pairs are presented for simplicity. Inbred selfing population is represented in blue and diverged donor is represented in red. A. Occasional outcrossing itroduces a new pool of alleles into a population through hybridization then propagation thanks to additional outcrossing events through generations. B. Crossover events taking place within homozygous regions do not contribute to the creation of novel alleles and genetic combinations. Whereas crossover event taking place within heterozygous regions allow for the production of novel genetic combinations that can be adavantegous for the adaptatbility of an organism. Adapted from Ziolkowski, 2022.

1.1 Intrinsic and extrinsic meiotic recombination effectors

Crossover recombination events are a limited outcome of the programmed double strand breaks (DSBs) that occur during early meiosis. Their primary role at the physical level is to ensure the proper pairing and segregation of homologous chromosomes (Hunter, 2007; Mercier et al., 2015). This role aligns with the fact that across species, crossover events tend to number from 1 to 3 events per pair of

homologs (Fernandes et al., 2018). However, when mapped along chromosomes, crossover events are not evenly distributed and tend to cluster toward distal regions (subtelomeres) of chromosomes (Haenel et al., 2018), sometimes toward pericentromeric regions as well, as seen in Arabidopsis (Figure 2), (Ziolkowski et al., 2017; Lawrence et al., 2019; Blackwell et al., 2020; Zhu et al., 2021). This behavior reflects the fact that crossovers are not randomly positioned and are subjected to regulation. These regulations are exercised at different levels, from sequence polymorphisms to chromatin state and structure to recombination factors all the way to environmental conditions.

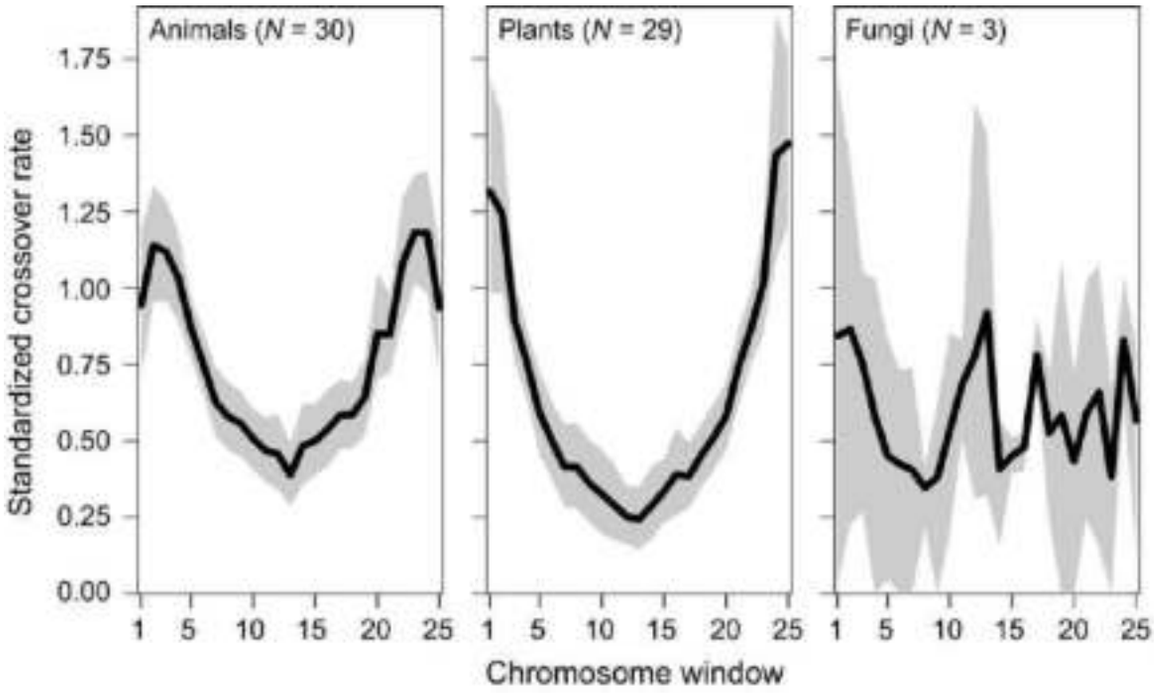


Figure 2. Representations of strandarized crossover distribution along chromosomes in three eukaryote types: animals, plants, and fungi. Chromosomes were scaled and divided into 25 windows. Average crossover number events are represented in black, and the 95% bootstrap confidence bands are shaded in grey. Both animal (N = 30 species) and plant (N = 29 species) kingdoms show a prevalence for crossover to take place in subtelomeric regions. The representation for fungi is less clear due to the low sample pool (N = 3 species). Figure rights belong to Haenel et al., 2018.

1.1.1 DNA sequence and chromatin

From the evolutionary standpoint, targeting crossover events to the more polymorphic regions is more advantageous as it is the only way to create

potentially more advantageous genetic combinations (Figure 1), (Felsenstein, 1974; Ziolkowski and Henderson, 2017; Wang and Copenhaver, 2018; Ziolkowski, 2022). Noticeably, crossover events taking place in homozygous regions will yield gametes with genetic sequences identical to the parental sequences. Whereas crossover events taking place in heterozygous regions will produce gametes with novel genetic combinations.

At the megabase scale, distribution of recombination events does not seem to be significantly affected by polymorphism density and distribution as is it similar between hybrid and inbred contexts (Lian et al., 2022). This is with the exception of large rearrangements such as inversions, where recombination is inhibited. The effect of sequence polymorphism is rather complex and varies according to the viewpoint selected. I will discuss it in more detail in section "*1.2 Meiotic crossover recombination and polymorphism*".

Crossover recombination events are a subset outcome of DSB repair processes. DSBs are programmed events that happen in early meiosis and are subjected to physical constraints introduced by the opportunistic nature of the SPO11 complex. In fact, DSBs are excluded from heterochromatin, which is characteristically enriched in repressive epigenetic marks, *e.g.* high nucleosome occupancy, repressive histone marks, DNA methylation, *etc.* (Takeda and Paszkowski, 2006; Yelina et al., 2015a; Ziolkowski and Henderson, 2017; Choi et al., 2018; Ono and Kinoshita, 2021). These marks are commonly found in centromeres, transposable element rich regions and highly repetitive sequences. Consequently, at the chromatin level, DSBs and recombination events are more abundant in euchromatic regions, which are enriched in activating epigenetic marks and coding genes (Choi et al., 2018).

1.1.2 Molecular effectors

Crossover recombination events are produced through several processes which can show interference. Crossover recombination interference is the peculiarity by which

the presence of a crossover event inhibits the occurrence of a second one in close vicinity (Jones and Franklin, 2006; Wang et al., 2015). Crossovers are commonly classified as interfering and non-interfering, respectively class I and class II. Class I crossovers are generated through the ZMM pathway, they represent around 90% of the crossover events in Arabidopsis (Figure 3A). Class II crossovers are generated by several pathways but are mostly attributed to the MUS81 pathway. In Arabidopsis' case, the majority of crossovers experience interference and their physical distribution is affected by the mode of action of the ZMM pathway. Additionally, recombination events correlate positively with the physical length of the axis and synaptonemal complex (Feldman et al., 1996; Ruiz-Herrera et al., 2017). These are the protein structures assembled onto chromatin to facilitate chromosome pairing and synapsis. Moreover, several meiotic factors show preferential positioning of crossovers. To enumerate a few examples, DMC1, a meiosis-specific recombinase, but not RAD51, a ubiquitous recombinase that is also involved in meiosis, shows a preference for mismatched nucleotides. Both are involved in the strand invasion stage of meiotic recombination, suggesting a bias towards more polymorphic regions (Lee et al., 2015). HEI10, the ZMM E3 SUMO/ubiquitin ligase, shows a preference for positioning recombination events in distal regions as demonstrated in the overexpression context by Ziolkowski et al., 2017. Plants overexpressing *HEI10* show a significant increase in crossover events in distal regions but not in pericentromeric regions. This preference for distal subtelomeric regions is also observed for the class II crossover repressive DNA helicases FANCM and RECQ4, as well as HCR1 and 2. HCR1 is a phosphatase that inhibits both class I and class II crossovers. HCR2 is a heat-shock binding protein that limits recombination by limiting *HEI10* transcription. Indeed, loss of function for their coding genes in Arabidopsis shows an increase in crossover rate with a clear preference for subtelomeres (Crismani et al., 2012; Fernandes et al., 2018; Mieulet et al., 2018; Serra et al., 2018b; Nageswaran et al., 2021; Kim et al., 2022). The absence of MSH2, the core component of MutS complexes in MMR, flattens

out the crossover distribution along chromosome arms, increasing the frequency of events in interstitial regions and weakening the preference for subtelomeric and pericentromeric regions observed in the wildtype context as shown in (Blackwell et al., 2020). SNI1, a component of the condensin-like SMC5/6 complex which is known for its role in DNA damage repair, is also involved in the control of meiotic crossover. The absence of SNI1 both increases and shifts crossovers towards subtelomeric regions (Zhu et al., 2021).

1.1.3 Environmental conditions

Meiotic crossover recombination as an evolution driver is also subjected to the pressure applied by the environmental conditions (Felsenstein, 1974; Dapper and Payseur, 2017; Dumont, 2020; Henderson and Bomblies, 2021; Protacio et al., 2022). Indeed, crossover rate responds to external stimuli at the short term as observed for the increased recombination frequency in response to higher and lower temperatures (Bomblies et al., 2015; Morgan et al., 2017; Lloyd et al., 2018). Also, when subjected to biotic stress, hosts showed an increase in recombination frequency that is believed to favor their adaptation to the pool of pathogens. This coevolutionary relationship is called the Red Queen dynamic, it is also believed to select for an increased crossover rate (Hamilton et al., 1990; Fischer and Schmid-Hempel, 2005; Salathé et al., 2008; Brockhurst et al., 2014; Lasky et al., 2014; de Storme and Geelen, 2014; Henderson and Bomblies, 2021). The Red Queen hypothesis remains however controversial. Indeed, the interpretation of the obtained data is made complicated by the large number of affected variables during host/pathogen interactions. More interestingly, interpopulation recombination frequency was shown to be subjected to natural selection. Significant differences at the whole genome scale were observed between two natural populations of *Drosophila pseudoobscura* from Utah and Arizona, USA (Samuk et al., 2020). A similar effect was observed for Arabidopsis, where chiasmata

count showed up to 22% variation between nine different accessions originating from different geographical and ecological origins (López et al., 2012).

1.2 Meiotic crossover recombination and polymorphism

Meiotic recombination is affected by many effectors such as chromatin structure, effectors activity and environmental conditions. One of the most studied effectors is the level of genetic divergence embodied by polymorphism density (Ziolkowski and Henderson, 2017; Dluzewska et al., 2018). Polymorphisms affect recombination frequency in different fashion depending on their nature (indels, inversions, SNPs, etc.), the chosen perspective (whole genome vs hotspot) and the level of heterozygosity. In this section I will briefly discuss three contexts: the full hybrid situation, the heterozygosity in cis and the hotspot scale.

1.2.1 Hybrid effect

Hybridization occurs through a cross between two individuals of the same species, closely related species, or genetically compatible species. It is an extensively studied phenomenon because of the hybrid vigor effect, also known as heterosis (Timberlake, 2013). Often, F1 hybrids characteristically show better fitness in comparison to their parents. This phenotype relies on the fact that, through outcrossing, the recessive deleterious alleles accumulated through inbreeding are dominated by the newly acquired alleles or both paternal alleles being over-dominated by the novel combinations. The hybrid context is of utmost interest for breeders, both for the heterosis effect and for the possibility of introgressing advantageous genetic traits into an organism of interest (Lippman et al., 2007; Lippman and Zamir, 2007; Ben-Ari and Lavi, 2012; Timberlake, 2013). Meiotic recombination is sensitive to heterozygosity and can be affected by it. Investigating the crossover rate at the whole genome scale relies on the presence of SNPs, which allows for discriminating between parent donors. This made the comparison of crossover rate and distribution, at the whole genome scale, between inbred parents and their hybrid progeny impossible for a long time. Recently however, Lian et al.,

2022 overcame this limitation in Arabidopsis. By introducing a limited but crossover mapping-sufficient number of SNPs in the parental lines they were able to quantify crossover event number and distribution in the parents and compare these to those of the hybrid progeny. Their results show that at the whole genome scale, SNP density and distribution do not affect the distribution of crossover events. However, the crossover rate in female meiocytes is significantly lower than in both parents, whereas the male meiocytes showed an intermediate rate between both parents (Lian et al., 2022). These observations are in agreement with previous findings, at lower scale, where meiotic recombination is downregulated by heterozygosity, predominantly by large genetic rearrangements such as large indels, translocations and inversions (Dooner, 1986; Lichten et al., 1987; Jeffreys and Neumann, 2005; Baudat and de Massy, 2007; Cole et al., 2010; Schwander et al., 2014; Thompson and Jiggins, 2014).

1.2.2 Heterozygosity in cis

Heterozygosity in cis or juxtaposition effect is a very fascinating phenomenon where a heterozygous region, at the Mbp scale, receives more crossover events at the expense of the homozygous region juxtaposed to it (Figure 3B). Indeed, Ziolkowski et al. showed in 2015 that when backcrossed to Col or Ct, Arabidopsis Col/Ct nearly isogenic lines (NILs), where the subtelomeric part of the north arm of chromosome 3 is fixed for Ct and the rest for Col and *vis versa*, favor recombination in the heterozygous part. They additionally showed that this phenomenon can be observed in different regions of Arabidopsis genome, like pericentromeres. It was also seen when using different Arabidopsis accessions (Ziolkowski et al., 2015; Lawrence et al., 2019). This is a very important discovery that comforts the evolutionary role of meiotic recombination as a driver for the creation of novels alleles and combinations of alleles. This is further confirmed by the fact that the heterozygosity in cis effect is lost in the absence of *MSH2* (Blackwell et al., 2020). MSH2 is a mismatch repair (MMR) protein. It is a component of all MMR MutS

heterodimers. MutS complexes are responsible for detecting post-replication mismatches and recruiting downstream machinery to repair them (further details in Chapter 3). The loss of juxtaposition effect in the *msh2* null mutant shows that the bias directing crossover events towards the heterozygous region is driven by the mismatches detected by the MMR system.

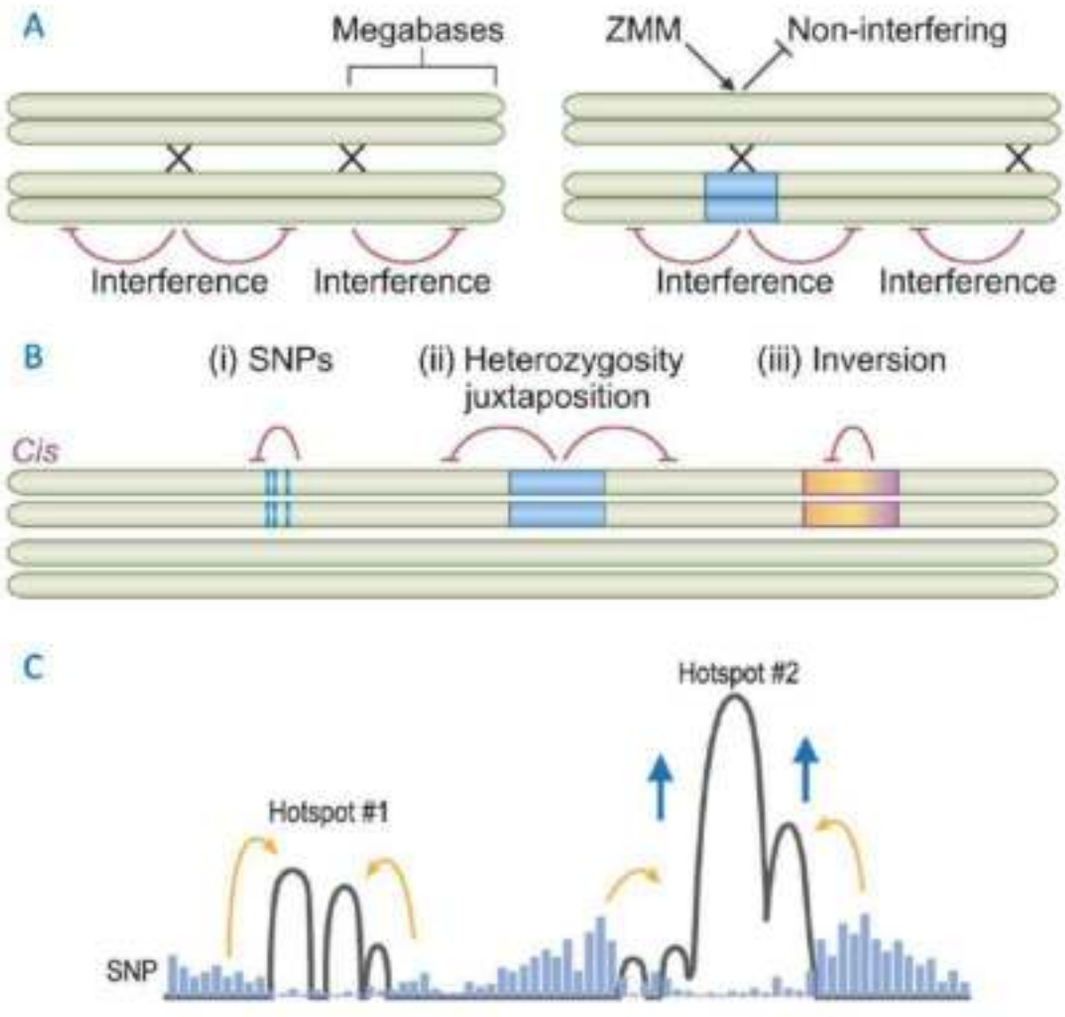


Figure 3. Representation of intrinsic and cis-acting meiotic recombination effectors. Chromosomes are represented in green, polymorphisms with blue regions and blue bars, inversion in a scale of orange hue A. Meiotic interference at the megabase scale. The occurance of a crossover event inhibits other crossover events to take place in close vicinity and ZMM as an interfering pathway that favors recombination events within more polymorphic regions. B. Polymorphisms can inhibit crossovers from taking place in adjascent homozygous regions as for SNPs and herozigosity in cis, and within themselves as in inversions. C. Representation of SNPs as drivers for recombination events to take place within hotspots. Adapted from Ziolkowski and Henderson, 2017 and Szymanska-Lejman et al., 2023.

1.2.3 Polymorphisms at the hotspot scale

Recombination hotspots are small stretches of DNA, within the kilobase scale, that predominantly experience crossover events with a rate that can be up to a thousand times higher than the adjacent regions (Figure 3C). They were first discovered in humans in 1982 (Orkin et al., 1982), but have since been found in many organisms of different taxa (Serrentino and Borde, 2012; Choi et al., 2013; Choi and Henderson, 2015; Brick et al., 2018; Protacio et al., 2022; Szymanska-Lejman et al., 2023).

Meiotic recombination hotspots are determined by their DNA sequence, and chromatin structure. Indeed, recombination hotspots tend to map to DSB hotspots. This shows that the accessibility of chromatin to SPO11 complexes and recombination machinery is part of their designation process. In mammals, DNA is made accessible thanks to zinc-finger protein, PRDM9, that recognizes and binds specific DNA sites (Dluzewska et al., 2018; Paiano et al., 2020). In yeast and plants, no PRDM9 homolog was identified. Instead, a similar process is observed where recombination hotspots map to low nucleosome occupancy in yeast (Berchowitz et al., 2009; Pan et al., 2011), and H2A.Z enriched nucleosomes in Arabidopsis gene promoter regions (Choi et al., 2013). Both low nucleosome occupancy and H2A.Z loading are signatures of open and active chromatin.

Recent work by Szymanska-Lejman et al., 2023 show that Arabidopsis meiotic recombination hotspot activity is negatively affected by genetic rearrangements such as indels within the hotspot but not by adjacent rearrangements within 7kb radius. A similar effect was observed in mice, where indels within the hostpot create a CO refractory zone but the rest of the hotspot remains active (Cole et al., 2010). Moreover, crossover hotspots characteristically show a higher SNP density than average suggesting again that they are driven by heterozygosity. This is further confirmed by decreased activity of a highly polymorphic hotspot with the loss of *MSH2*, as observed previously with the heterozygosity in cis effect (Ziolkowski et

al., 2015; Blackwell et al., 2020). Finally, and interestingly, whilst preference for polymorphic regions is favored for recombination events, the actual crossover break point seem to happen within the almost homozygous stretches within the heterozygous regions. This intriguing phenomenon was observed in other Arabidopsis hotspots and for other organisms such as mice (Cole et al., 2010; Choi et al., 2016; Serra et al., 2018a).

1.3 Natural variability of meiotic factors

Several forward genetics investigations unraveled that meiotic molecular factors were subjected to natural selection and developed genetic variation that translated into quantitative effects at the recombination rate and distribution level. *e.g.* using a classic genetic marker genotyping approach, the cross between Col and L*er* allowed for isolating two strong QTLs that showed up to be the meiosis-specific E3 ligase HEI10 (Chromosome 1) and the ubiquitous SNI1 (Chromosome 4) a subunit of the condensin-like SMC5/6. The Col alleles for *HEI10* and *SNI1* showed up to be semi-dominant and dominant respectively, with the L*er* allele conferring a lower recombination rate for *HEI10* and a higher recombination rate for *SNI1*. Interestingly, these genes did show SNPs discriminating the Col allele from the L*er* allele. For HEI10 the observed effects were linked to an R264G substitution in its last exon (Ziolkowski et al., 2017). Expression level quantification did not show a significant difference between the Col and L*er* alleles. Nevertheless, we know that *HEI10* displays a dosage effect where the expression level correlates positively with the crossover rate (Ziolkowski et al., 2017). As for *SNI1,* the L*er* allele displays a similar phenotype as the *sni1* weak mutant suggesting that the increased recombination rate observed is due to a decreased activity. Here again, the expression levels of the Col and L*er* alleles did not show any significant differences (Zhu et al., 2021).

Using a combination of genotyping by sequencing (GBS) and classic genetic marker mapping, *TAF4b* was identified as causative for a recombination QTL on

chromosome 1 of Arabidopsis Col/Bur cross. Bur is an Arabidopsis accession originating from the British Isles. TAF4b is a subunit of the TFIID complex, a generic transcription factor involved in the recruitment of RNA-polymerase II. The Col allele of *TAF4b* showed a dominant behavior with the Bur allele conferring a lower recombination rate. The Bur *TAF4b* allele shows a similar phenotype to that of the *taf4b* mutant, suggesting a decreased functioning. As a subunit of a large transcription factor, the current model is that TAF4b influences meiotic recombination through its ability to bind DNA and making chromatin more accessible for SPO11 complexes and recombination machinery. It was also shown to directly and/or indirectly influence the expression level of major meiotic genes such as *MSH5, REC8* and *ATM* (Lawrence et al., 2019).

Natural variants influencing meiotic recombination were also identified using Genome Wide Association Studies (GWAS). A recent study on barley, using crosses between domesticated barley and 25 wild accessions, suggested *REC8* as a candidate natural modifier. *REC8* codes for a meiosis specific cohesion subunit (Dreissig et al., 2020). An investigation in cattle identified multiple genes, *HFM1, MLH3, MSH4, MSH5, RNF212*, and *RNF212B* as recombination quantitative loci. A similar study identified *REC8* and *RNF212B* in wild populations of red deer. *HFM1* codes for a germ-cell specific helicase involved in spindle assembly. *MLH3* codes for a subunit of the main class I crossover resolvase MutLγ. *MSH4* and *MSH5* code for the two subunits forming the MutSγ heterodimer, which is a core component of the ZMM proteins that are responsible for the class I crossovers. *RNF212* and *RNF212B* code for E3 ligases homologs of *HEI10*. Finally, *REC8* codes for a meiosis specific cohesion subunit (Kadri et al., 2016a; Johnston et al., 2018a).

Natural variation and selection can be observed at different levels of regulation, it has been shown to affect meiotic factors directly but also chromatin structure and accessibility factors. This shows the vast extent of adaptability and evolution of meiotic recombination to intrinsic and extrinsic variables.

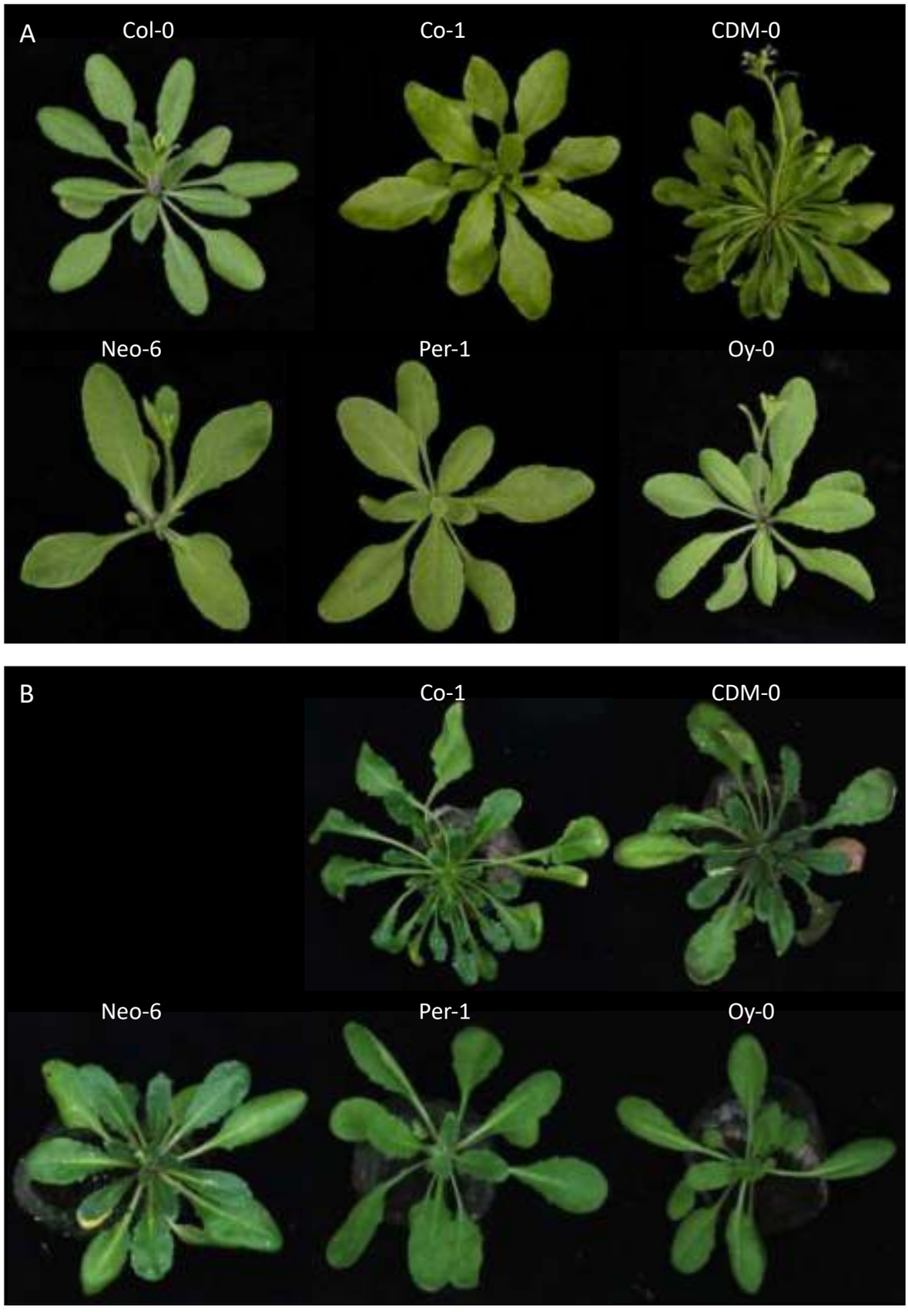


Figure 4. Representative rosette phenotypes displaying the used diversity. A. The six accessions used as parents for F1 hybrids and B. The F1s obtained from the cross of the Col-0 to the other five accessions. Image rights from panel A belong to ABRC. Image rights from panel B belong to Alexandre Pelé.

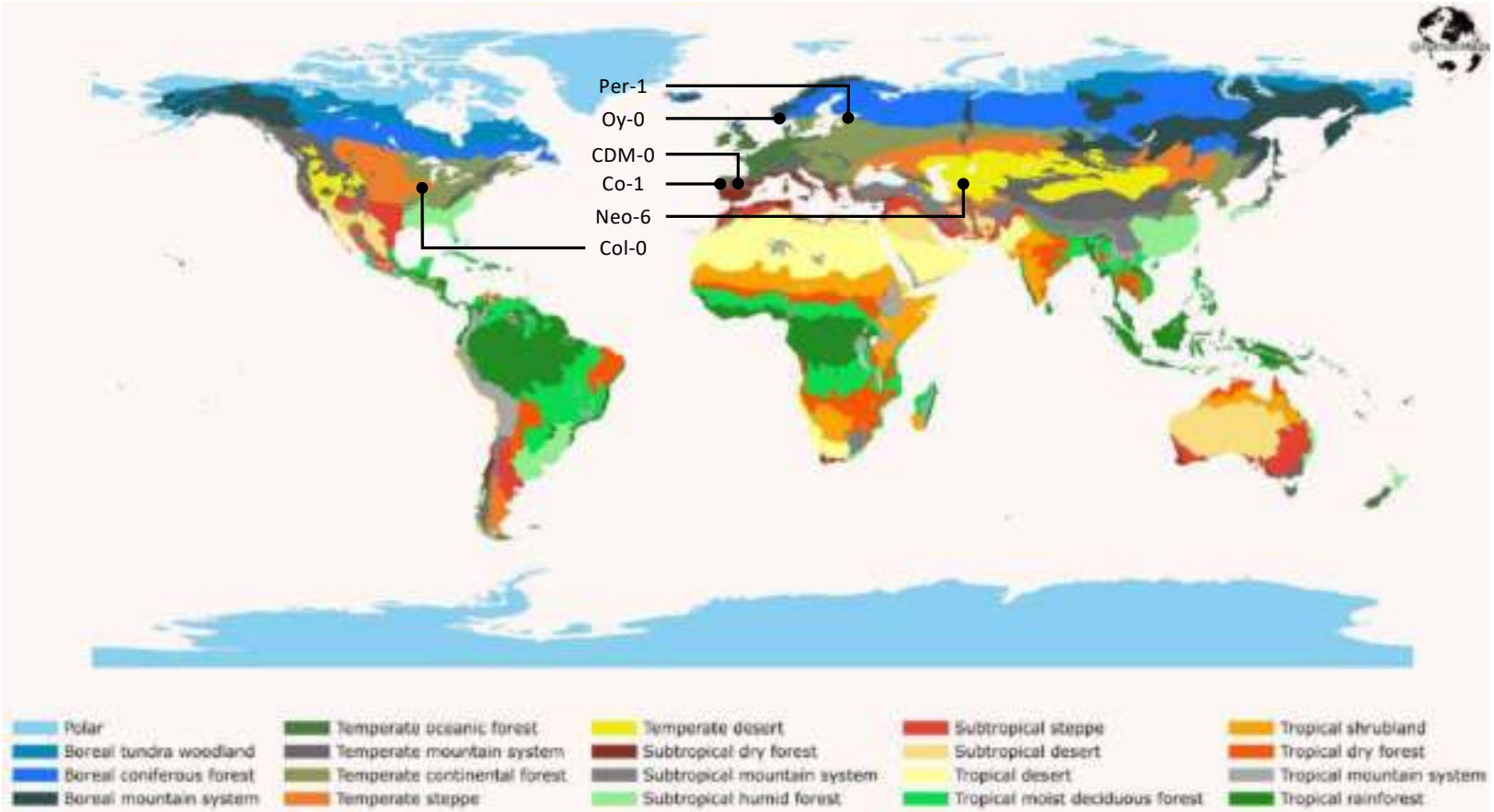


Figure 5. Positioning on a climate world map of the six accessions used for recombination QTL mapping. Per-1 originates from Russian temperate continental forest climate region. Oy-0 originates from Norwegian boreal coniferous forest climate region. CDM-0 originates from Spanish subtropical dry forest climate region. Co-1 originates from Portugese temperate forest climate region. Neo-6 originates from Tajikistan temperate desert climate region. The afore mentioned accessions were crossed to Col-0. Col-0 originates from temperate steppe climate Columbia, Missouri. Round arrow heads point to the respective geographical origin. Adapted, rights to the map belong to @PythonMaps (Dr. Adam Symington, https://python-maps.github.io).

1.4 Aims and research hypothesis

Five Arabidopsis accessions (Figure 4) were selected based on the climate of their geographical origin (Figure 5), flowering time and admixture group (Table 1). The aim of this investigation is to study the crossover landscape of the selected ecotypes and identify candidate QTLs involved in meiotic crossover recombination. The underlying hypothesis is that the genetic divergence introduced through adaptation to the original environment of each accession affected meiotic recombination (Feldman et al., 1996; Otto and Lenormand, 2002; Henderson and Bomblies, 2021). The latter is known to be affected by environmental conditions such as temperature. Although *Arabidopsis thaliana* only grows in temperate climates, the selected regions are rather distant geographically and experience different climatological conditions. *e.g.,* differences in temperature minima and maxima, daily temperature fluctuations, daily hours of light, different season transition, *etc.*

1.4.1 Experimental setting

The selected test accessions come from five different climates: CDM-0 originates from the Spanish subtropical dry forest climate, Co-1 from the Portugese temperate forest climate, Neo-6 from Tajikistan temperate desert climate, Oy-0 from Norwegian boreal coniferous forest climate and Per-1 originates from Russian temperate continental forest climate (Figure 5). These accessions were crossed to Col-420 which is a Col-0 reference line variant that harbors fluorescent markers on chromosome 3. Col-0 originates from the temperate steppe climate of Columbia, Missouri, USA. These accessions can be classified into five admixture groups: CDM-0 belongs to the “Spain” group, Co-1 “Italy_Balkan_caucasus”, Neo-6 and Per-1 to the “Asia” group, Col-0 to “Germany”. Oy-0 belongs to the admixed group, meaning that it has a multiple origin ancestry. The yearly average temperature of these regions varies from 0 to 15°C and single day fluctuations can be very different. Both high and low temperatures can induce chromosome missegregation (Lloyd et al., 2018). The rather large variation can

trigger adaptative strategies to warrant a successful outcome for meiosis and chromosome segregation.

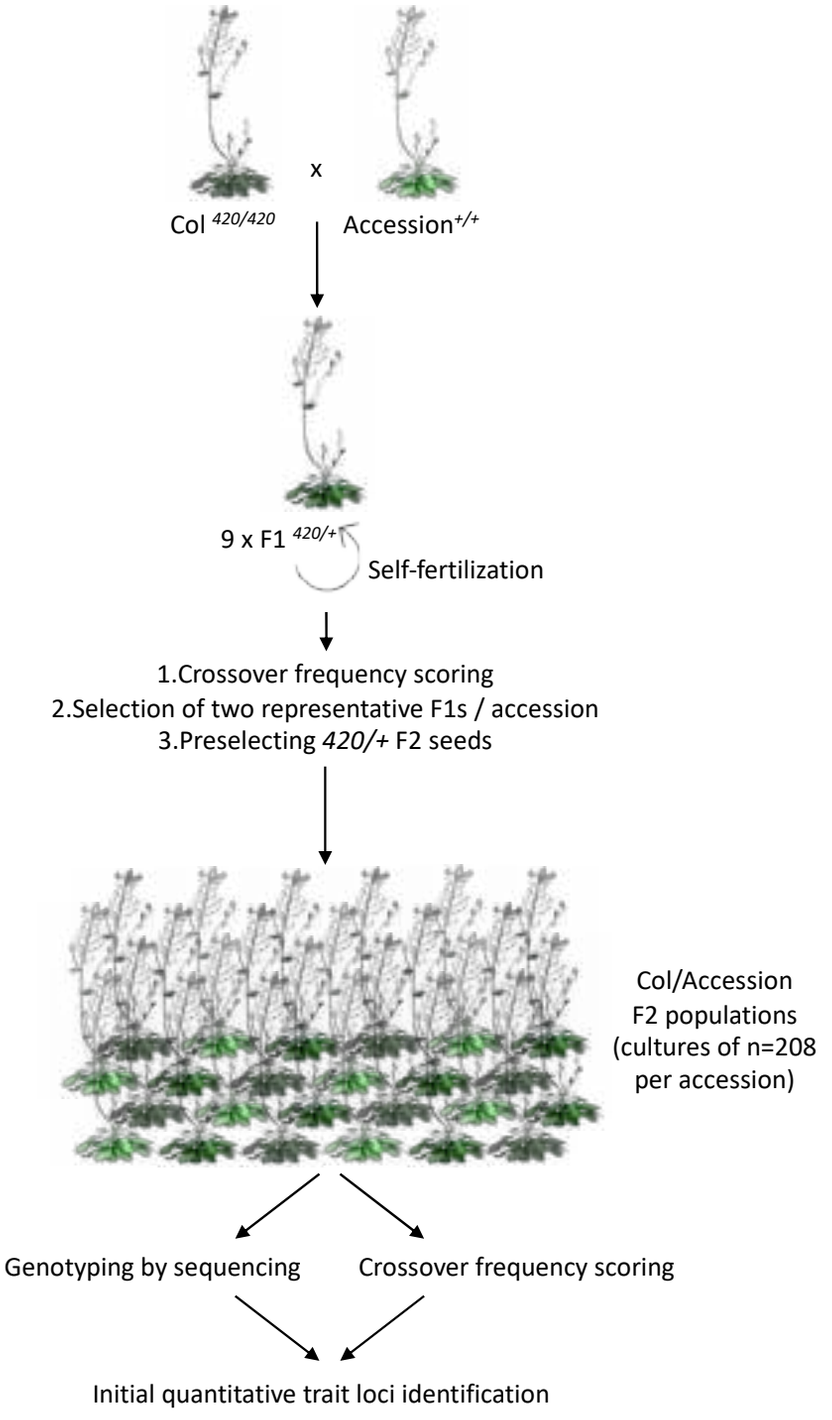


Figure 6. Strategy for the initial quantitative trait loci (QTL) identification. The five selected accessions, CDM-0, Co-1, Neo-6, Oy-0 and Per-1, were crossed to Col-420. Nine F1s for each cross were scored and two representative individuals were chosen for preselecting *420/+* F2 seeds. 208 F2s per accession were sown. These populations were subjected to whole genome sequencing (marker) and crossover frequency measurements (trait). The resulting data was used to compute the LOD (logarithm of the odds) of potential meiotic recombination QTLs. Rights to the Arabidopsis plant drawing belong to @_HETAKA, DOI: 10.7875/togopic.2021.057.

For practical and time management reasons, these lines were also selected for the fact that they do not require vernalization and their flowering times were under 50 culture days (Table 1).

The five accessions were crossed to Col-420. The F1s were grown to seed, and crossover rate was measured on chromosome 3 in the 420 subtelomeric interval. All lines showed appropriate fluorescence intensity and mendelian segregation for the fluorescent tags. Two representative F1 individuals, for each accession, were selected to be propagated to the F2 generation. The F2 seeds were selected for being hemizygous for eGFP and dsRed (GR/++) and grown. Leaf tissue was collected and used for genotyping by sequencing (GBS). The obtained sequencing data was used for mapping crossover distribution. In combination with the recombination frequency measured in 420, which is the chosen trait, the GBS data was also used for QTL mapping (Figure 6). Promising candidate QTLs would be further investigated, and their confidence interval narrowed down by backcrossing twice and mapping candidate causative genes in the BC2F2 populations. This work is still ongoing and so I will only discuss the obtained results up until the BC1 step.

1.4.2 Biological significance

This project is generating knowledge regarding the effects of the genetic polymorphism landscape on crossover frequency and distribution at the whole genome scale. The QTL mapping can give way to the identification of multiple genetic modifiers responsible for regulating meiotic recombination. Although core meiotic factors directly acting in recombination regulation have been largely identified and characterized, still very little is known about potential adaptative variants and many indirect and crucial actors. The high throughput nature of this investigation is, itself, promising in regards of the knowledge that can be harvested from such approaches.

Table 1. General information about the used accessions. Seed bank references, geographic origin, admixture group, climate and flowering time in days are provided. Acc = Accession.

Accession ID	Name	CS Number	Country	Lattitude	Longitude	Admixture Group	Climate	Flowering day	
								Accession	Col/Acc F1
9943	**Cdm-0**	CS76410	Spain	39.73	-5.74	Spain	Subtropical dry forest	45,0	51,0
7077	**Co-1**	CS76468	Portugal	40.12	-8.25	Italy_Balkan_caucasus	Temperate forest	37,0	51,1
6909	**Col-0**	CS76778	USA	38.3	-92.3	Germany	Temperate steppe	30,0	30,0
772	**Neo-6**	CS76560	Tajikistan	37.35	72.4667	Asia	Temperate desert	30,0	41,6
7288	**Oy-0**	CS77156	Norway	60.385543	6.193019	Admixed	Boreal coniferous forest	30,0	35,0
8354	**Per-1**	CS76571	Russia	58	56.3167	Asia	Temperate continental forest	41,0	37,6

2 Material and methods

1.1 Biological material and culture

2.1.1 Plant material

Arabidopsis thaliana seeds for the accessions Col-0 (N1092), Cdm-0 (N76410), Co-1 (N76468), L*er*-0 (N24238), Neo-6 (N76560), Per-1 (N76571) and Oy-0 (N77156) were purchased from the Nottingham Arabidopsis Stock Centre (NASC). The fluorescent tagged line (FTL) Col-420 was generously shared by Professor Avraham Levy (Melamed-Bessudo et al., 2005).

2.1.2 Growth conditions and seed collection

Seeds were sown on rehydrated Jiffy™ pellets. These are made from either 100% peat or a mixture of peat and Jiffy's own-manufactured RHP-certified coconut substrate, or 100% coconut substrate. After 3 days of stratification in the dark at 4°C the trays were transferred to controlled conditions culture chambers. The conditions used were 21°C, long day (16h day: 8h night), 70% humidity, and 130uM light intensity.

Seeds were collected when the plants were fully dry. They were then cleaned using a sieve and stored in glassine paper bags to keep them in a dry environment. This is important for maintaining a good fluorescence quality which is later used for crossovers frequency scoring.

2.1.3 Fertilizers and pesticide treatments

Plants were watered three times a week. Once per week, fertilizers were added to the water (5mM KNO_3, 2mM Ca $(NO_3)_2$, 2,5mM KH_2PO_4, 2mM $MgSO_4$, 50uM Fe-EDTA, 70uM H_3BO_3, 14uM $MnCl_2$, 0,5uM $CuSO_4$, 1uM $ZuSO_4$, 0,2uM Na_2MoO_4, 10mM NaCl and 0,01uM $CaCl_2$). In addition, once per month, or if needed, the plants were watered with the insecticide Substral Polysect 005 SL (Acetamiprid - 5

g/l, used at 1:100 dilution) and sprayed with the fungicide Syngenta Amistar OPTI 480 SC (32 % azoxystrobin, 0,5 % chlorothalonil, used at 1:200 dilution).

2.2 Molecular biology material

2.2.1 Quibit

Qubit 4 fluorometer was used to quantify genomic DNA and whole genome tagmented libraries. The samples were prepared according to ThermoFisher 1X dsDNA HS (high sensitivity) assay kit. Catalog numbers: Machine: Q33238, Reagent: Q33231, Tubes: Q32856.

2.2.2 KAPA2G Robust PCR Kit

The enzyme was purchased from MERCK. KAPA2G Robust was used for constructing and amplifying the whole genome sequencing libraries. Recommended proportions were maintained for a final reaction volume of 25uL. Catalog number: KK5024.

2.2.3 Electrophoresis

2.2.3.1 50X Tris Acetate EDTA (TAE)

The 50X TAE stock solution of 50 mM EDTA, 2M Tris base and 1M glacial acetic acid solution was periodically prepared by the laboratory manager. This solution was diluted 100 times for use as buffer for electrophoresis.

2.2.3.2 Agarose

Powder agarose was purchased from ABO Sp. z o.o. It was melted in 0,5X TAE at concentrations from 1 to 2% according to the size of nucleic acid to be resolved. Catalog number: BLE1.

2.2.3.3 Nucleic acid dye

SimpliSafe, the DNA stain, was purchased from EURX Sp. z o.o. It was used to visualize nucleic acid after resolution by electrophoresis and UV exposure. Catalog number: E4600-01

2.2.4 Clean-Up Concentrator

This kit was used to purify the pooled libraries PCR products. DNA was purified according to A&A Biotechnology recommendations. Catalog number: 021-250C.

2.2.5 Gel-Out Concentrator

This kit was used to purify the libraries after size selection. DNA was purified according to A&A Biotechnology recommendations. Catalog number: 023-250C.

2.3 Molecular biology methods

2.3.1 Leaf sampling and DNA extraction:

Two leaves of about 2 cm in length were collected from each 4 to 6 weeks old plant and placed separately in 1,2 mL volume 96 wells plates. Each well contained two 3 mm glass beads that serve to grin the -80°C frozen plant tissue using the QIAGEN TissueLyser II. The samples were grinded for 2 x 2 min at 30 shakes /sec to obtain a fine grain powder. The samples were incubated with shaking for 30 min in a 65°C water-bath after adding 350 µL of CTAB Buffer (140 mM Sorbitol, 220 mM Tris pH=8.0, 22 mM EDTA, 800 mM NaCl, 0,1% w/v Sarcosyl (N-Lauryl sarcosine sodium salt), 0,8% w/v Cetyltrimethylammonium bromide (CTAB). The DNA was isolated by the addition of an equal volume of chloroform to each tube. Samples were mixed vigorously then span down at maximum speed at 4°C for 20 min and the aqueous phases were transferred to fresh plates. The DNA was then precipitated by the addition of an equal volume of isopropanol to each well and an incubation for 5 min at room temperature before a maximum speed at 4°C for 20 min spin. The pellets were washed with 70% ethanol, span down at maximum speed at 4°C for

10 min, dried under a laminar hood for 5 min then resuspended in 100 µl of TE with RNase A (10 mM TRIS pH 7.5, 1mM EDTA pH 8, 100 ug/ml RNAseA) and incubated at 37°C for 30 min. DNA was purified by adding 0.1v of 3M NaAc and 2.5v 100% EtOH and incubating the samples at -20°C for at least 30 min followed by a 30 min, 4°C and maximum speed spin. The DNA pellets were washed once again with 70% ethanol, span down, dried and resuspended in 100 uL TE and stored at -20°C.

Qualitative DNA concentration and quality were checked by running 2ul of each sample in 1% agarose | 0,5X TAE gels. The samples were sorted into 3 groups: high, average, and low concentration. 8 representative samples from each group were tested using Qubit™ 1X dsDNA High Sensitivity (HS) assay Kit (ThermoFisher Scientific). The obtained values were used to make approximately 5ng/uL dilutions for all the samples.

2.3.2 Whole genome sequencing libraries construction

2.3.2.1 Loaded Tn5 preparation

Equimolar quantities of linker oligonucleotides Tn5ME-A (5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-3') or Tn5ME-B (5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-3') were mixed to Tn5Merev (5'-[phos]CTGTCTCTTATACACATCT-3') in a 50 mM NaCl, 40 mM Tris-HCl pH 8.0 annealing buffer (Picelli et al., 2014). The Tn5ME-A/Tn5MErev and Tn5ME-B/Tn5Merev linkers were annealed using the following program: 95°C | 5 min, -0.1°C / sec to reach 65°C, 65°C | 5 min, -0.1°C / sec to reach 4°C. The annealed linkers were mixed in a 1:1 ratio, diluted 5X with MilliQ sterile water then 1V of glycerol was added. The diluted oligonucleotides were afterwards loaded onto in-house produced Tn5 tranposase, prepared according to Hennig et al., 2018, in a 4V:1V (oligonucleotide: Tn5) ratio and incubated at 23°C with 340 rpm shaking for 30 to 40 min.

2.3.2.2 DNA tagmentation and amplification

5ng of DNA were tagmented in 10mM Tris-HCl pH=7.5, 10mM MgCl2, 0.025U Tn5, 10% DMF through a 1min 30 sec | 55°C incubation followed by a 10 min | 65°C inactivation after the addition of 1/6 V of 0,1% SDS. The tagmented DNA was then amplified using the Sigma-Aldrich KAPA2G Robust PCR kit in a 1X KAPA2G GC buffer with $MgCl_2$, 2mM dNTPs, 0.625U KAPA2G Robust enzyme and 0.2uM N7 and S5 index oligonucleotides (Supplemental table 11) and according to the following program: 3 min | 72°C, 1 min | 95°C, 14 x (10 sec | 95°C, 20 sec | 65°C, 3min | 72°C), 5min | 72°C and infinite hold at 4°C.

The tagmentation and amplification of the libraries are then checked by running 2 uL of each reaction on 1.5% agarose | 0.5X TAE gel. The desired fragment sizes range from 450 to 700 bp. The libraries were then pooled in equal amounts and concentrated using Clean-Up Concentrator (A&A Biotechnology Kit). At this stage pools consist of 96 libraries corresponding to 96 individuals. Our indexes allow for preparing 576 libraries, and so 6 pools are obtained.

2.3.2.3 Size selection and sequencing pools preparation

The concentrated pools are run on 1% agarose | 0.5X TAE gel for 2h and the sections corresponding to 450-700bp smears were cut and purified using the Gel-Out Concentrator (A&A Biotechnology Kit). The libraries were then eluted in 30 uL of TE. The concentration of DNA in the different pools was measured using Qubit™ 1X dsDNA High Sensitivity (HS) assay Kit (ThermoFisher Scientific) and pooled to equal molarity to obtain one final pool of at least C=[20ng/uL] and V= 30uL, as recommended by Macrogen, the contracted sequencing company. The libraries quality was checked using TapeStation and Agilent High Sensitivity D1000 kit at the Molecular Biology Techniques Laboratory, Adam Mickiewicz University in Poznan. The libraries were sequenced on HiSeq X-10 instrument (Illumina). The sequencing was outsourced in Macrogen Europe.

2.4 Bio-informatic methods

2.4.1 Seed based crossover rate scoring

2.4.1.1 Fluorescent tagged lines

The seed-based Fluorescent Tagged Line (FTL) consists of two fluorescent cassettes eGFP and dsRed that are present at known positions of Arabidopsis genome (Wu et al., 2015). They are expressed under the seed specific *napin* promoter and translate in coloring the seeds in green and/or red when excited with UV light and observed throw adequate filters. The FTLs are maintained at a homozygous state, *i.e.*, the fluorescent cassettes are present on both homolog chromosomes. To measure recombination in the region determined by the two fluorescent cassettes, the line of interest is crossed to the FTL and recombination is measured in the progeny. This system allows for the quantification of crossover events through the quantification of the frequency of segregation of the two fluorescent cassettes. For the QTL mapping Col-420 was used. In this line, meiotic recombination is measured in the region located on the north arm of Arabidopsis chromosome 3. The eGFP cassette is at 256516 bp (0.25 Mbp) and the dsRed cassette is at 5361637 bp (5.36 Mbp). 420 interval in 5.1Mbp big.

2.4.1.2 Cell profiler _ Automatic seed scoring

Seeds from each individual were verified for proper segregation of the fluorescent eGFP and dsRed tags (GR/++). Samples showing proper segregation were placed in a monolayer and pictured through 3 optic paths: Bright field, green field and red field (Figure 6).

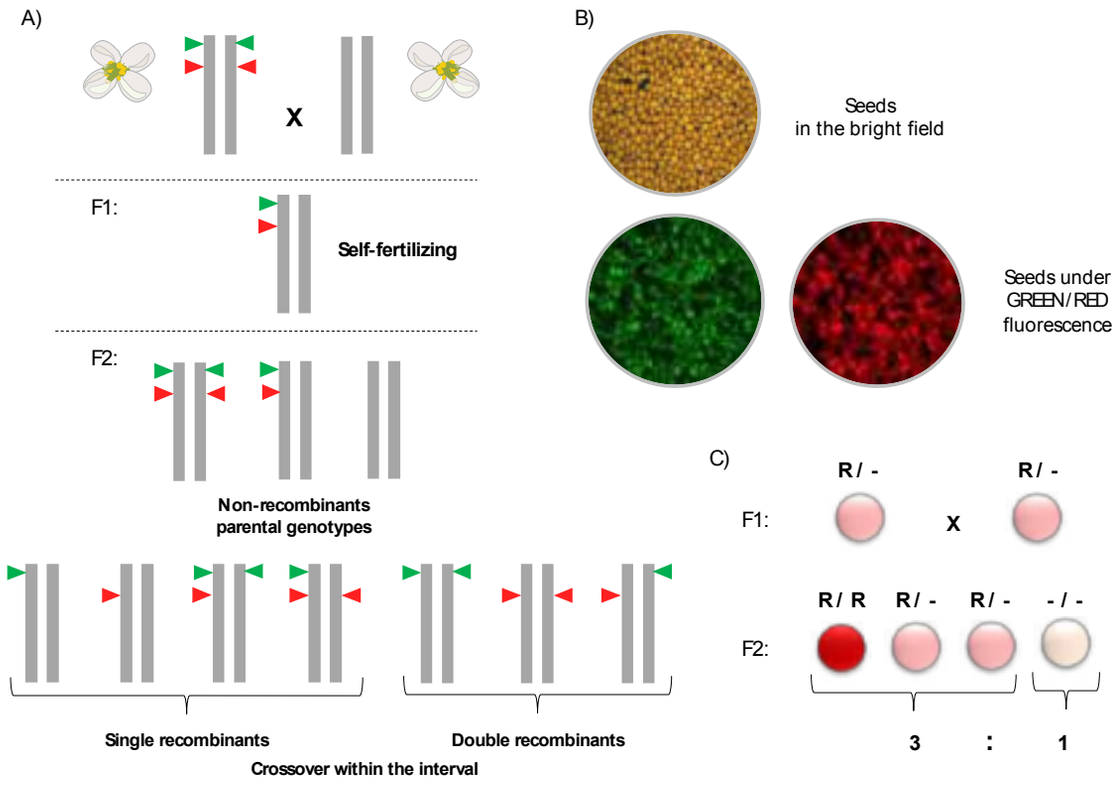


Figure 7. Seed-based system used to quantify crossover frequency in Arabidopsis thaliana. A*)* Crossing scheme with FTLs in order to obtain F2 seeds that show segregation of fluorescent reporters. Grey bars represent chromosomes. A. non-color line of interest is crossed with an FTL line containing fluorescent transgenes marking a specific chromosomal region. The F1 plant is self-fertilized to produce F2 offspring seeds, in which one can observe presence or absence of seed-expressed fluorescent proteins (see Note 14). As F2 seeds are diploid and generated through selfing, each seed is a product of both male and female meioses. Possible genotypes of F2 seeds are shown. B. Dosage-dependent expression of dsRed/eGFP genes under the control of seed-specific napin promoter. Images of Arabidopsis seeds in the bright field, under green and red fluorescence are shown. Strong green or red fluorescence indicate the presence of the fluorescent transgene copy on both chromosomes (FTL homozygous), whereas medium fluorescence values indicate hemizygote transgenes. Non-fluorescent seed is that which do not inherit FTL transgenes. C. Green/red fluorescent transgenes should show Mendelian inheritance, with the ratio of color to non-color seeds close to 3:1. Figure from Kbiri et al., 2022.

Because of the big size of the populations, *AutomaticSeedScoring* CellProfiler pipeline was used. This is the automated variant of the *SeedScoring* pipeline (Kbiri et al., 2022). *SeedScoring* pipelines use the afore mentioned set of three pictures as input (bright field, red field and green field). Single seeds objects are recognized, and an intensity of fluorescence value is attributed. Based on this value, the identified objects are later categorized as non-color or colored seeds. In the

automatic pipeline the threshold that categorizes the identified objects as color or non-color is estimated for each sample by the pipeline and not by the experimenter as for the manual *SeedScoring* pipeline. This allows for scoring up to 100 samples at once, with a significant decrease in the time consumption. The frequency of dissociation of the two colors is used to calculate the recombination frequency in the used interval.

Recombination Frequency (RF) in centiMorgan (cM) is calculated as follows:

Equation 1: $RF = 100 \times (1 - [1 - 2(N_G + N_R) / N_T]1/2)$

N_G = green-only fluorescent seeds, N_R = red-only fluorescent seeds and N_T = total number of seeds. Equation 1 is adapted from the genetic linkage equation:

Equation 2: $RF = (Total\ offspring / Recombinants) \times 100$

Equation 1 applies a correction that accounts for the fact that *SeedScoring* and *AutomaticSeedScoring* pipelines cannot discriminate between GR/++ and GR/GR seeds, which are all categorized as colored seeds.

2.4.2 Sequencing data computation

The sequencing results were received after being demultiplexed by Macrogen. The paired end reads are pooled giving two "*.fasta"* files. These sequences are aligned to the reference genome, Col-0, using the BOWtie2. The resulting Bam file sorted and indexed then single-nucleotide polymorphisms (SNPs) are called using SAMtools and mpileup. After which we obtain a text file listing the SNPs identified between the reference genome and the into sequences. Then the SNPs fist is used to identify the genotype of each individual at each of the identified genetic marker, Col/Col, Col/Acc and Acc/Acc. The mitochondrial and chloroplast genomes are then filtered out followed by a quality selection of each SNP is made. This selection is based on the sufficient coverage for each SNP (at least 5 reads) and the absence of bias towards the reference (Col-0) of the variant (Acc, accession). Centromeric and repetitive sequences are then covered but only on the SNP list file. A text file

is compiled and input into the Trained Individual GenomE Reconstruction (TIGER) software.

2.4.3 Genome wide crossover mapping

Using the Hidden Markov Model, TIGER generates haplotypes for each sequenced individual. Based on the number of SNPs it assigns a genotype "Ref", "Var" or "Het" at each genetic marker. The output text file can be used to compute crossovers maps and genetic distance maps, using mstmap and Kosambi function. The genetic distance map can then be used for QTL mapping.

2.4.4 Quantitative Trait Loci (QTL) mapping

The formatted ".csv" file with the genotypes associated to their respective phenotype is imported into R. QTL mapping was performed using version 1.5 of R/qtl statistical package (Broman et al., 2003). Single QTL analysis was performed using "*scanone*" function and the Haley-Knott regression with a 1cM step (Broman and Sen, 2009). The Logarithm of the odds (LOD) threshold was set by using 1000 permutations and a 0.05 significance level. An initial LOD confidence interval is estimated for each QTL using the "*lodint"* and the "*bayesint*" functions. These functions give an interval with the marker mapping to the highest LOD and supporting markers within +/- 1.5 LOD from the maximum (Broman and Sen, 2009). The effect of the identified QTLs is then visualized at the position of the highest LOD using the *"plotPXG"* and "*effectplot*" functions. The existence of additional QTLs and their linkage is verified using the "*scantwo*" function. Here we used the Haley-Knott regression with a 1cM step. The significant QTLs were combined, two by two, and their interaction tested using "*fitqtl"* function. Then, doublet interaction significance was tested using "*addint*" function. Finally, the contribution in phenotype variability and a refined position of each QTL were estimated using "*refineqtl*". New, narrower, confidence intervals for each identified QTL were estimated using "*lodint"* and the "*bayesint*" functions (Broman and Sen, 2009).

3 Results

Five *Arabidopsis thaliana* accessions, Cdm-0, Co-1, Neo-6, Oy-0 and Per-1 were crossed to Col-420 FTL which is in the Col-0 reference accession background. They were propagated to the second filial generation (F2). In an account for unavoidable plant losses over 200 *420/+* F2s were sown for each cross. Rosette leaves were collected for all five populations. DNA was extracted and used to construct genomic DNA libraries using Tn5 tagmentation. These libraries were amplified and sequenced using next generation Illumina sequencing. The resulting data was pre-demultiplexed by the contracted sequencing company. The reads were aligned to the Col-0 reference genome using BOWtie script. The SNPs were identified using SAMtools and mpileup. The resulting list of SNPs was used as an input file for the TIGER software to assign a genotype for each identified marker and for each individual. The obtained haplotypes were then used to compute crossovers maps, with 300 kb bins, and genetic maps, using mstmap and Kosambi function. In this approach, a crossover was defined as every haplotype switch between two marker/SNPs (Rowan et al., 2015; Blackwell et al., 2020).

3.1 Natural variability of crossover distribution in Arabidopsis

Here, I investigate how crossovers are distributed along the genomes for the five studied populations: Col-420 x Cdm-0, Col-420 x Co-1, Col-420 x Neo-6, Col-420 x Oy-0 and Col-420 x Per-1. Additionally, a Col-420 x L*er*-0 F2 population was grown in the same conditions to be used as a reference. Col x L*er* F2 populations were extensively used for studying crossover distribution and the effects of heterozygosity on crossover frequency in Arabidopsis (Ziolkowski et al., 2017; Rowan et al., 2019a; Blackwell et al., 2020; Nageswaran et al., 2021; Zhu et al., 2021; Kim et al., 2022). The north arm of chromosome 3 was filtered out for the used accessions because all individuals were preselected for being hemizygous for the 420 fluorescent makers.

The sequencing data was filtered for the quality of its coverage, at least 80,000 ~ 100,000 reads per individual. The presented data is based on, n = 167 individuals for Col-420 x L*er*-0, n = 177 for Col-420 x Cdm-0, n = 171 individuals for Col-420 x Co-1, n = 163 individuals for Col-420 x Neo-6, n= 163 individuals for Col-420 x Oy-0 and n = 141 individuals for Col-420 x Per-1.

3.1.1 Similarities and differences in the crossover frequency and chromosomal distribution in the studied *A. thaliana* populations

When compared to the Col-420 x L*er*-0 F2 population, all the five tested F2 populations follow an overall similar distribution along chromosomes (Figure 9 and Figure 10). They, however, show significantly lower crossover counts, except for the Col-420 x Oy-0 population (Figure 8). Generally, distal regions (sub-telomeres) and pericentromeric regions receive more recombination events than the interstitial regions. However, with a closer look the five tested accessions do not perfectly follow the Col-420 x L*er*-0 (hereafter L*er*) crossover distribution.

Col-420 x Cdm-0, (hereafter Cdm) shows overall more recombination event on distal regions than L*er* apart from the north arms of the fourth and fifth chromosomes which show less crossovers. In the interstitial part of the chromosomes, Cdm shows a slightly less active recombination. In the percentromeric regions, Cdm has less recombination events except for the north arm of the first chromosome and the south arm of the fifth chromosome (Figure 9A). When averaged along the chromosome arm, Cdm does show a similar but slightly lower recombination activity in subtelomeric regions and much less recombination activity in interstitial and pericentromeres regions (Figure 10A). At the genome wide scale Cdm-0 recombines less than Ler-0 (Figure 8).

Col-420 x Co-1 (hereafter Co) shows less recombination activity in distal regions apart from the south arm of chromosome 1. Interstitial regions are quite variable, they are sometimes more and sometimes less active than the L*er* reference. Pericentromeric regions are overall less active with a slightly more active regions

on chromosomes 1 and 5 north arms (Figure 9B). When averaged along the chromosome arm Co shows less recombination activity in subtelomeres and interstitial regions adjacent to the pericentromeres. The more distal interstitial part of the arm and the more proximal pericentromeric regions show higher recombination activity. On the whole genome scale Co shows a slightly decreased crossover rate in comparison with L*er* (Figure 8 and Figure 10B).

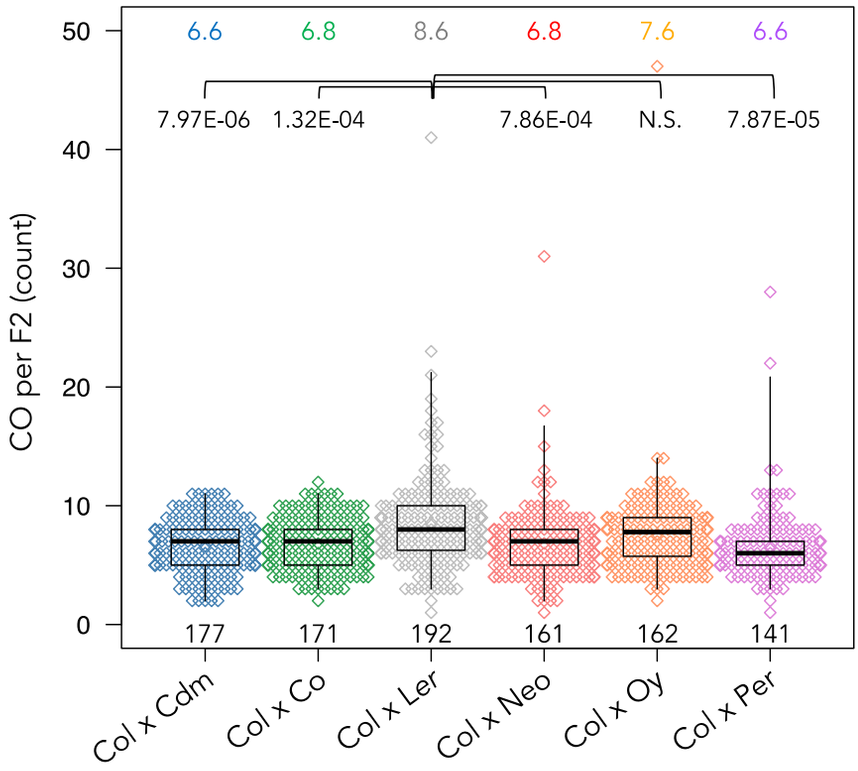


Figure 8. Crossover count in the six used F2 populations. Every diamond represents a single individual. The number of plotted individuals is indicated for each population. The thick black bar in the boxplots represents the average number of crossovers events. The exact values are indicated at the top of the graph.

Col-420 x Neo-6 (hereafter Neo) shows a similar or higher recombination rate in subtelomeric regions in comparison to L*er* except for the north arm of chromosome 4. Interstitial and pericentromeric regions are overall less active (Figure 9C). The recombination rate averaged along the chromosome arm shows the same trends. At the whole genome level, Neo shows a similar recombination rate but lower crossover count to that of L*er* (Figure 8 andFigure 10C).

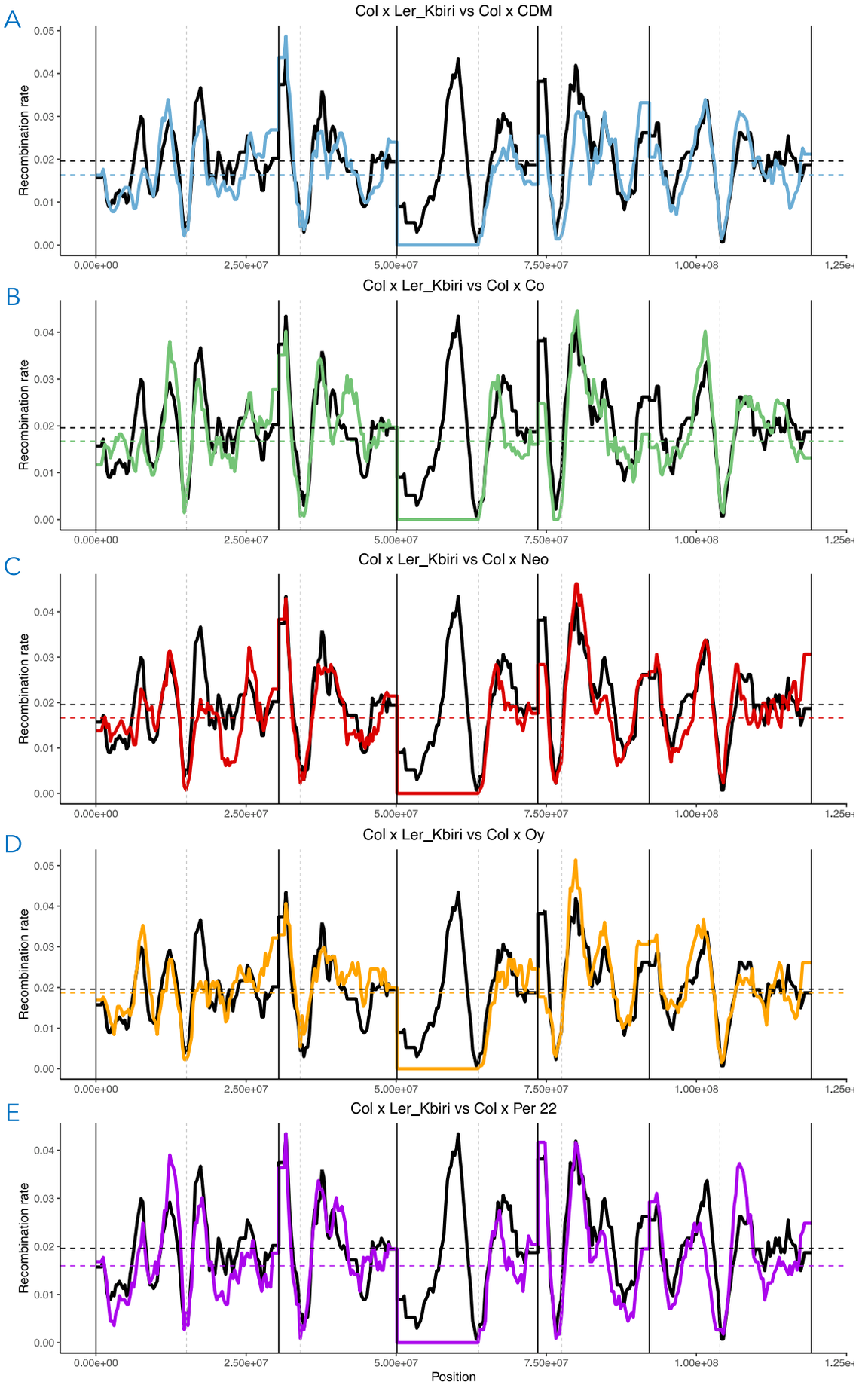

A
Col x Ler_Kbiri vs Col x CDM
B
Col x Ler_Kbiri vs Col x Co
C
Col x Ler_Kbiri vs Col x Neo
D
Col x Ler_Kbiri vs Col x Oy
E
Col x Ler_Kbiri vs Col x Per 22
Recombination rate
0.00
0.01
0.02
0.03
0.04
0.05
0.00e+00
2.50e+07
5.00e+07
7.50e+07
1.00e+08
1.25e+
Position

Figure 9. Recombination rate along the five chromosomes for the five F2 populations compared to a reference Col x L*er* F2 population. Crossovers were counted in 300kb bins. The whole genome mean crossover rate values are shown by horizontal dashed lines. Data are shown for Col x L*er* in black and the different accessions in the different colors. A. Col-420 x Cdm-0 in blue. B Col-420 x Co-1 in green. C. Col-420 x Neo-6 in red. D. Col-420 x Oy-0 in orange. And E. Col-420 x Per-1 in purple. The positions of telomeres and separations between chromosomes are represented by full vertical lines. Centromeres are labelled with vertical dashed lines.

Col-420 x Oy-0 (hereafter Oy) shows an overall similar or higher than L*er* along the chromosomes. However, significantly less recombination activity is observed in the pericentromeric region of chromosome 1 and chromosome 4 north arms (Figure 9D). The recombination rate along the chromosome arm shows higher recombination rate on distal and interstitial regions and a slightly decreased in pericentromeric regions. At the whole genome scale, Oy show a higher crossover recombination rate but similar crossover count in comparison to L*er* (Figure 8 and Figure 10D).

Col-420 x Per-1 (hereafter Per) does not show a very specific pattern of distribution of crossovers when the three different types of regions, distal, interstitial and pericentromeric, are observed separately for the five chromosomes (Figure 9E). When averaged along the chromosome arm, Per-1 shows higher recombination activity in subtelomeric and pericentromeric regions. The interstitial region experiences similar or slightly higher crossovers rate. At the whole genome scale, the average recombination rate is higher in Per than the L*er*, but the crossover count is significantly lower (Figure 8 and Figure 10E).

Overall, all five investigated populations show a crossover distribution that is consistent with all previously published maps. Subtelomeric and pericentromic regions are the most active regions of the genome when it comes to crossing-over. When compared to the L*er* population, differences in behavior can be observed. This can suggest the existence of modifiers that influence crossover rate and distribution. This is notably observed in the Cdm and Per populations, which show the most divergent behaviors. Cdm shows a lower recombination rate along the

chromosome arm and a lower crossover count. On the other hand, Per shows a higher recombination rate in pericentromeric and subtelomeric regions but a lower crossover count.

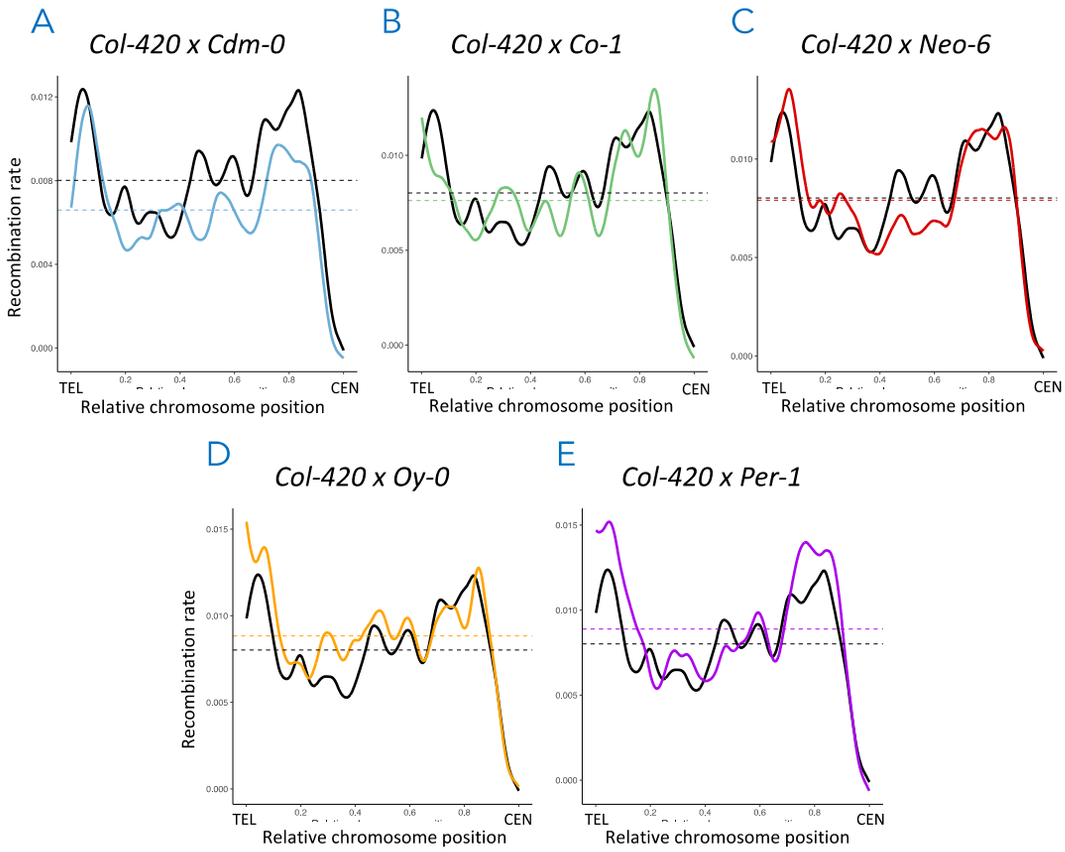


Figure 10. Recombination rate averaged along a chromosome arm for the five F2 populations compared to the Col x L*er* reference population. The telomeres (TEL) and centromeres (CEN) are represented (left to right). The whole genome mean crossover rate values are represented with horizontal dashed lines. Data are shown for Col x L*er* in black and the different accessions in the different colors. A. Col-420 x Cdm-0 in blue. B Col-420 x Co-1 in green. C. Col-420 x Neo-6 in red. D. Col-420 x Oy-0 in orange. And E. Col-420 x Per-1 in purple.

3.1.2 Single Nucleotide Polymorphisms and crossover rate correlate positively in different Arabidopsis accessions

The F2 populations crossover distribution reflects the recombination activity in the F1 mothers. As the F1 individuals were full hybrids, the meiotic recombination activity is probably influenced by the Single Nucleotide Polymorphism (SNP) density. Whole genome crossover events and SNPs counts were binned in 100 kb windows for each of the five chromosomes. The counted events were then summed

into percentiles and sorted in an ascending fashion according to the SNP count. The correlation of the SNP count to the crossover count was then plotted as shown in Figure 11, and Spearman's rank correlation coefficient *(r_s)* calculated.

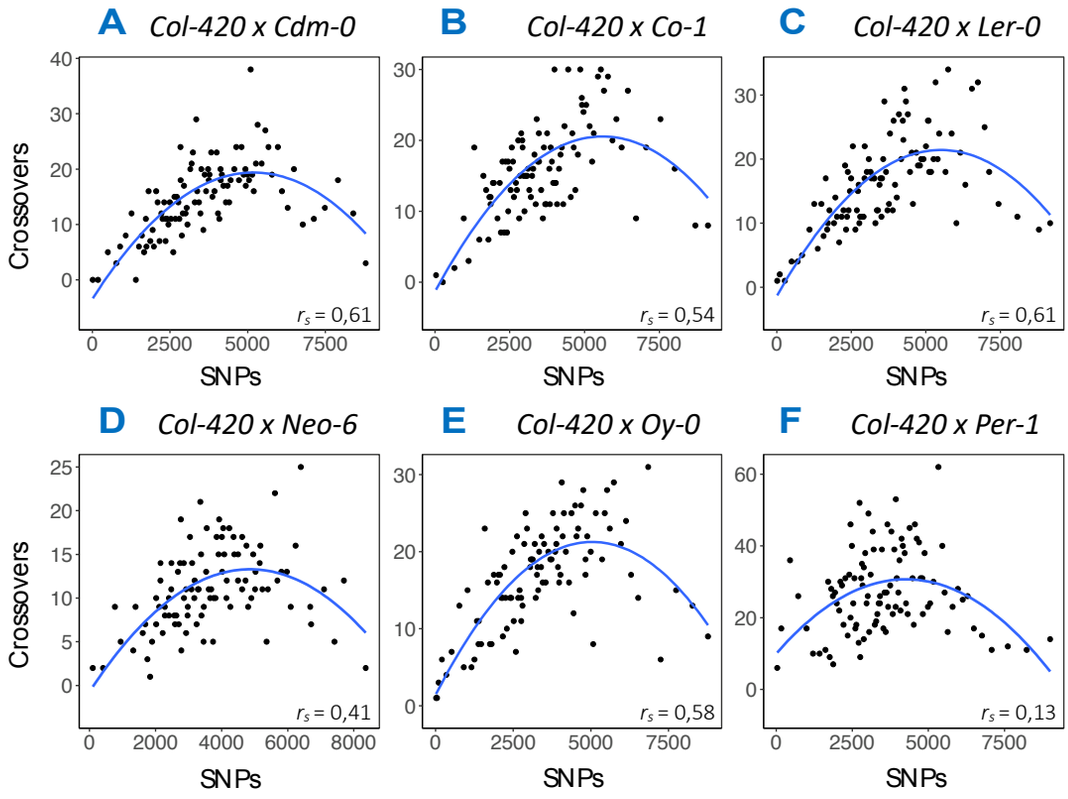


Figure 11. Single nucleotide polymorphisms to crossover number correlation in the five F2 populations and the Col x L*er* F2 reference population. Crossover count and Single Nucleotide Polymorphisms (SNPs) were binned in 100 kb windows. Spearman's rank correlation coefficient (r_s) was calculated for each population and is displayed on the bottom right of each graph. A. Col-420 x Cdm-0. B. Col-420 x Co-1. C. Col-420 x L*er*-0. D. Col-420 x Neo-6. E. Col-420 x Oy-0. And F. Col-420 x Per-1.

All populations show a positive correlation between SNP number and crossover number. However, on one hand, Cdm, Co, L*er* and Oy show relatively stronger correlation than the remaining populations, suggesting that the presence of SNP could favors crossover recombination events. r_s =0.61, 0.54, 0.61 and 0.58 respectively. On the other hand, Neo shows a relatively lower correlation with r_s =0.41 and there is virtually no correlation for Per, r_s =0.13. However, the correlation coefficient values must be treated with caution because the observed relationship is not linear. It is also interesting to note that Neo shows the lowest maximum number of crossovers (n ~ 25/ 100 kb) coupled to the highest maximum number

of SNPs (n ~ 6500/ 100 kb), whereas Per shows the highest maximum of crossover number (n ~ 60/ 100 kb) for a similar amount of SNP as the highly correlating populations (n ~ 5000/ 100 kb). Moreover, for all the six inverted parabolas, the crossover count to SNP count correlation is positive (upwards) till around 5000 SNP/100kb where it breaks downwards and becomes negative.

3.1.3 Discussion

In this section, I investigated the recombination profiles of five novel Arabidopsis hybrids created from crosses between Col and divergent accessions in comparison to the extensively studied Col x L*er* population (Ziolkowski et al., 2017; Serra et al., 2018b; Blackwell et al., 2020; Zhu et al., 2021). My results show that the global distribution of crossovers in all Arabidopsis hybrids follows the same pattern. Pericentromeres and subtelomeres receive more crossover events than the interstitial regions. Nevertheless, along chromosome arms, two out of five hybrids, Col x Cdm and Col x Co, showed a globally lower crossover recombination frequency than Col x L*er*. One hybrid, Col x Neo-6, showed a similar activity level, and Col x Oy and Col x Per, showed a higher rate than Col x L*er*. The five investigated lines show very similar detected SNP density, 3.03 SNP/ kb on average (3.04 for Cdm, 3.07 for Co, 3.08 for L*er*, 3.1 for Neo, 2.87 for Oy and 3.01 for Per). The consistency of the observed behavior is compelling because it shows that the genome-wide crossover distribution is not affected by our detected level of divergence in the different accessions. It is also consistent with the Lian et al., 2022 study that showed a similar result with the exception of major genetic indels and rearrangements. These types of polymorphisms were not considered in this study. This observation raised the question of how SNP density and crossover number correlate in these populations. To investigate this matter, I counted the number of crossovers events and SNPs in bins of 100 kb and sorted the crossover count following the ascending number of SNPs. This approach allows us to correlate the SNP density to the crossover count in standardized genomic portions

independently from their chromosome position. Just like the Col x L*er* reference population, all five tested populations showed a positive correlation between the crossover count and SNP density. This translates in most crossovers mapping to the regions with SNPs densities under 5000 SNP /100 kb (50 SNP/kb), the ascending part of the inverted parabolas. The parabola shape is consistent with some of the known effects of SNPs on meiotic recombination. Studies have shown that heterozygous regions tend to receive more crossovers at the expense of the homozygous regions (Ziolkowski et al., 2015; Blackwell et al., 2020). This is only true to some extent, as recombination is inhibited in highly polymorphic regions such as centromeres and repetitive elements (Higgins et al., 2012; Choi et al., 2016; Underwood and Choi, 2019). Crossovers are subjected to many intrinsic and extrinsic regulations and effectors other than SNPs. Some crossovers are subjected to interference, phenomenon by which the existence of a crossover inhibit the formation of a second one in close vicinity (Jones and Franklin, 2006; Yelina et al., 2013; Wang et al., 2015; Ziolkowski et al., 2015). Crossovers are also a byproduct of double strand breaks (DSBs) repair, which means, they are directed by the distribution of DSBs. DSB formation is believed to be a product of opportunity, where chromatin is enough accessible (Culligan and Britt, 2008; Keeney, 2008; Gray and Cohen, 2016; Smeenk and Mailand, 2016; Tian and Loidl, 2018; Xue et al., 2018). As such, heterochromatin characteristically enriched in DNA methylation, nucleosomes and generally more compacted, receives less DSBs and so less crossovers (Yelina et al., 2015b; Yelina et al., 2015c; Ziolkowski and Henderson, 2017; Fernandes et al., 2019; Rowan et al., 2019b). Although SNP to crossover count and crossover distribution can be very informative, SNPs cannot explain all the observed differences.

The initial hypothesis of this study is that the original environmental conditions of the different accessions may have given rise to genetic variability that could affect meiotic recombination distribution and frequency. The chosen approach is not ideal due to the need to generate hybrids in order to map crossover events. This

does not allow to map crossover in the chosen accession, but a combined effect of the studied accession and the reference accession. A better setting would be using a pure line of the accession, or an almost pure line with sufficient introduced SNPs as in Lian et al., 2022. Although the overall distribution of crossovers and SNP to crossovers counts behavior are similar in all accessions, some local differences in crossover frequencies and detected crossover count are observed. This can suggest the existence of meiotic recombination natural modifiers. This hypothesis is investigated in the following section.

3.1.4 Acknowledgments

F1 and F2 seeds from Cdm-0 x Col-420, Co-1 x Col-420, Neo-6 x Col-420, Oy-0 x Col-420 and Per-1 x Col-420 crosses were provided by Dr. Alexandre Pelé. MSc. Wojciech Dzięgielewski helped with NGS data analysis and created the R interface to generate the genome-wide and chromosome arm crossover distribution graphs.

Table 2. Number of individuals and genetic markers used in the QTL mapping for each population. The average segregation of the markers is also presented. Acc = Accession.

F2 population	No. individuals			No. of markers / Chromosome					Total markers	Average marker segregation (%)		
	Sequenced	Scored	Mapped	Chr1	Chr2	Chr3	Chr4	Chr5		Col/Col	Col/Acc	Acc/Acc
Cdm-0	192	141	141	218	55	65	99	189	626	24.2	51.3	24.6
Co-1	192	176	168	250	181	190	151	216	988	24.1	55.5	20.4
Neo-6	192	171	138	192	188	183	162	195	930	25.6	51.1	23.4
Oy-0	192	175	151	184	171	149	172	209	885	23.4	52.7	23.9
Per-1	192	181	123	275	318	260	293	282	1428	22.7	55.4	21.9

Table 3. Estimated locations and effect sizes of rQTLs identified in a Col-420 × Acc F2 population using single and multiple QTL mapping

Single QTL Mapping

Accession (Acc)	Chr	rQTL	Position (cM)	Proximal Marker (bp)	+/- 1.5 LOD units (cM)	+/- 1.5 LOD markers	420 cM			Mode of action	LOD	Variance (%)	Total model	
							Col/Col	Col/Acc	Acc/Acc				LOD	Variance (%)
CDM	1	Cdm-rQTL1	55.3	19190812	40.64...62.1	13648291...22180700	21.85	18.42	16.2	Semi-dominant	8.18	19.94	17.36	43.04
	3	Cdm-rQTL3	17	c3.loc17	14.65...22.37	8302212…9710429	23.96	17.98	21.3	Cis-effect	8.1	19.74		
Co	1	Co-rQTL1	61.89	20144289	61.29...69.03	19871338...23137052	20.20	17.13	14.91	Semi-dominant	12.80	22.53	23.79	49.79
	3	Co-rQTL3	14.3	8165118	10.73...24.52	7482009...9832264	19.57	16.96	21.36	Cis-effect	7.03	11.34		
	4	Co-rQTL4	6.33	736241	0...45.94	54125...11326312	19.40	17.15	16.71	Dominant Co	4.34	6.72		
Neo	3	Neo-rQTL3	25	c3.loc25	22.22...31.45	9575108...11434768	21.71	17.64	18.98	Cis-effect	6.25	18.96	6.25	18.96
Oy	-	None	-	-	-	-	-	-	-	-	-	-	-	-
Per	3	Per-rQTL3	15.5	8537404	7.17...38.24	6687487...11800636	18.3	17.3	20.60	Cis-effect	4.79	15.48	4.79	15.48

Multiple QTL mapping

Accession (Acc)	Chr	rQTL	Position (cM)	Proximal Marker (bp)	+/- 1.5 LOD units (cM)	+/- 1.5 LOD markers	420 cM			Mode of action	LOD	Variance (%)	Total Model	
							Col/Col	Col/Acc	Acc/Acc				LOD	Variance (%)
CDM	1	Cdm-rQTL1a	40.9	13824673	40.64...41	13482375...c1.loc41	21.80	17.99	16.98	Dominant Cdm	9.159	14.033	27.81	59.42
	1	Cdm-rQTL1b	41	"5"	40.9 … 42.73	"4" … "9"	21.85	18.35	16.94	Semi-dominant	8.848	13.485		
	1	Cdm-rQTL1c	55.6	19335839	51.64 … 63.94	17757861 … 22693711	22.44	18.4	16.18	Semi-dominant	6.401	9.361		
	3	Cdm-rQTL3	17	c3.loc17	14.65 … 20.8	8302212 …9466711	22.19	17.84	20.63	Cis-effect	11.475	18.294		
Co	1	Co-rQTL1	61.59	20009215	60.7…65.46	19629753...21579056	20.19	17.18	14.91	Semi-dominant	14.114	24.852	24.48	50.78
	3	Co-rQTL3	12	c3.loc12	11.62…24.51	7790821...9832264	19.59	17.03	21.5	Cis-effect	7.942	12.729		
	4	Co-rQTL4	1	c4.loc1	0…7.21	54125...823078	19.35	17.17	16.64	Dominant Co	4.586	6.991		
Neo	3	Neo-rQTL3	25	c3.loc25	24.70...31.45	9754374...11434768	21.71	17.64	18.98	Cis-effect	6.25	18.96	6.25	18.96
Oy	-	None	-	-	-	-	-	-	-	-	-	-	-	-
Per	3	Per-rQTL3	15.49	8537404	7…37.16	c3.loc7...11676355	18.3	17.03	20.6	Cis-effect	4.78	16.4	4.78	16.4

3.2 High throughput mapping for novel quantitative trait loci involved in meiotic crossover recombination

In this part of the project, the aim is to identify natural meiotic crossover recombination modifiers. The whole genome sequencing data was used to identify genetic markers that will be used for mapping. The SNPs were filtered to keep only the SNPs with at least five reads coverage. Then they were binned to the relative size of each chromosome to obtain an overall similar coverage for all five chromosomes. This process yields a list of SNPs with the highest causative relation to the genotype. About 1000 markers were selected for each population. The quantitative trait used for mapping is the recombination frequency in the subtelomeric 420 interval. This region is present on the north arm of chromosome 3. It stretches from 256,516 bp to 5,361,637 bp, making it 5.11 Mbp big. The choice was made to map for natural modifiers using the 420 interval because:

i. Subtelomeric regions are very active when it comes to meiotic crossover recombination (Ziolkowski et al., 2017; Serra et al., 2018c; Blackwell et al., 2020). Meiotic recombination has indeed been extensively studied and the biggest players have already been characterized. However, in this study, we are also open to investigating allelic variants of already known factors. Allelic variants can present different expression or activity profiles, which affects meiotic recombination in a way that is not observed in the commonly used reference ecotype Col-0.

ii. As the 420 interval was previously used in multiple QTL mapping and hybrid crosses (Ziolkowski et al., 2015; Ziolkowski et al., 2017; Lawrence et al., 2019; Blackwell et al., 2020; Zhu et al., 2021), it proved to be reliable when it comes to crossover frequency quantification. The reliability of the FTL is very crucial to quantification. From practical experience, fluorescent-tagged lines need to be tested for their segregation and fluorescence resistance. In fact, many FTLs experience silencing or sectorization of the fluorescence after a few generations or

because of hybrid genetic backgrounds. These problems make the crossover frequency measurements unreliable.

iii. These populations come from F1s generated for a bigger Genome-Wide Association Study (GWAS) that is currently ongoing. QTL mapping allows for a more profitable usage of an already existing unique biological material.

3.2.1 Single QTL analysis

For each population, a csv file where the haplotype for each F2 individual was associated with its measured recombination frequency in 420 was generated. This csv file is used as input for R/qtl package on R to detect Quantitative trait loci (QTLs) (Broman et al., 2003; Broman and Sen, 2009). Single QTLs were detected using the "scanone" function with 1 cM step Haley-Knott regression. Logarithm of odds (LOD) plots representing genetic markers on the X axis and the corresponding LOD score on the Y axis were obtained (Figure 13, Figure 15, Figure 17, and Figure 19). The LOD threshold of significance is determined using 1000 permutations and 0.05 accepted error. The lowest commonly accepted threshold is usually at least LOD 3. This value represents a 1:1000 chance for two loci to be genetically linked and bound to co-segregate. The interval of confidence for each QTL was calculated using "*lodint*" function. This function marks the limits of the confidence interval at the markers within +/- 1.5 LOD score from the marker with the highest estimated LOD value of each QTL (Broman et al., 2003; Broman and Sen, 2009; Ziolkowski et al., 2017; Lawrence et al., 2019; Zhu et al., 2021).

3.2.1.1 Col-420 x Cdm-0 F2 population

For the Col-420 x Cdm-0 F2 population, both genetic maps and recombination frequency scores were obtained for 141 individuals. The genetic map used for the mapping contains 626 markers in total. 218 markers on chromosome 1, 55 on chromosome 2, 65 on chromosome 3, 99 on chromosome 4, and 189 markers on chromosome 5. The average segregation of the markers follows a mendelian segregation with Col/Col representing 24.2%, Col/Cdm 51.3 % and Cdm/Cdm

24.6% (Table 2). Recombination frequency measurements average 18.31 cM.with the lower and upper whiskers of 10.76 cM and 29.94 cM, respectively (Figure 12). The recombination frequency of the Col-420 x Cdm-0 F2 population (n=141) has a 15.43 variance. The F2 population average recombination frequency is higher than that of the F1 population (n=9), 13.8 cM. Although biased by the significant difference in population size between the F1 and F2 populations, the relatively high variance in the F2 population suggests the presence of possibly segregating trans-modifiers (Figure 12).

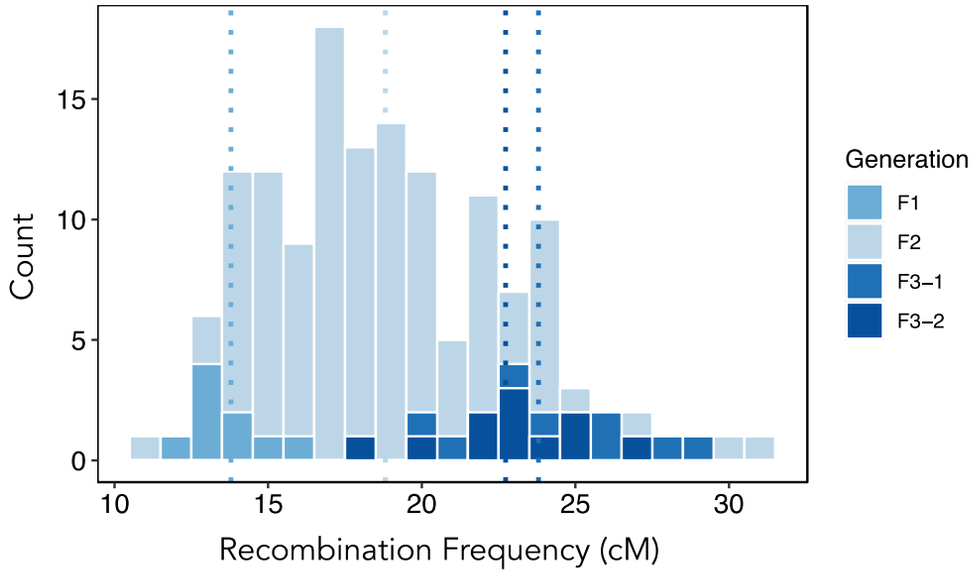


Figure 12. Distribution of recombination frequency in the 420 interval of three filial generations of Col-420 x Cdm-0. The F1, F2 and 2 selected F3s are represented. On the Y axis is represented the count of occurrence of a recombination frequency value, X axis, using a 1 cM bin. The mean of each population is represented with a vertical dotted line of the same color.

Single QTL detection for the Col-420 x Cdm-0 F2 population brought out two potential QTLs with LOD scores above threshold. The calculated threshold for this population after 1000 permutations and an accepted error of 0.05 is 3.84 (Red horizontal line, Figure 13). The first QTL peak is located on chromosome 1, hereafter labeled Cdm-rQTL1. The second QTL peak is on chromosome 3, hereafter labeled Cdm-rQTL3.

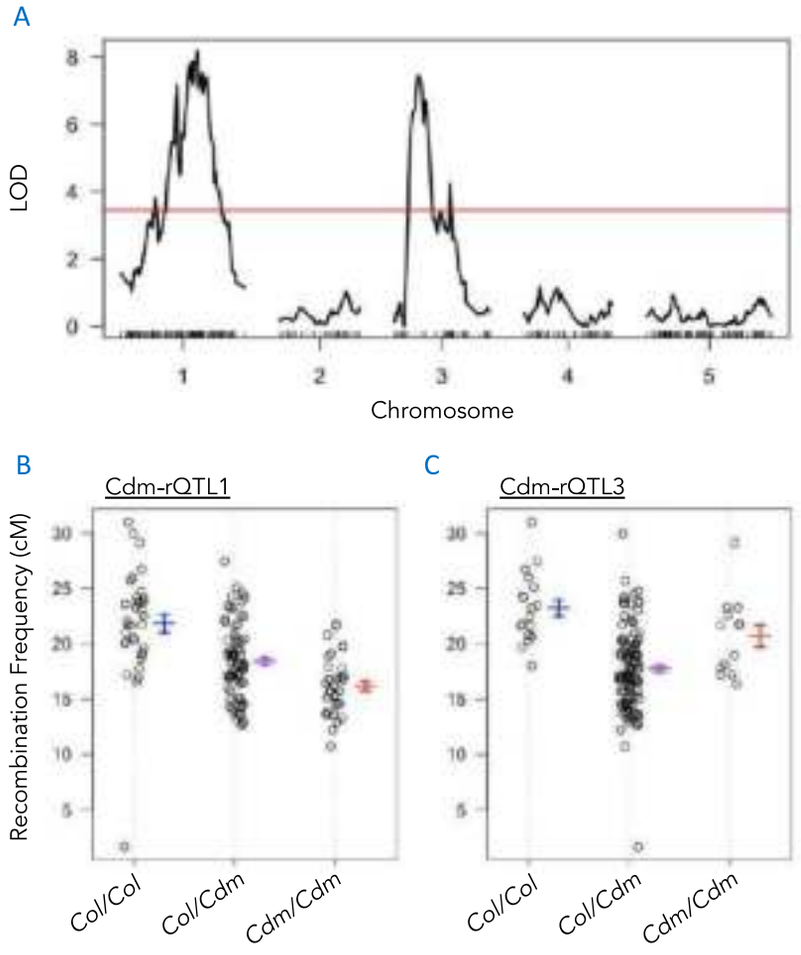


Figure 13. One-dimensional rQTL mapping in Col-420 × Cdm F2 population. A. LOD scores (Y axis) are represented for the selected genetic markers (X axis). The ticks on the X axis represent genetic markers distanced in cM. All five chromosomes are presented. The red horizontal line indicates the LOD significance threshold. B,C. Zygosity effect plot at the identified potential QTLs on chromosomes 1 and 3 respectively.

The highest peak of Cdm-rQTL1 shows a LOD score of 8.18. It corresponds to the genetic marker at 55.5 cM on the genetic map and 19,190,812 bp on the physical map. The confidence interval is 8,532,409 bp (8.5 Mbp) big. It spans from 40.64 cM to 62.1 cM on the genetic map and from 13,648,291 bp to 22,180,700 bp on the physical map (Table 3). It shows a semi-dominant mode of action with the Col/Col averaging a 21.85 cM recombination frequency, the Col/Cdm 18.42 cM and the Cdm/Cdm 16.2 cM (Figure 13B, Table 3).

The highest peak of Cdm-rQTL3 shows an LOD score of 8.10. It corresponds to the genetic marker at 17 cM on the genetic map and c3.loc17 on the physical map, a

calculated physical position. The confidence interval is 1,408,217 bp (1.41 Mbp) big. It spans from 14.65 cM to 22.37 cM on the genetic map and from 8,302,212 bp to 9,710,429 bp on the physical map (Table 3). It shows a Cis-effect with Col/Col averaging a 23.96 cM recombination frequency, the Col/Cdm 17.98 cM and the Cdm/Cdm 21.30 cM (Figure 13C and Table 3). In this QTL mapping we use the crossover frequency in the 420 interval as mapping trait. This introduces a segregation bias on chromosome 3 as it is preselected for being hemizygous for the green and red fluorescent markers, and so almost 100% heterozygous Col/Cdm. This segregation bias very often translates into a peak that mimics the presence of a quantitative locus. This does not mean that no recombination affecting loci are present on chromosome 3. However, with the chosen mapping setting, we cannot determine how much of the observed effect can be attributed to a quantitative locus and how much to the cis effect (Ziolkowski et al., 2017; Lawrence et al., 2019; Zhu et al., 2021).

3.2.1.2 Col-420 x Co-1 F2 population

For the Col-420 x Co-1 F2 population, both genetic maps and recombination frequency scores were obtained for 168 individuals. The genetic map used for the mapping contains 988 markers in totals. 250 markers on chromosome 1, 181 on chromosome 2, 190 on chromosome 3, 151 on chromosome 4 and 216 markers on chromosome 5. The average segregation of the markers follows a mendelian segregation with Col/Col representing 24.1%, Col/Co 55.5 % and Co/Co 20.4% (Table 2). Recombination frequency measurements average 19.35 cM and the lower and upper whiskers are 11.36 cM and 28.29 cM respectively (Figure 14). Recombination frequency of the Col-420 x Co-1 F2 population (n=176) has a 12.56 variance. The F1 population (n=9) average recombination frequency is 12.47 cM. This is closer to the lower whisker of the F2 population. Again, because of the significant difference in population size, no mathematical comparison can be made,

however, the F2 population still shows a relatively high variance again suggesting the presence of trans-modifiers (Figure 14).

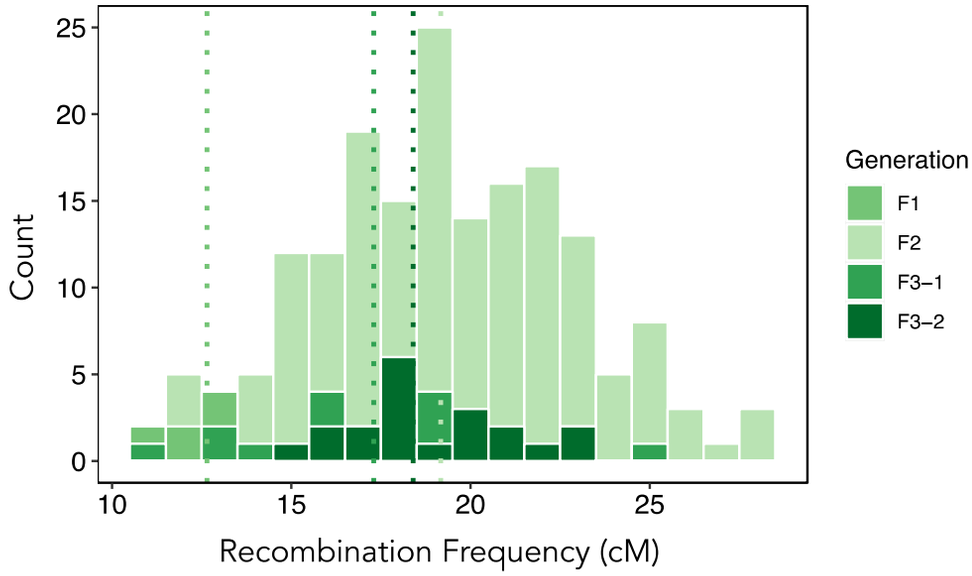


Figure 14. Distribution of recombination frequency in the 420 interval of three filial generations of Col-420 x Co-1. The F1, F2, and 2 selected F3s are represented. On the Y axis is represented the count of occurrence of a recombination frequency value, X axis, using a 1 cM bin. The mean of each population is represented with a vertical dotted line of the same color.

Single QTL detection for Col-420 x Co-1 F2 population brought out three potential QTLs with LOD scores above threshold (Figure 15). The calculated threshold for this population after 1000 permutations and an accepted error of 0.05 is 3.27. The first peak is on chromosome 1, hereafter labeled Co-rQTL1. The second peak is on chromosome 3, hereafter labeled Co-rQTL3. And the third peak is on chromosome 4, hereafter labeled Co-rQTL4 (Table 3).

Co-rQTL1 peaks at a LOD score of 12.8. This corresponds to the 61.89 cM position on the genetic map and the marker at 20,144,289 bp on the physical map. The confidence interval of Co-rQTL1 is 3,265,714 bp (3.26 Mbp) big. It spans from 19,871,338 bp to 23,137,052 bp on the physical map and from 61.29 cM to 69.03 cM on the genetic map. It shows a semi-dominant mode of action with Col/Col averaging 20.2 cM recombination frequency in the 420 interval, Col/Co 17.13 cM and Co/Co 14.91 cM (Figure 15B, Table 3).

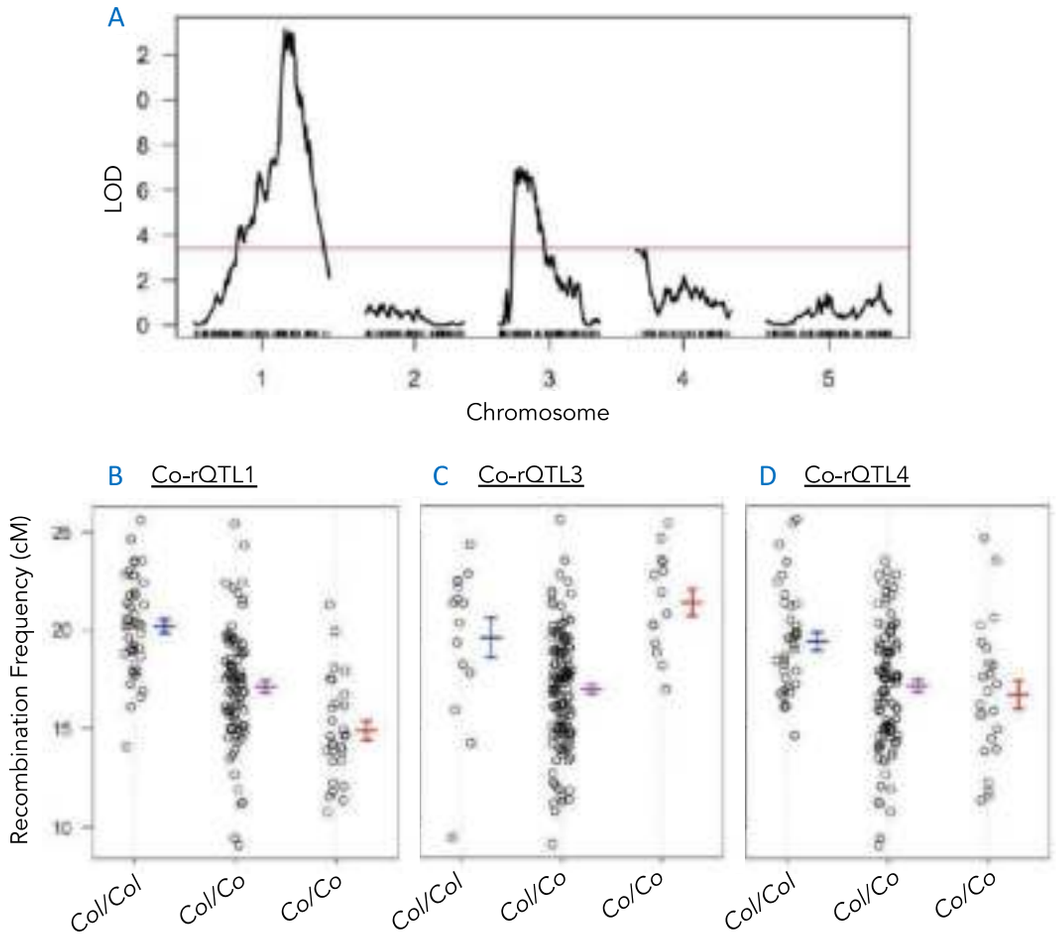


Figure 15. One-dimensional rQTL mapping in Col-420 × Co F2 population. A. LOD scores (Y axis) are represented for the selected genetic markers (X axis). The ticks on the X axis represent genetic markers distanced in cM. All five chromosomes are presented. The red horizontal line indicates the LOD significance threshold. B-D. Zygosity effect plot of the identified potential QTLs on chromosomes 1, 3 and 4 respectively.

Co-rQTL3 peaks at a LOD score of 7.03. This LOD correspond to the 14.3 cM position on the genetic map and the marker 8,165,118 bp on the physical map. The confidence interval for Co-rQTL3 is 2,350,255 bp (2.35 Mbp) big. Its span from 7,482,009 bp to 9,832,264 bp on the physical map and from 10.73 cM to 24.52 cM on the genetic map. It shows a Cis-effect with Col/Col averaging RF = 19.57 cM, Col/Co 16.96 cM and Co/Co 21.36 cM (Figure 15C, Table 3). As stated previously this rQTL is probably an artefact generated by the preselected seeds for hemizygous 420 fluorescent tags.

Co-rQTL4 peaks at a LOD score of 3.36. It corresponds to the 6.33 cM position on the genetic map and the marker at 736,241 bp on the physical map. Its confidence interval is 11,272,187 bp (11.2 Mbp) big. It spans from 54,125 bp to 11,326,312 bp on the physical map and from 0 cM to 45.94 cM on the genetic map. Co-rQTL4 shows a dominant mode of action for the Co-1 allele. Col/Col averages RF = 19.40 cM, Col/Co 17.16 cM and Co/Co 16.71 cM (Figure 15D and Table 3).

3.2.1.3 Col-420 x Neo-6 and Col-420 x Per-1 F2 populations

For the Col-420 x Neo-6 F2 population, both genetic maps and recombination frequency scores were obtained for 138 individuals. The genetic map used for the mapping contains 930 markers in totals. 192 markers on chromosome 1, 188 on chromosome 2, 183 on chromosome 3, 162 on chromosome 4 and 195 markers on chromosome 5. The average segregation of the markers follows a mendelian segregation with Col/Col representing 25.6%, Col/Neo 51.1% and Neo/Neo 23.4% (Table 2). Recombination frequency measurements of the F2 population (n=171) average 18.55 cM and the lower and upper whiskers are 12.16 and 26.06 cM respectively with the calculated variance of 10.66. The average recombination frequency of the F2 population is higher than that of the F1 (n=9), 15.86 cM (Figure 16A).

As for Col-420 x Per-1 F2 population, 123 individuals were used for QTL mapping. The genetic map has 1428 markers, which include 275 markers on chromosome 1, 318 on chromosome 2, 260 on chromosome 3, 293 on chromosome 4 and 282 markers on chromosome 5. The average segregation of the markers follows a mendelian segregation with Col/Col representing 22.7%, Col/Per 55.4 % and Per/Per 21.9 % (Table 2). Recombination frequency measurements of the F2 population (n=181) average 17.37 cM, the lower and upper whiskers are 12.54 and 23.39 cM, respectively, and the calculated variance is 5.96. The average recombination frequency of the F2 population is higher than that of the F1 (n=8), 13,45 cM (Figure 16B).

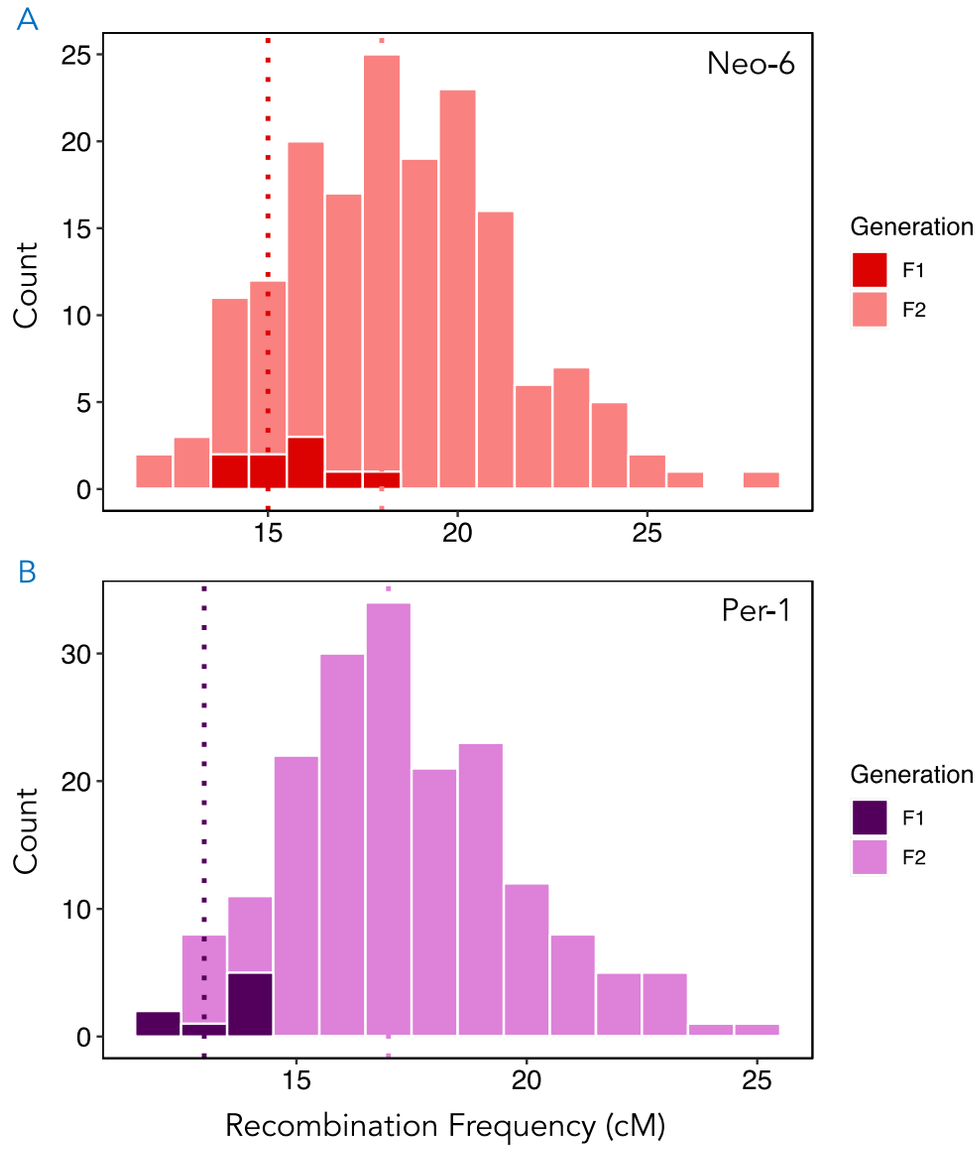


Figure 16. Distribution of recombination frequency in the 420 interval of two filial generations of Col-420 x Neo-6 (A) and Col-420 x Per-1 (B). the F1, and F2 populations are represented. On the Y axis is represented the count of occurrence of a recombination frequency value, X axis, using a 1 cM bin. The mean of each population is represented with a vertical dotted line of the same color.

Single QTL detection for Col-420 x Neo-6 and Col-420 x Per-1 F2 populations brought out one potential QTL with LOD score above the calculated thresholds. The calculated thresholds for these populations after 1000 permutations and an accepted error of 0.05 are 3.91 for Col-420 x Neo-6 and 3.44 for Col-420 x Per-1 (Figure 17).

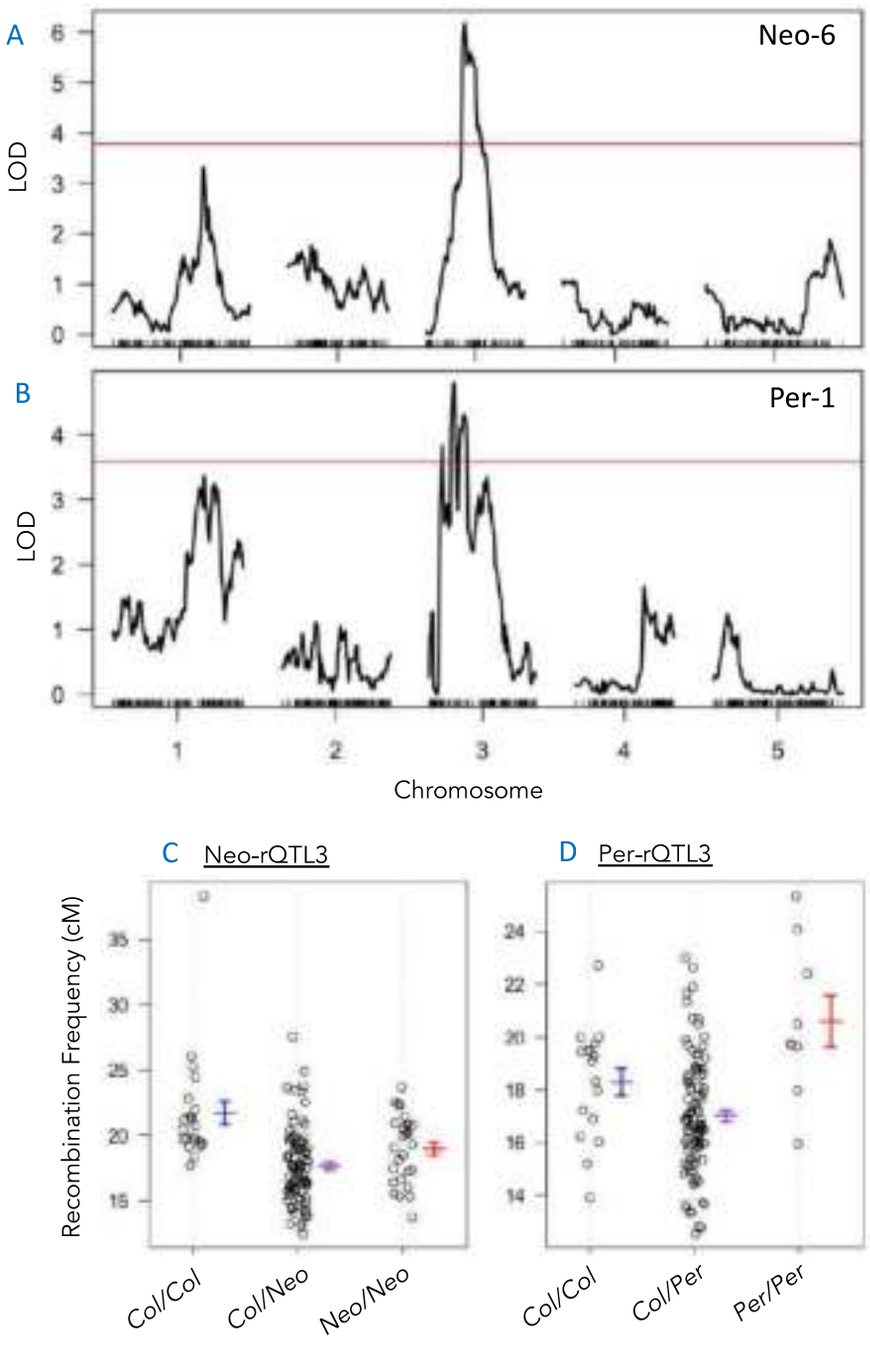


Figure 17. One-dimensional rQTL mapping in Col-420 × Neo-6 and Col-420 × Per-1 F2 populations. A & B. LOD scores (Y axis) are represented for the selected genetic markers (X axis). The ticks on the X axis represent genetic markers distanced in cM. All five chromosomes are presented. The red horizontal line indicates the LOD significance threshold. C & D. Zygosity effect plot at the identified potential QTLs on chromosomes 3 for Neo-6 and Per-1 respectively.

The Col-420 x Neo-6 peak on chromosome 3, hereafter labeled Neo-rQTL3, showed a LOD score of 6.25. Its corresponds to the 25 cM position on the genetic map and the c3.loc25 calculated position on the physical map. Neo-rQTL3

confidence interval is 1,859,660 bp (1.86 Mbp) big. It spans from 22.22 cM to 31.45 cM on the genetic map and from 9,575,108 bp to 11,434,768 bp on the physical map. Neo-rQTL3 likely corresponds to a cis-effect with RF = 21.71 cM for Col/Col, 17.64 cM for Col/Neo and 18.98 for Neo/Neo (Figure 17 A & C).

The Col-420 x Per-1 peak on chromosome 3, hereafter labeled Per-rQTL3, showed a LOD score of 4.79. It corresponds to the 15.49 cM position on the genetic map and the marker at 8,537,404 bp position on the physical map. Per-rQTL3 confidence interval is 5,113,149 bp (5.11 Mbp) big. It spans from 7.17 cM to 38.24 cM on the genetic map and from 6,687,487 bp to 11,800,636 bp on the physical map. Per-rQTL3 shows a cis-effect with RF = 18.3 cM for Col/Col, 17.03 cM for Col/Per and 20.6 for Per/Per (Figure 17 B & D).

Both these populations only show potential rQTLs only on chromosome 3. As discussed before, a bias was introduced by the preselection of individuals hemizygous for the fluorescent markers flanking the 420 interval. Theses QTLs cannot be investigated with confidence for harboring meiotic effectors. Therefore, Neo-6 and Per-1 populations were dropped from any further QTL mapping experimentations.

3.2.1.4 Col-420 x Oy-0 F2 population

Col-420 x Oy-0 F2 population yielded 151 individuals with both genetic maps and recombination frequency scores. The genetic map used for the mapping contains 885 markers in totals, including 184 markers on chromosome 1, 171 on chromosome 2, 149 on chromosome 3, 172 on chromosome 4 and 209 markers on chromosome 5. The average segregation of the markers follows a mendelian segregation with Col/Col representing 23.4%, Col/Oy 52.7% and Oy/Oy 23.9% (Table 2). Recombination frequency measurements of the F2 population (n=175) average 21.3 cM, the lower and upper whiskers are 13.9 and 28.29 cM respectively, and the calculated variance is 9.6. The average recombination frequency of the F2 population is higher than that of the F1 (n=8), 18.03 cM (Figure 18).

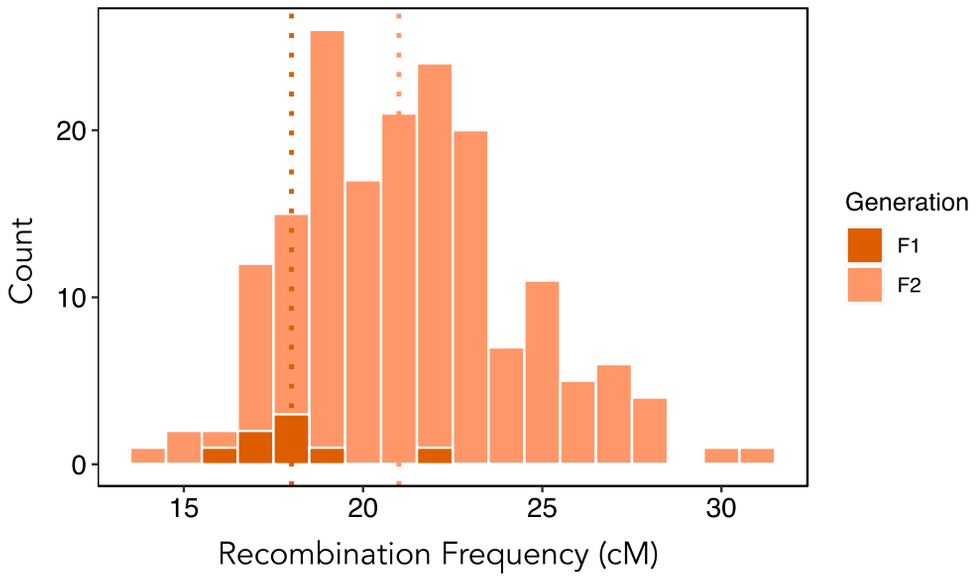


Figure 18. Distribution of recombination frequency in the 420 interval of two filial generations of Col-420 x Oy-0. the F1 and F2 populations are represented. On the Y axis is represented the count of occurrence of a recombination frequency value, X axis, using a 1 cM bin. The mean of each population is represented with a vertical dotted line of the same color.

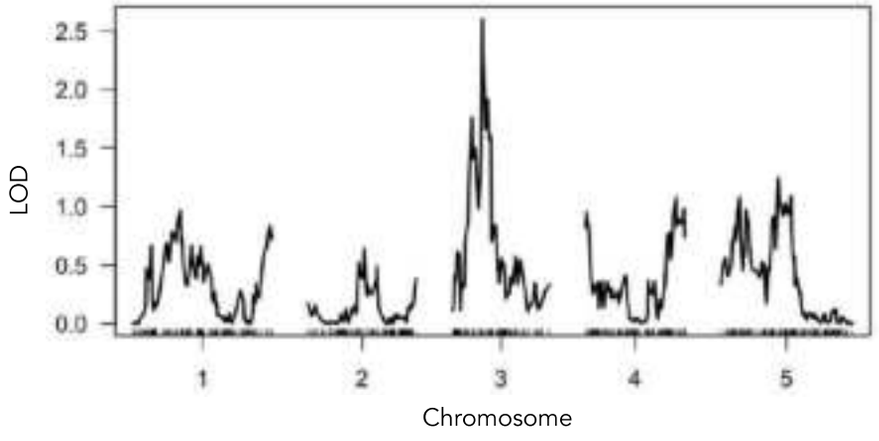


Figure 19. One-dimensional rQTL mapping in Col-420 × Oy-0 F2 population. LOD scores (Y axis) are represented for the selected genetic markers (X axis). The ticks on the X axis represent genetic markers distanced in cM. All five chromosomes are presented. The LOD threshold in not represented because it is much higher than the highest computed LOD value.

Single QTL mapping for Col-420 x Oy-0 F2 population did not show any potential QTLs. All detected peaks scored a LOD lower than the calculated threshold of 3.32, or the commonly accepted LOD score 3 (Figure 19). Therefore, this population was also dropped from any further QTL mapping experimentations.

3.2.2 Multiple QTL analysis

In this section, I attempt to predict weather or not the QTLs detected via "scanone" are single QTLs or fused multiple QTLs. This is useful for preselecting the lines that will be used for the further mapping. If the confidence interval possibly contains multiple QTLs, I can select lines in a manner that can separate the possible effect of each detected locus. The QTLs are detected using 1 cM step Haley-Knott regression. To separate the different QTLs, I used the functions "makeqtl" and "fitqtl". To test the linkage between the different single QTLs I used the formula: y~Q1*Q2*…*Qn, where n is the order of the QTL in the "makeqtl" step. Finally, the obtained QTLs were refined using "refineqtl" function. This function recalculates the probability for each QTLs using a Haley-Knott regression, and improves the LOD scores, when possible. The refined maps were also plotted with LOD scores on the Y axis and the chromosomes of interest on the X axis (Figure 20A and Figure 21A). The confidence interval for the newly identified QTLs were computed using "lodint" and "bayesint" functions. On Table 3 are reported the values obtained from the bayesint function. The effect of each QTL was plotted using the marker corresponding to the marker/physical position with the highest LOD score.

3.2.2.1 Col-420 x Cdm-0 F2 population

Using "scanone" function, two rQTLs were identified for the Col-420 x Cdm F2 population: The first, Cdm-rQTL1, is present on chromosome 1 with LOD score of 8.18 LOD score and a confidence interval of 8.5 Mbp (Table 3), and the second, Cdm-rQTL3, is present on chromosome 3 with an 8.1 LOD score. Cdm-rQTL3 cannot be used for further mapping because of the bias introduced by the pre-selection of the 420 region. However, it was maintained in the refined mapping because the linkage analysis showed that the QTL(s) on chromosome 1 and chromosome 3 are interdependent. This interdependence is expected as the mapping relies on the genetic link between the recombination frequency in the 420 region and the population haplotype segregation.

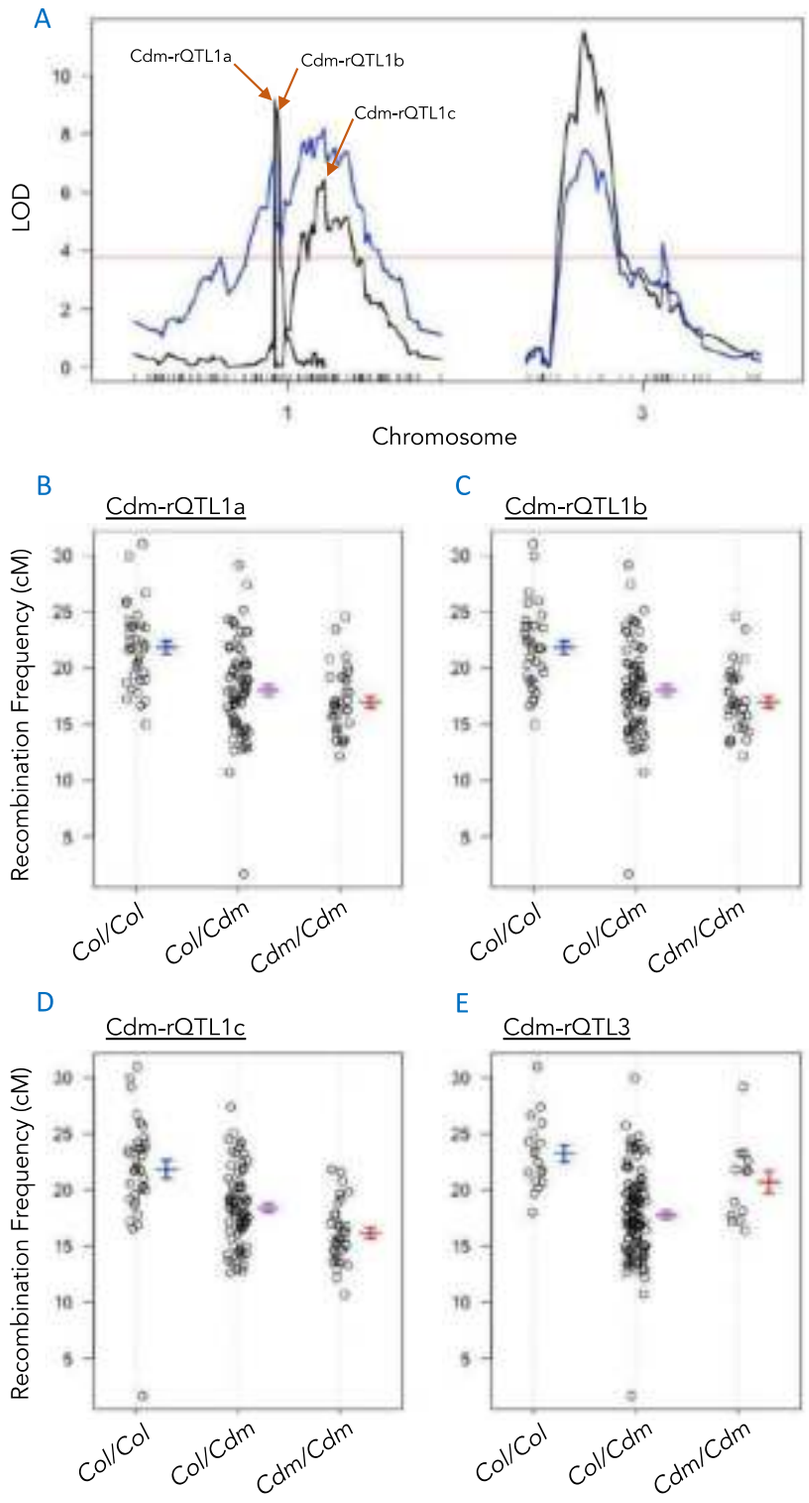


Figure 20. Refined multiple rQTL mapping in Col-420 × Cdm-0 F2 population. A. LOD scores (Y axis) are represented for the selected genetic markers (X axis). The ticks on the X axis represent genetic markers distanced in cM. Chromosomes 1 and 3 are presented. The red horizontal line indicates the LOD significance threshold, the blue graphs are the identified single QTLs. The black graphs are the same QTLs after the multiple QTL and refinement calculations. B-D. Zygosity effect plot of the identified potential QTLs on chromosomes 1 and 3.

The multiple QTL analysis on Cdm-rQTL1 showed that it could be causal of three rQTLs that will hereafter be labelled Cdm-rQTL1a to 1c. Cdm-rQTL1a shows a 9.16 LOD scoreand localizes at 40.9 cM on the genetic map and 13,824,673 bp on the physical map. The confidence interval of Cdm-rQTL1a spans from 40.64 cM to 41 cM on the genetic map and 13,482,375 bp to c1.loc41 of the physical map. Because of the lack of genetic markers in this region, the location of the limits of the confidence interval were estimated computationally. This rQTL displays a dominant Cdm allele mode of action. The zygosity effect for this QTL was plotted for the marker corresponding to 13,824,673 bp of the physical map. The recombination frequencies of the Col/Col, Col/Cdm, and Cdm/Cdm are 21.8 cM, 17.99 cM and 16.98 cM respectively (Table 3 & Figure 20B). The recombination frequency of the heterozygous Col/Cdm and homozygous Cdm/Cdm are statistically similar with a *p-value*= 0.15. On the other hand, the recombination frequency of the Col/Col was significantly different from Col/Cdm and Cdm/Cdm with *p-values* of 3.83E-06 and 2.82E-08, respectively. This suggests that the Col allele for Cdm-rQTL1a is dominant.

Cdm-rQTL1b shows an 8.85 LOD score and localizes at 41 cM on the genetic map and position "5" on the physical map. The confidence interval of Cdm-rQTL1b spans from 40.9 cM to 42.73 cM on the genetic map. The zygosity effect for this QTL was plotted for the marker corresponding to 15574,085 bp of the physical map, as it is the marker corresponding to the 41 cM position. The recombination frequencies of the Col/Col, Col/Cdm, and Cdm/Cdm are 21.85 cM, 18.35 cM and 16.94 cM respectively (Table 3 & Figure 20C). The recombination frequency of the Col/Col was significantly different from Col/Cdm and Cdm/Cdm with *p-values* of 0.001 and 1.7E-05 respectively. Additionally, Col/Cdm and Cdm/Cdm show a statistical difference, albite weak, with a p-value of 0.046. Thus, Cdm-rQTL1b displays a semi-dominant mode of action.

Cdm-rQTL1c shows a 6.4 LOD score and localizes at 55.6 cM on the genetic map and 19,335,839 bp on the physical map. The confidence interval of Cdm-rQTL1c spans from 51.64 cM to 63.94 cM on the genetic map and 17,757,861 bp to 22,693,711 bp on the physical map. It is 4,935,850 bp (4.9 Mbp) big. The zygosity effect for this QTL was plotted for the marker corresponding to 19,335,839 bp of the physical map. The recombination frequencies of the Col/Col, Col/Cdm and Cdm/Cdm are 21.8 cM, 17.99 cM and 16.98 cM respectively (Table 3 and Figure 20D). The recombination frequency of the Col/Col was significantly different from Col/Cdm and Cdm/Cdm with *p-values* of 2.48E-04 and 1.19E-07 respectively. Additionally, Col/Cdm and Cdm/Cdm show a statistical difference with a p-value of 2.08E-04. This rQTL displays a semi-dominant mode of action.

3.2.2.2 Col-420 x Co-1 F2 population

Three rQTLs were identified for the Col-420 x Co F2 population. The first, Co-rQTL1, present of chromosome 1 with a 12.8 LOD score (Table 3), the second, Co-rQTL3, present on chromosome 3, with a 7.03 LOD score, and the third, Co-rQTL4, present on chromosome 4, with a 4.34 LOD score. Co-rQTL3, which again is likely an artifact of seed preselection, was maintained in the refined mapping because of the genetic linkage to the other QTLs. The multiple QTL analysis on Co-rQTLs did not unravel the presence of any composite QTLs. However, it allowed for the improvement of the confidence intervals (Figure 21A).

After fitting and refining, Co-rQTL1, shows a 14.114 LOD score with a 1.95 Mbp confidence interval. It spans from 19,629,753 bp to 21,579,056 bp on the physical map and from 60.7 cM to 65.46 cM on the genetic map (Table 3). The proximal marker, used for the effect plot, is at 20,009,215 bp on the physical map and 61.59 cM on the genetic map (Figure 21B). Co-rQTL4, shows a 4.586 LOD score with a 0.77 Mbp confidence interval. It spans from 54125 bp to 823078 bp on the physical map and from 0 cM to 7.21 cM on the genetic map (Table 3). The proximal marker,

used for the effect plot, is at 54,125 bp on the physical map and 1 cM on the genetic map (Figure 21D).

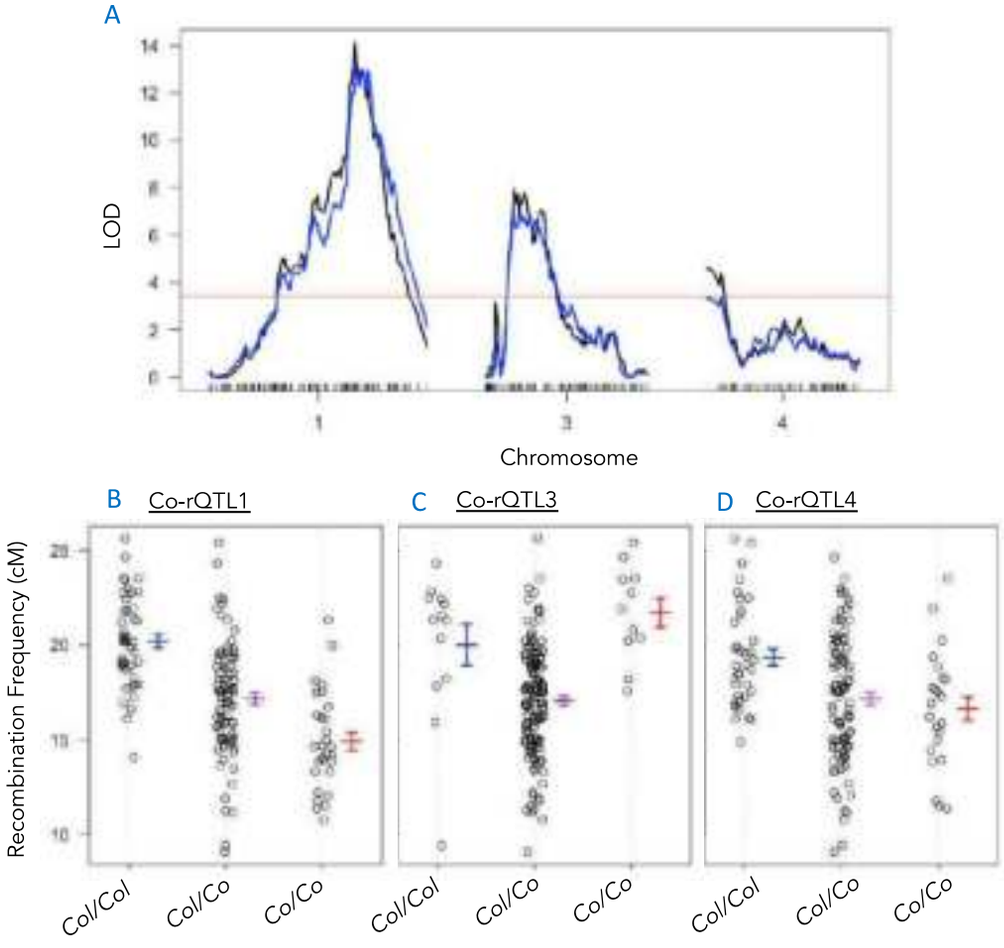


Figure 21. Refined multiple rQTL mapping in Col-420 × Co-1 F2 population. A. LOD scores (Y axis) are represented for the selected genetic markers (X axis). The ticks on the X axis represent genetic markers distanced in cM. Chromosomes 1, 3 and 4 are presented. The red horizontal line indicates the LOD significance threshold, the blue graphs are the identified single QTLs. The black graphs are the same QTLs after the multiple QTL and refinement calculations. B-D. Zygosity effect plot of the identified potential QTLs on chromosomes 1, 3 and 4 respectively.

3.2.2.3 The other populations

For Col-420 x Neo-6, Col-420 x Per-1 and Col-420 x Oy-0 F2 populations, the multiple QTL mapping did not bring any improvement to the results obtained with the single-QTL mapping.

3.2.3 The selected lines for backcross 1

To narrow down confidence intervals, individuals showing a heterozygous state for the intervals obtained during the initial selection were selected. In this section, I will shortly present the selected lines and the specific criteria that were chosen for each population.

A

bp	RF cM	25.97 1D	29.15 5G	16.92 11B
76768	0	H	A	H
10647209	24.92	H	A	H
10849776	26.23	H	H	H
11801086	32.53	H	H	H
12015362	33.84	A	H	H
13482375	40.64	A	H	H
13824673	40.90	A	H	H
13816332	40.90	A	H	H
13710767	40.90	A	H	H
15574085	41.43	A	H	H
16066362	41.69	A	H	H
16531299	43.26	A	H	H
16583673	44.04	A	A	H
18000948	51.90	A	A	H
18154191	52.16	A	A	A
19696851	56.34	A	A	A
20003939	57.39	A	A	A
20573769	58.69	A	A	A
20681479	58.95	H	A	A
22180700	62.10	H	A	A
26294359	74.66	H	A	A
26676602	76.77	H	H	A
28760429	83.04	H	H	A
30162582	89.25	A	H	H
30170843	89.51	A	H	H

B

Figure 22. Representation of the chromosome 1 haplotype for Cdm-0 x Col-420 selected lines. A. Haplotype representation of the three selected lines. Are presented, the physical map position in bp, the genetic map position in cM, the recombination frequency of the selected plants (RF) and their genotype for each position. The genotype for each marker is represented by three letters: "A" for Col, "H" for heterozygous and "B" for Cdm-0. The map was constructed by keeping the first and last marker with the same genotype for at least one line and one marker upstream and downstream from markers of interest. The crossed cells represent the marker(s) proximal to the centromere. The solid black rectangle represents the closest marker to *HEI10*. The dashed line rectangle represents the confidence interval for Cdm-rQTL1. B. Representation to scale of the confidence interval (beginning, end) and the position of *HEI10* gene on Arabidopsis chromosome 1.

3.2.3.1 Cdm-0 x Col-420 F3 selected lines

For Cdm-0 population, the focus was directed toward Cdm-rQTL1, chromosome 1. The confidence interval obtained through initial mapping spans from 13,648,291bp to 22,180,700 bp. The Cdm-rQTL1c in the refined mapping is positioned at the proximity of the genetic marker at the physical position 19,335,839 bp (Table 3). This position happens to be in very close proximity of a well characterized strong recombination modifier, *HEI10*, which physical coordinates are 19,963,267 to 19,966,952 bp. (Ziolkowski et al., 2017). To subtract its effect for the subsequent mapping, the lines were chosen for being fixed for the confidence interval of Cdm-rQTL1c and segregating for the before or after it (Table 3, Figure 22). With these criteria, the lines 1D and 5G from the sub-population CDM-0 1.1.3, and 11B from CDM-0 1.1.1 were chosen.

3.2.3.2 Co-1 x Col-420 F3 selected lines

As the Co-rQTL1 peaks at 20,009,215 bp on the physical map, which, again, is very close to the characterized recombination modifier *HEI10*, any further mapping of this QTL were put on hold. For this reason, I focused on Co-rQTL4, which is located on chromosome 4 and does not overlap with any known recombination QTL.

Co-rQTL4 spans from 54,125 bp to 11,326,312bp on the physical map on the initial mapping and from 54,125 bp to 823,078 bp on the physical map with the refined mapping. Lines with a heterozygous haplotype for parts of the initial confidence interval and the whole refined interval were selected for further mapping (Figure 23). The lines 11G, 7D and 3B all were descendants of the Co 1.1.2 subpopulation (Supplemental table 3).

3.2.4 Backcross 1

Seeds from the six selected F3 lines were selected for being fixed for the 420 reporter tags, GR/GR. The obtained plants were backcrossed to Col-0 and sequenced to check for their haplotype. The choice was made to sequence the F3s in parallel to genotyping them because the quality of the available sequences is

not good enough to design enough reliable SSLP primers and the dCAPS primers were inefficient.

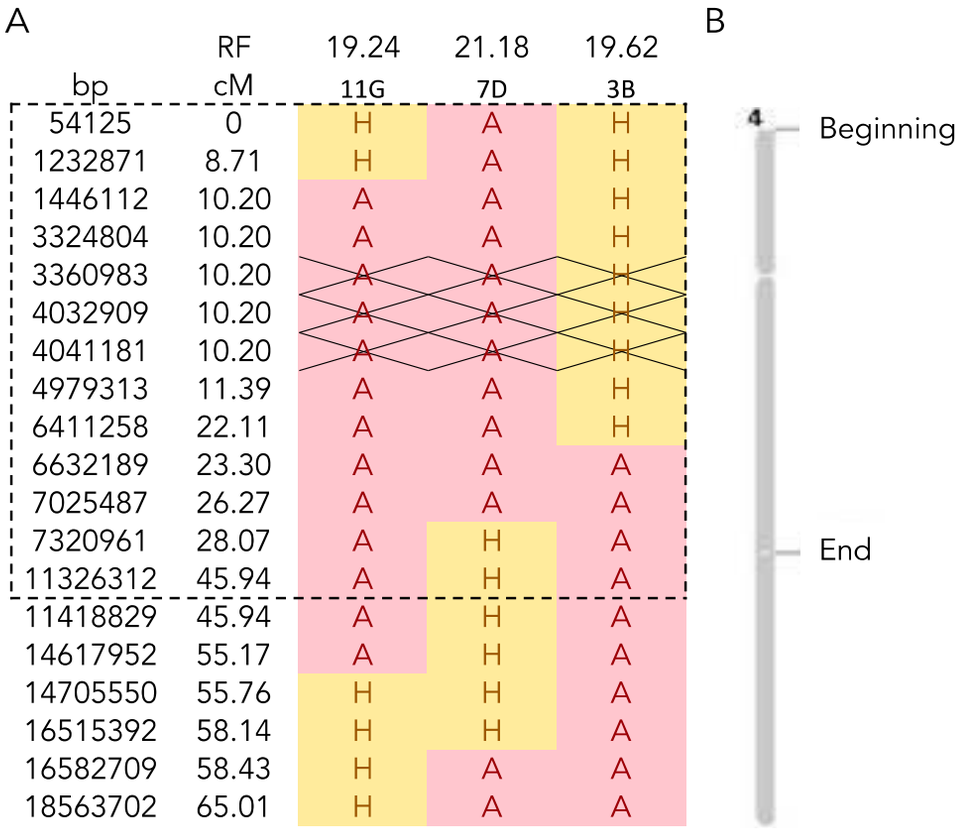


Figure 23. Representation of chromosome 4 haplotype for Co-1 x Col-420 selected lines. A. Haplotype representation of the three selected lines. Are presented, the physical map position in bp, the genetic map position in cM the recombination frequency of the selected plants (RF) and their genotype for each position. The genotype for each marker is represented by three letters: "A" for Col, "H" for heterozygous and "B" for Cdm-0. The map was constructed by keeping the first and last marker with the same genotype for at least one line and one marker upstream and downstream from markers of interest. The crossed cells represent the marker(s) proximal to the centromere. The dashed line rectangle represents the broader confidence interval for Cdm-rQTL1. B. Representation to scale of the confidence interval (beginning, end) on Arabidopsis chromosome 4.

3.2.4.1 Cdm-0 x Col-420 F3 haplotypes for the confidence interval

Twenty-four seeds for each of the selected lines were sown. Many of the plants had a very delayed flowering time. To cross-out this phenotype these plants were not used for crossing. Ultimately, six F3s from the 1D F2 individual, five from 5G and one from 11B were obtained, backcrossed to Col-0, and sequenced (Table 4). Lines 1D-5 and 6, 5G-1 and 5 and 11B-1 were selected for the second backcross (BC2). These lines present the longer stretches of heterozygosity or fixed Cdm-0. 1D-6

and 11B-1 show a heterozygous state for *HEI10*, so the effect of this modifier will still have to be crossed-out in the BC2.

Table 4. Haplotype of the backcrossed Cdm-0 x Col-420 F3 s within the Cdm-rQTL1 confidence interval. The physical map positions are presented in bp. The genotype for each marker is represented by three letters: "A" for Col, "H" for heterozygous and "B" for Cdm-0. The double-line represents where the centromere is. The solid black rectangle represents the closest marker to *HEI10*.

bp	1D-1	1D-2	1D-3	1D-4	1D-5	1D-6	5G-1	5G-2	5G-3	5G-4	5G-5	11B-1
13317896	A	A	A	A	A	A	B	H	A	A	H	H
16508910	A	A	A	A	A	A	B	H	A	A	H	H
16509262	A	A	A	A	A	A	B	H	A	A	H	A
16550296	A	A	A	A	A	A	B	H	A	A	H	A
16550315	A	A	A	A	A	A	B	A	A	A	H	A
16565957	A	A	A	A	A	A	B	A	A	A	H	A
16566118	A	A	A	A	A	A	H	A	A	A	H	A
16580334	A	A	A	A	A	A	H	A	A	A	H	A
16580539	A	A	A	A	A	A	H	A	A	A	A	A
16580836	A	A	A	A	A	A	H	A	A	A	A	A
16581312	A	A	A	A	A	A	A	A	A	A	A	A
17906806	A	A	A	A	A	A	A	A	A	A	A	A
17908685	A	A	A	A	A	A	A	A	A	A	A	H
19578103	A	A	A	A	A	A	A	A	A	A	A	H
19578554	A	A	A	A	A	H	A	A	A	A	A	H
20683567	A	A	A	A	A	H	A	A	A	A	A	H
20683620	A	A	A	A	A	B	A	A	A	A	A	H
20689856	A	A	A	A	A	B	A	A	A	A	A	H
20690064	A	A	A	A	H	B	A	A	A	A	A	H
20723405	A	A	A	A	H	B	A	A	A	A	A	H
20723534	H	A	A	A	H	B	A	A	A	A	A	H
25948659	H	A	A	A	H	B	A	A	A	A	A	H

3.2.4.2 Co-1 x Col-420 F3 haplotypes for the confidence interval

The same number of seeds were sown for the Co-1 x Col-420 F3 selected lines. Similarly delayed flowering plants were not used for backcrossing. Six F3s were maintained, backcrossed, and sequenced for 11G, six for 7D and 3 for 3B (Table 5). Finally, lines 11G-5 and 11G-6 and 3B-2 and 3B-3 were selected for BC2. Lines from 7D were put on hold for the moment and will be investigated if evidence of the presence of a QTL is seen on the south arm of chromosome 4.

Table 5.Haplotype of the backcrossed Co-1 x Col-420 F3 s within the Co-rQTL4 confidence interval. The physical map positions are presented in bp. The genotype for each marker is represented by three letters: "A" for Col, "H" for heterozygous and "B" for Cdm-0. The dashed line rectangle shows the position on the refined shorter confidence interval. The double-line represents where the centromere is.

bp	11G-1	11G-2	11G-3	11G-4	11G-5	11G-6	7D-1	7D-2	7D-3	7D-4	7D-5	7D-6	3B-1	3B-2	3B-3
1	B	A	A	A	H	B	A	A	A	A	A	A	A	H	H
219312	B	A	A	A	H	B	A	A	A	A	A	A	A	H	H
220197	B	A	A	A	H	B	A	A	A	A	A	A	A	H	B
303323	B	A	A	A	H	B	A	A	A	A	A	A	A	H	B
314810	B	A	A	A	H	B	A	A	A	A	A	A	A	B	B
461518	B	A	A	A	H	B	A	A	A	A	A	A	A	B	B
461978	B	A	H	A	H	B	A	A	A	A	A	A	A	B	B
994237	B	A	H	A	H	B	A	A	A	A	A	A	A	B	B
996847	H	A	H	A	H	B	A	A	A	A	A	A	A	B	B
1273068	H	A	H	A	H	B	A	A	A	A	A	A	A	B	B
1273139	H	A	A	A	H	H	A	A	A	A	A	A	A	B	B
1290881	H	A	A	A	H	H	A	A	A	A	A	A	A	B	B
1299006	A	A	A	A	H	H	A	A	A	A	A	A	A	B	B
1303449	A	A	A	A	H	H	A	A	A	A	A	A	A	B	B
1303689	A	A	A	A	H	A	A	A	A	A	A	A	A	B	B
1419663	A	A	A	A	H	A	A	A	A	A	A	A	A	B	B
1421326	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B
6298404	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B
6298561	A	A	A	A	A	A	A	A	A	A	A	A	A	B	H
6595204	A	A	A	A	A	A	A	A	A	A	A	A	A	B	H
6596009	A	A	A	A	A	A	A	A	A	A	A	A	A	B	A
7150798	A	A	A	A	A	A	A	A	A	A	A	A	A	B	A
7150887	A	A	A	A	A	A	A	A	A	A	A	A	A	H	A
7215169	A	A	A	A	A	A	A	A	A	A	A	A	A	H	A
7215610	A	A	A	A	A	A	A	A	A	A	H	A	A	H	A
7215979	A	A	A	A	A	A	A	A	A	A	H	A	A	H	A
7216096	A	A	A	A	A	A	A	H	A	A	H	A	A	H	A
7221666	A	A	A	A	A	A	H	H	A	A	H	A	A	H	A
7657372	A	A	A	A	A	A	H	H	A	A	H	A	A	H	A
7657859	A	A	A	A	A	A	H	H	A	A	H	A	A	A	A
8384810	A	A	A	A	A	A	H	H	A	A	H	A	A	A	A
8384846	A	A	A	A	A	A	H	H	A	H	H	A	A	A	A
8417782	A	A	A	A	A	A	H	H	A	H	H	A	A	A	A
8418348	A	A	A	A	A	A	H	H	H	H	H	A	A	A	A
8438728	A	A	A	A	A	A	H	H	H	H	H	A	A	A	A
8438745	A	A	A	A	A	A	H	H	H	H	B	A	A	A	A
9680941	A	A	A	A	A	A	H	H	H	H	B	A	A	A	A
9681066	A	A	A	A	A	A	H	H	H	H	H	A	A	A	A
11143111	A	A	A	A	A	A	H	H	H	H	H	A	A	A	A
11143173	A	A	A	A	A	A	H	H	H	B	H	A	A	A	A
11227176	A	A	A	A	A	A	H	H	H	B	H	A	A	A	A
11227272	A	A	A	A	A	A	B	H	H	B	H	A	A	A	A
11828337	A	A	A	A	A	A	B	H	H	B	H	A	A	A	A
11828416	A	A	A	A	A	A	H	H	H	B	H	A	A	A	A
11989045	A	A	A	A	A	A	H	H	H	B	H	A	A	A	A
11989152	A	A	A	A	A	A	H	H	H	H	H	A	A	A	A

3.2.5 Discussion

In this section, I used the five segregating populations to map for recombination QTLs. The selected trait for the mapping is recombination frequency in the subtelomeric interval 420. This interval is flanked from each side by two fluorescent tags, eGFP and dsRed. The frequency by which these two tags are separated allows us to measure recombination in the given 420 interval. The haplotype data obtained from the GBS allows to correlate different recombination frequencies with given genotypes. On R, using r/qtl package, I generated LOD maps for all five populations. The single QTL mapping yielded two QTLs for the Cdm, one on chromosome 1 and the second on chromosome 3, and three QTLs, on chromosomes 1, 3, and 4, for the Co populations. The other three populations did not yield any QTLs that can be considered.

The QTLs mapped to chromosome 3 were not considered for further investigation as they are most probably an artifact generated by the preselection for hemizygous fluorescent reporters to allow for recombination measurement. The obtained QTLs were further refined to check if they were a composite of multiple QTLs or not. The QTL on chromosome 1 of the Cdm population could be divided into three, Cdm-rQTL1a, 1b, and 1c. The QTLs on chromosomes 1 and 4 of the Co population seem to contain single QTLs. The size of the populations is relatively small but sufficient for an initial mapping, but it does not allow for very reliable refined QTL mapping. This was a compromise between the population sizes and the cost of the GBS. The chosen method already proved its effectiveness in Lawrence et al., 2019, where the initial mapping was performed similarly. Using GBS also allowed me to screen multiple populations simultaneously, screening of which would have been far more tedious and much more time consuming if a classic mapping by genotyping approach was adopted (Ziolkowski et al., 2017; Zhu et al., 2021).

After identifying the confidence intervals for the mapped QTLs, F2 lines that showed a heterozygous state for the whole or parts of the confidence intervals

were selected. For the three identified QTLs on chromosome 1 of the Cdm population, Cdm-rQTL1a and 1b were selected for further investigation, but not Cdm-rQTL1c. Cdm-rQTL1c maps in the very close vicinity to the already known and strong recombination modifier *HEI10* (Ziolkowski et al., 2017). Therefore, an effort was made to choose three lines that are homozygous Col for this region, in order to subtract its effect for the future mapping. A similar situation was seen with the Co population, the Co-rQTL1 maps at very close vicinity to *HEI10*, and so it was not prioritized for further mapping. On the other hand, three lines with a heterozygous genotype for the confidence interval of Co-rQTL4 were selected.

The recurrent mapping of *HEI10* as a recombination QTL in two of the tested populations and the published Col x L*er* and Col x Bur populations positions *HEI10* as a major natural crossover recombination modifier in Arabidopsis. it is however important to keep in mind that the chosen strategy relies on the FTLs which are only available in Col and L*er* backgrounds. This limitation does not allow to identify other potential natural modifiers that would only be detected from more diverse Arabidopsis accessions. This limitation can be overcome by using a GWAS approach for example.

F3 seeds from the selected lines were preselected for fluorescence and backcrossed to Col-0. These same F3 plants were sequenced to select the backcrosses that retained the heterozygosity for the confidence intervals allowing future fine mapping. Because of time limitations, causal genes for the different potential QTLs are yet to be identified. As of the writing of this manuscript, backcross 2 was produced for all the candidate QTLs. The obtained plant material will be used by other group members and should ultimately lead to discoveries of new recombination modifiers.

4 Conclusion

The studied accessions were sequenced with a 0.5X coverage. They show about 0.3% single nucleotide polymorphism. Longer indels and inversions were not considered. The availability of the different accessions and all the genetic and molecular biology tools makes *Arabidopsis thaliana* an accessible organism to study genomic evolution and adaptation in correlation to meiotic recombination or any other process of interest. Moreover, even with such a low considered diversity (0.3%), we are able to map recombination causative loci. This shows that a prospective much higher divergence can provide a very wide range of potential novel natural modifiers to be identified and characterized. Only two of the five tested accessions suggested a presence of possible meiotic recombination modifiers. This may suggest that the observed differences in crossover distribution, activity and count could be at least partially due to these potential modifiers. On the other hand, the differences observed in the other three populations could be due to the accumulated polymorphisms, especially structural variants (Cao et al., 2011; Lian et al., 2023).

The identification of recombination modifiers through hybrid segregation is a well-established approach as it was used in many studies (Dumont and Payseur, 2011; Fledel-Alon et al., 2011; Sandor et al., 2012; Kong et al., 2013; Ziolkowski et al., 2015; Hunter et al., 2016; Johnston et al., 2016; Kadri et al., 2016b; Wang and Payseur, 2017; Ziolkowski et al., 2017; Johnston et al., 2018b). The power of this study relies in the fact that it allows for a high throughput mapping for recombination frequency natural modifiers. We expect to analyze the BC2F2 and short list candidate genes before the end of 2023.

5 Bibliography

Baudat F, de Massy B (2007) Cis- and trans-acting elements regulate the mouse Psmb9 meiotic recombination hotspot. PLoS Genet **3**: 1029–1039

Ben-Ari G, Lavi U (2012) Marker-assisted selection in plant breeding. Plant Biotechnology and Agriculture. Elsevier, pp 163–184

Berchowitz LE, Hanlon SE, Lieb JD, Copenhaver GP (2009) A positive but complex association between meiotic double-strand break hotspots and open chromatin in Saccharomyces cerevisiae. Genome Res **19**: 2245–2257

Blackwell AR, Dluzewska J, Szymanska-Lejman M, Desjardins S, Tock AJ, Kbiri N, Lambing C, Lawrence EJ, Bieluszewski T, Rowan B, et al (2020) MSH 2 shapes the meiotic crossover landscape in relation to interhomolog polymorphism in Arabidopsis. EMBO J. doi: 10.15252/embj.2020104858

Bomblies K, Higgins JD, Yant L (2015) Meiosis evolves: Adaptation to external and internal environments. New Phytologist **208**: 306–323

Brick K, Pratto F, Sun C-Y, Camerini-Otero RD, Petukhova G (2018) Analysis of Meiotic Double-Strand Break Initiation in Mammals. Methods Enzymol. Academic Press Inc., pp 391–418

Brockhurst MA, Chapman T, King KC, Mank JE, Paterson S, Hurst GDD (2014) Running with the Red Queen: The role of biotic conflicts in evolution. Proceedings of the Royal Society B: Biological Sciences. doi: 10.1098/RSPB.2014.1382

Broman KW, Sen S (2009) A Guide to QTL Mapping with R/qtl. doi: 10.1007/978-0-387-92125-9

Broman KW, Wu H, Sen Ś, Churchill GA (2003) R/qtl: QTL mapping in experimental crosses. Bioinformatics **19**: 889–890

Cao J, Schneeberger K, Ossowski S, Günther T, Bender S, Fitz J, Koenig D, Lanz C, Stegle O, Lippert C, et al (2011) Whole-genome sequencing of multiple Arabidopsis thaliana populations. Nat Genet **43**: 956–965

Choi K, Henderson IR (2015) Meiotic recombination hotspots - A comparative view. Plant Journal **83**: 52–61

Choi K, Reinhard C, Serra H, Ziolkowski PA, Underwood CJ, Zhao X, Hardcastle TJ, Yelina NE, Griffin C, Jackson M, et al (2016) Recombination Rate Heterogeneity within Arabidopsis Disease Resistance Genes. PLoS Genet **12**: e1006179

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. Nat Genet **45**: 1327–1338

Choi K, Zhao X, Tock AJ, Lambing C, Underwood CJ, Hardcastle TJ, Serra H, Kim J, Cho HS, Kim J, et al (2018) Nucleosomes and DNA methylation shape meiotic DSB frequency in *Arabidopsis thaliana* transposons and gene regulatory regions. Genome Res **28**: 532–546

Cole F, Keeney S, Jasin M (2010) Comprehensive, Fine-Scale Dissection of Homologous Recombination Outcomes at a Hot Spot in Mouse Meiosis. Mol Cell **39**: 700–710

Crismani W, Girard C, Froger N, Pradillo M, Santos JL, Chelysheva L, Copenhaver GP, Horlow C, Mercier R (2012) FANCM limits meiotic crossovers. Science (1979) **336**: 1588–1590

Culligan KM, Britt AB (2008) Both ATM and ATR promote the efficient and accurate processing of programmed meiotic double-strand breaks. Plant Journal **55**: 629–638

Dapper AL, Payseur BA (2017) Connecting theory and data to understand recombination rate evolution. Philosophical Transactions of the Royal Society B: Biological Sciences. doi: 10.1098/RSTB.2016.0469

Dluzewska J, Szymanska M, Ziolkowski PA (2018) Where to Cross Over ? Defining Crossover Sites in Plants. **9**: 1–20

Dooner HK (1986) Genetic Fine Structure of the BRONZE Locus in Maize. Genetics **113**: 1021

Dreissig S, Maurer A, Sharma R, Milne L, Flavell AJ, Schmutzer T, 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

Dumont BL (2020) Evolution: Is Recombination Rate Variation Adaptive? Current Biology **30**: R351–R353

Dumont BL, Payseur BA (2011) Genetic Analysis of Genome-Scale Recombination Rate Evolution in House Mice. PLoS Genet **7**: e1002116

Egel R, Lankenau D-H, eds (2008) Recombination and Meiosis. Recombination and Meiosis. doi: 10.1007/978-3-540-75373-5

Feldman MW, Otto SP, Christiansen FB (1996) POPULATION GENETIC PERSPECTIVES ON THE EVOLUTION OF RECOMBINATION. Annu Rev Genet **30**: 261–295

Felsenstein J (1974) THE EVOLUTIONARY ADVANTAGE OF RECOMBINATION. Genetics **78**: 737–756

Fernandes JB, Séguéla-Arnaud M, Larchevêque C, Lloyd AH, Mercier R (2018) Unleashing meiotic crossovers in hybrid plants. Proceedings of the National Academy of Sciences **115**: 2431–2436

Fernandes JB, Wlodzimierz P, Henderson IR (2019) Meiotic recombination within plant centromeres. Curr Opin Plant Biol **48**: 26–35

Fischer O, Schmid-Hempel P (2005) Selection by parasites may increase host recombination frequency. Biol Lett **1**: 193–195

Fledel-Alon A, Leffler EM, Guan Y, Stephens M, Coop G, Przeworski M (2011) Variation in Human Recombination Rates and Its Genetic Determinants. PLoS One. doi: 10.1371/JOURNAL.PONE.0020321

Gray S, Cohen PE (2016) Control of Meiotic Crossovers: From Double-Strand Break Formation to Designation. Annu Rev Genet **50**: 175–210

Haenel Q, Laurentino TG, Roesti M, Berner D (2018) Meta-analysis of chromosome-scale crossover rate variation in eukaryotes and its significance to evolutionary genomics. Mol Ecol **27**: 2477–2497

Hamilton WD, Axelrodtt R, Tanese R (1990) Sexual reproduction as an adaptation to resist parasites (A Review) (evolution/recombination/population genetics/evolutionary genetics/disease resistance). Proc Nail Acad Sci USA **87**: 3566–3573

Henderson IR, Bomblies K (2021) Evolution and Plasticity of Genome-Wide Meiotic Recombination Rates. Annu Rev Genet **55**: 23–43

Hennig BP, Velten L, Racke I, Tu CS, Thoms M, Rybin V, Besir H, Remans K, Steinmetz LM (2018) Large-scale low-cost NGS library preparation using a robust Tn5 purification and tagmentation protocol. G3: Genes, Genomes, Genetics **8**: 79–89

Higgins JD, Perry RM, Barakate A, Ramsay L, Waugh R, Halpin C, Armstrong SJ, Franklin FCH (2012) Spatiotemporal Asymmetry of the Meiotic Program Underlies the Predominantly Distal Distribution of Meiotic Crossovers in Barley. Plant Cell **24**: 4096–4109

Hunter CM, Huang W, Mackay TFC, Singh ND (2016) The Genetic Architecture of Natural Variation in Recombination Rate in Drosophila melanogaster. PLoS Genet **12**: e1005951

Hunter N (2015) Meiotic recombination: The essence of heredity. Cold Spring Harb Perspect Biol **7**: 1–35

Hunter N (2007) Meiotic recombination. Molecular Genetics of Recombination. Springer, Berlin, Heidelberg, pp 381–442

Jeffreys AJ, Neumann R (2005) Factors influencing recombination frequency and distribution in a human meiotic crossover hotspot. Hum Mol Genet **14**: 2277–2287

Johnston SE, Bérénos C, Slate J, Pemberton JM (2016) Conserved Genetic Architecture Underlying Individual Recombination Rate Variation in a Wild Population of Soay Sheep (Ovis aries). Genetics **203**: 583–598

Johnston SE, Huisman J, Pemberton JM (2018a) A Genomic Region Containing *REC8* and *RNF212B* Is Associated with Individual Recombination Rate Variation in a Wild Population of Red Deer (*Cervus elaphus*). G3 Genes|Genomes|Genetics **8**: 2265–2276

Johnston SE, Huisman J, Pemberton JM (2018b) A Genomic Region Containing REC8 and RNF212B Is Associated with Individual Recombination Rate Variation in a Wild Population of Red Deer (Cervus elaphus). G3 Genes|Genomes|Genetics **8**: 2265–2276

Jones GH, Franklin FCH (2006) Meiotic Crossing-over: Obligation and Interference. Cell **126**: 246–248

Kadri NK, Harland C, Faux P, Cambisano N, Karim L, Coppieters W, Fritz S, Mullaart E, Baurain D, Boichard D, et al (2016a) Coding and noncoding variants in *HFM1* , *MLH3* , *MSH4* , *MSH5* , *RNF212* , and *RNF212B* affect recombination rate in cattle. Genome Res **26**: 1323–1332

Kadri NK, Harland C, Faux P, Cambisano N, Karim L, Coppieters W, Fritz S, Mullaart E, Baurain D, Boichard D, et al (2016b) Coding and noncoding variants in HFM1, MLH3, MSH4, MSH5, RNF212, and RNF212B affect recombination rate in cattle. Genome Res **26**: 1323–1332

Kbiri N, Dluzewska J, Henderson IR, Ziolkowski PA (2022) Quantifying Meiotic Crossover Recombination in Arabidopsis Lines Expressing Fluorescent Reporters in Seeds Using SeedScoring Pipeline for CellProfiler. Methods in Molecular Biology. Humana Press Inc., pp 121–134

Keeney S (2008) Spo11 and the Formation of DNA Double-Strand Breaks in Meiosis. Recombination and Meiosis. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 81–123

Kim J, Park J, Kim H, Son N, Kim E, Kim J, Byun D, Lee Y, Park YM, Nageswaran DC, et al (2022) *Arabidopsis* HEAT SHOCK FACTOR BINDING PROTEIN is required to limit meiotic crossovers and *HEI10* transcription. EMBO J. doi: 10.15252/embj.2021109958

Kong A, Thorleifsson G, Frigge ML, Masson G, Gudbjartsson DF, Villemoes R, Magnusdottir E, Olafsdottir SB, Thorsteinsdottir U, Stefansson K (2013) Common and low-frequency variants associated with genome-wide recombination rate. Nature Genetics 2013 46:1 **46**: 11–16

Lasky JR, des Marais DL, Lowry DB, Povolotskaya I, McKay JK, Richards JH, Keitt TH, Juenger TE (2014) Natural variation in abiotic stress responsive gene expression and local adaptation to climate in Arabidopsis thaliana. Mol Biol Evol **31**: 2283–2296

Lawrence EJ, Gao H, Tock AJ, Lambing C, Blackwell AR, Feng X, Henderson IR (2019) Natural Variation in TBP-ASSOCIATED FACTOR 4b Controls Meiotic Crossover and Germline Transcription in Arabidopsis. Current Biology **29**: 2676-2686.e3

Lee JY, Terakawa T, Qi Z, Steinfeld JB, Redding S, Kwon Y, Gaines WA, Zhao W, Sung P, Greene EC (2015) Base triplet stepping by the Rad51/RecA family of recombinases. Science (1979) **349**: 977–981

Lian Q, Hüttel B, Walkemeier B, Lopez-Roques C, Gil L, Roux F, Mercier R (2023) A pan-genome of 72 Arabidopsis thaliana accessions reveals a conserved genome structure throughout the global species range. doi: 10.21203/RS.3.RS-2976609/V1

Lian Q, Solier V, Walkemeier B, Durand S, Huettel B, Schneeberger K, Mercier R (2022) The megabase-scale crossover landscape is largely independent of sequence divergence. Nature Communications 2022 13:1 **13**: 1–11

Lichten M, Borts RH, Haber JE (1987) Meiotic Gene Conversion and Crossing over between Dispersed Homologous Sequences Occurs Frequently in Saccharomyces cerevisiae. Genetics **115**: 233

Lippman ZB, Semel Y, Zamir D (2007) An integrated view of quantitative trait variation using tomato interspecific introgression lines. Curr Opin Genet Dev **17**: 545–552

Lippman ZB, Zamir D (2007) Heterosis: revisiting the magic. Trends in Genetics **23**: 60–66

Lloyd A, Morgan C, Franklin FCH, Bomblies K (2018) Plasticity of meiotic recombination rates in response to temperature in arabidopsis. Genetics **208**: 1409–1420

López E, Pradillo M, Oliver C, Romero C, Cuñado N, Santos JL (2012) Looking for natural variation in chiasma frequency in Arabidopsis thaliana. J Exp Bot **63**: 887–94

Melamed-Bessudo C, Yehuda E, Stuitje AR, Levy AA (2005) A new seed-based assay for meiotic recombination in Arabidopsis thaliana. Plant Journal **43**: 458–466

Mercier R, Mézard C, Jenczewski E, Macaisne N, Grelon M (2015) The Molecular Biology of Meiosis in Plants. Annu Rev Plant Biol **66**: 297–327

Mieulet D, Aubert G, Bres C, Klein A, Droc G, Vieille E, Rond-Coissieux C, Sanchez M, Dalmais M, Mauxion J-P, et al (2018) Unleashing meiotic crossovers in crops. Nat Plants **4**: 1010–1016

Morgan CH, Zhang H, Bomblies K (2017) Are the effects of elevated temperature on meiotic recombination and thermotolerance linked via the axis and synaptonemal complex? Philosophical Transactions of the Royal Society B: Biological Sciences. doi: 10.1098/RSTB.2016.0470

Nageswaran DC, Kim J, Lambing C, Kim J, Park J, Kim E-J, Cho HS, Kim H, Byun D, Park YM, et al (2021) HIGH CROSSOVER RATE1 encodes PROTEIN PHOSPHATASE X1 and restricts meiotic crossovers in Arabidopsis. Nat Plants **7**: 452–467

Ono A, Kinoshita T (2021) Epigenetics and plant reproduction: Multiple steps for responsibly handling succession. Curr Opin Plant Biol. doi: 10.1016/j.pbi.2021.102032

Orkin SH, Kazazian HH, Antonarakis SE, Goff SC, Boehm CD, Sexton JP, Waber PG, Giardina PJ v. (1982) Linkage of β-thalassaemia mutations and β-globin gene polymorphisms with DNA polymorphisms in human β-globin gene cluster. Nature **296**: 627–631

Otto SP, Lenormand T (2002) Resolving the paradox of sex and recombination. Nat Rev Genet **3**: 252–261

Paiano J, Wu W, Yamada S, Sciascia N, Callen E, Paola Cotrim A, Deshpande RA, Maman Y, Day A, Paull TT, et al (2020) ATM and PRDM9 regulate SPO11-bound recombination intermediates during meiosis. Nat Commun **11**: 1–15

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

Picelli S, Björklund ÅK, Reinius B, Sagasser S, Winberg G, Sandberg R (2014) Tn5 transposase and tagmentation procedures for massively scaled sequencing projects. Genome Res **24**: 2033–2040

Protacio RU, Davidson MK, Wahls WP (2022) Adaptive Control of the Meiotic Recombination Landscape by DNA Site-dependent Hotspots With Implications for Evolution. Front Genet **0**: 1578

Rowan BA, Heavens D, Feuerborn TR, Tock AJ, Henderson IR, Weigel D (2019a) An ultra high-density arabidopsis thaliana crossover map that refines the influences of structural variation and epigenetic features. Genetics **213**: 771–787

Rowan BA, Heavens D, Feuerborn TR, Tock AJ, Henderson IR, Weigel D (2019b) An ultra high-density arabidopsis thaliana crossover map that refines the influences of structural variation and epigenetic features. Genetics **213**: 771–787

Rowan BA, Patel V, Weigel D, 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

Ruiz-Herrera A, Vozdova M, Fernández J, Sebestova H, Capilla L, Frohlich J, Vara C, Hernández-Marsal A, Sipek J, Robinson TJ, et al (2017) Recombination correlates with synaptonemal complex length and chromatin loop size in bovids-insights into mammalian meiotic chromosomal organization. Chromosoma **126**: 615–631

Salathé M, Kouyos RD, Regoes RR, Bonhoeffer S (2008) Rapid parasite adaptation drives selection for high recombination rates. Evolution (N Y) **62**: 295–300

Samuk K, Manzano-Winkler B, Ritz KR, Noor MAF (2020) Natural Selection Shapes Variation in Genome-wide Recombination Rate in Drosophila pseudoobscura. Current Biology **30**: 1517-1528.e6

Sandor C, Li W, Coppieters W, Druet T, Charlier C, Georges M (2012) Genetic Variants in REC8, RNF212, and PRDM9 Influence Male Recombination in Cattle. PLoS Genet **8**: e1002854

Schwander T, Libbrecht R, Keller L (2014) Supergenes and complex phenotypes. Current Biology **24**: R288–R294

Serra H, Choi K, Zhao X, Blackwell AR, Kim J, Henderson IR (2018a) Interhomolog polymorphism shapes meiotic crossover within the Arabidopsis RAC1 and RPP13 disease resistance genes. PLoS Genet **14**: e1007843

Serra H, Lambing C, Griffin CH, Topp SD, Nageswaran DC, Underwood CJ, Ziolkowski PA, Séguéla-Arnaud M, Fernandes JB, Mercier R, et al (2018b) Massive crossover elevation via combination of HEI10 and recq4a recq4b during Arabidopsis meiosis. Proceedings of the National Academy of Sciences **115**: 2437–2442

Serra H, Lambing C, Griffin CH, Topp SD, Nageswaran DC, Underwood CJ, Ziolkowski PA, Séguéla-Arnaud M, Fernandes JB, Mercier R, et al (2018c) Massive crossover elevation via combination of HEI10 and recq4a recq4b during Arabidopsis meiosis. Proceedings of the National Academy of Sciences **115**: 2437–2442

Serrentino M-E, Borde V (2012) The spatial regulation of meiotic recombination hotspots: Are all DSB hotspots crossover hotspots? Exp Cell Res **318**: 1347–1352

Smeenk G, Mailand N (2016) Writers, readers, and erasers of histone ubiquitylation in DNA double-strand break repair. Front Genet **7**: 1–14

de Storme N, Geelen D (2014) The impact of environmental stress on male reproductive development in plants: Biological processes and molecular mechanisms. Plant Cell Environ **37**: 1–18

Szymanska-Lejman M, Dziegielewski W, Dluzewska J, Kbiri N, Bieluszewska A, Poethig RS, Ziolkowski PA (2023) The effect of DNA polymorphisms and natural variation on crossover hotspot activity in Arabidopsis hybrids. Nat Commun **14**: 33

Takeda S, Paszkowski J (2006) DNA methylation and epigenetic inheritance during plant gametogenesis. Chromosoma **115**: 27–35

Thompson MJ, Jiggins CD (2014) Supergenes and their role in evolution. Heredity 2014 113:1 **113**: 1–8

Tian M, Loidl J (2018) A chromatin-associated protein required for inducing and limiting meiotic DNA double-strand break. **46**: 11822–11834

Timberlake WE (2013) Heterosis. Brenner's Encyclopedia of Genetics. Elsevier, pp 451–453

Underwood CJ, Choi K (2019) Heterogeneous transposable elements as silencers, enhancers and targets of meiotic recombination. Chromosoma **128**: 279–296

Wang RJ, Payseur BA (2017) Genetics of genome-wide recombination rate evolution in mice from an isolated Island. Genetics **206**: 1841–1852

Wang S, Zickler D, Kleckner N, Zhang L (2015) Meiotic crossover patterns: Obligatory crossover, interference and homeostasis in a single process. Cell Cycle **14**: 305–314

Wang Y, Copenhaver GP (2018) Meiotic Recombination: Mixing It Up in Plants. Annu Rev Plant Biol **69**: 577–609

Wu G, Rossidivito G, Hu T, Berlyand Y, Poethig RS (2015) Traffic lines: New tools for genetic analysis in Arabidopsis thaliana. Genetics. doi: 10.1534/genetics.114.173435

Xue M, Wang J, Jiang L, Wang M, Wolfe S, Pawlowski WP, Wang Y, He Y (2018) The Number of Meiotic Double-Strand Breaks Influences Crossover Distribution in Arabidopsis. Plant Cell **30**: 2628–2638

Yelina N, Diaz P, Lambing C, Henderson IR (2015a) Epigenetic control of meiotic recombination in plants. Sci China Life Sci **58**: 223–231

Yelina N, Diaz P, Lambing C, Henderson IR (2015b) Epigenetic control of meiotic recombination in plants. Sci China Life Sci **58**: 223–231

Yelina NE, Lambing C, Hardcastle TJ, Zhao X, Santos B, Henderson IR (2015c) DNA methylation epigenetically silences crossover hot spots and controls chromosomal domains of meiotic recombination in Arabidopsis. Genes Dev **29**: 2183–2202

Yelina NE, Ziolkowski PA, Miller N, Zhao X, Kelly KA, Muñoz DF, Mann DJ, Copenhaver GP, Henderson IR (2013) High-throughput analysis of meiotic crossover frequency and interference via flow cytometry of fluorescent pollen in Arabidopsis thaliana. Nat Protoc **8**: 2119–2134

Zamariola L, Tiang CL, de Storme N, Pawlowski W, Geelen D (2014) Chromosome segregation in plant meiosis. Front Plant Sci **5**: 279

Zhu L, Fernández-Jiménez N, Szymanska-Lejman M, Pelé A, Underwood CJ, Serra H, Lambing C, Dluzewska J, Bieluszewski T, Pradillo M, et al (2021)

Natural variation identifies SNI1, the SMC5/6 component, as a modifier of meiotic crossover in Arabidopsis. Proc Natl Acad Sci U S A. doi: 10.1073/PNAS.2021970118/-/DCSUPPLEMENTAL

Zickler D, Kleckner N (2015) Recombination, Pairing, and Synapsis of Homologs during Meiosis. Cold Spring Harb Perspect Biol **7**: a016626

Ziolkowski PA (2022) Why do plants need the ZMM crossover pathway? A snapshot of meiotic recombination from the perspective of interhomolog polymorphism. Plant Reprod **1**: 3

Ziolkowski PA, Berchowitz LE, Lambing C, Yelina NE, Zhao X, Kelly KA, Choi K, Ziolkowska L, June V, Sanchez-Moran E, et al (2015) Juxtaposition of heterozygous and homozygous regions causes reciprocal crossover remodelling via interference during Arabidopsis meiosis. Elife **4**: 1–29

Ziolkowski PA, Henderson IR (2017) Interconnections between meiotic recombination and sequence polymorphism in plant genomes. New Phytologist **213**: 1022–1029

Ziolkowski PA, Underwood CJ, Lambing C, Martinez-Garcia M, Lawrence EJ, Ziolkowska L, Griffin C, Choi K, Franklin FCH, Martienssen RA, et al (2017) Natural variation and dosage of the HEI10 meiotic E3 ligase control Arabidopsis crossover recombination. Genes Dev **31**: 306–317

6 Supplemental data

6.1 Recombination frequency measurements in 420 for the five F2 segregating populations

Recombination frequency was measured in 420 for all individual that provided enough seeds. Only individuals with mendelaen segregations of the green and red fluorescent tags were kept for the QTL mapping (G/nonG and R/nonR between 2.6 and 3.4, the values that fit within the interval are colored in green).

Supplemental table 1. Recombination frequency measurements for CDM-0 x 420 1.1.1

Individual	Green	Red	Both	None	Total	None/Total	RF(%)	G/non G	R/non R	G/T	R/T	G/R
1A												
1B	149	139	1124	266	1678	0,17	**18,96**	3,14	3,04	0,09	0,08	1,07
1C	155	122	1154	329	1760	0,16	**17,22**	2,90	2,64	0,09	0,07	1,27
1D	119	158	1312	389	1978	0,14	**15,15**	2,62	2,89	0,06	0,08	0,75
1E	149	144	1253	312	1858	0,16	**17,26**	3,07	3,03	0,08	0,08	1,03
1F	182	179	1357	284	2002	0,18	**20,04**	3,32	3,30	0,09	0,09	1,02
1G	174	154	1095	279	1702	0,19	**21,61**	2,93	2,76	0,10	0,09	1,13
1H	195	182	1158	299	1834	0,21	**23,26**	2,81	2,71	0,11	0,10	1,07
2A	89	112	741	192	1134	0,18	**19,66**	2,73	3,04	0,08	0,10	0,79
2B	117	100	1190	303	1710	0,13	**13,62**	3,24	3,07	0,07	0,06	1,17
2C	190	102	919	173	1384	0,21	**23,97**	4,03	2,81	0,14	0,07	1,86
2D	193	162	1098	247	1700	0,21	**23,69**	3,16	2,86	0,11	0,10	1,19
2E	127	156	1198	321	1802	0,16	**17,18**	2,78	3,02	0,07	0,09	0,81
2F	117	128	1056	288	1589	0,15	**16,84**	2,82	2,92	0,07	0,08	0,91
2G	111	108	1080	279	1578	0,14	**15,00**	3,08	3,05	0,07	0,07	1,03
2H	106	109	1078	332	1625	0,13	**14,25**	2,68	2,71	0,07	0,07	0,97
3A	134	127	1121	285	1667	0,16	**17,12**	3,05	2,98	0,08	0,08	1,06
3B												
3C	147	151	1125	279	1702	0,18	**19,39**	2,96	3,00	0,09	0,09	0,97
3D	175	148	1083	258	1664	0,19	**21,78**	3,10	2,84	0,11	0,09	1,18
3E	127	147	1177	332	1783	0,15	**16,77**	2,72	2,88	0,07	0,08	0,86
3F	161	137	1009	218	1525	0,20	**21,95**	3,30	3,02	0,11	0,09	1,18
3G	198	160	1097	223	1678	0,21	**24,28**	3,38	2,99	0,12	0,10	1,24
3H	159	141	1265	286	1851	0,16	**17,79**	3,33	3,16	0,09	0,08	1,13
4A	129	124	1081	301	1635	0,15	**16,90**	2,85	2,80	0,08	0,08	1,04
4B	103	120	1118	332	1673	0,13	**14,36**	2,70	2,85	0,06	0,07	0,86
4C	255	270	525	20	1070	0,49	**13,67**	2,69	2,89	0,24	0,25	0,94
4D	179	130	1090	277	1676	0,18	**20,55**	3,12	2,68	0,11	0,08	1,38
4E	108	112	962	281	1463	0,15	**16,38**	2,72	2,76	0,07	0,08	0,96
4F	160	154	1201	303	1818	0,17	**19,09**	2,98	2,93	0,09	0,08	1,04
4G	162	150	1077	295	1684	0,19	**20,66**	2,78	2,68	0,10	0,09	1,08
4H	93	108	1136	326	1663	0,12	**12,92**	2,83	2,97	0,06	0,06	0,86
5A	136	152	1135	262	1685	0,17	**18,87**	3,07	3,23	0,08	0,09	0,89
5B												

5C	51	49	441	107	648	0,15	**16,85**	3,15	3,10	0,08	0,08	1,04
5D												
5E	82	104	931	260	1377	0,14	**14,57**	2,78	3,03	0,06	0,08	0,79
5F	226	267	1152	237	1882	0,26	**31,00**	2,73	3,06	0,12	0,14	0,85
5G	126	165	1115	290	1696	0,17	**18,95**	2,73	3,08	0,07	0,10	0,76
5H	170	141	1220	335	1866	0,17	**18,35**	2,92	2,70	0,09	0,08	1,21
6A	131	126	1174	298	1729	0,15	**16,17**	3,08	3,03	0,08	0,07	1,04
6B	93	97	709	175	1074	0,18	**19,61**	2,95	3,01	0,09	0,09	0,96
6C	77	119	840	219	1255	0,16	**17,08**	2,71	3,24	0,06	0,09	0,65
6D	0	56	420	0	476	0,12	**12,55**	7,50	#DIV/0!	0,00	0,12	0,00
6E	62	62	630	203	957	0,13	**13,93**	2,61	2,61	0,06	0,06	1,00
6F	90	84	1014	330	1518	0,11	**12,21**	2,67	2,61	0,06	0,06	1,07
6G												
6H	90	113	1163	334	1700	0,12	**12,75**	2,80	3,01	0,05	0,07	0,80
7A	145	180	1017	223	1565	0,21	**23,54**	2,88	3,25	0,09	0,12	0,81
7B	314	139	1376	243	2072	0,22	**24,98**	4,42	2,72	0,15	0,07	2,26
7C	110	96	676	147	1029	0,20	**22,57**	3,23	3,00	0,11	0,09	1,15
7D	98	107	1140	274	1619	0,13	**13,58**	3,25	3,35	0,06	0,07	0,92
7E	103	92	1064	310	1569	0,12	**13,31**	2,90	2,80	0,07	0,06	1,12
7F	102	126	1125	327	1680	0,14	**14,64**	2,71	2,92	0,06	0,08	0,81
7G	139	130	981	230	1480	0,18	**20,22**	3,11	3,01	0,09	0,09	1,07
7H	390	79	1122	55	1646	0,28	**34,42**	11,28	2,70	0,24	0,05	4,94
8A	107	105	1038	278	1528	0,14	**15,00**	2,99	2,97	0,07	0,07	1,02
8B	65	65	694	202	1026	0,13	**13,59**	2,84	2,84	0,06	0,06	1,00
8C	62	39	539	149	789	0,13	**13,75**	3,20	2,74	0,08	0,05	1,59
8D	31	34	283	82	430	0,15	**16,47**	2,71	2,81	0,07	0,08	0,91
8E	141	135	1029	291	1596	0,17	**19,12**	2,75	2,69	0,09	0,08	1,04
8F	94	88	670	152	1004	0,18	**20,16**	3,18	3,08	0,09	0,09	1,07
8G												
8H	142	136	1127	308	1713	0,16	**17,82**	2,86	2,81	0,08	0,08	1,04
9A	161	147	1320	290	1918	0,16	**17,61**	3,39	3,25	0,08	0,08	1,10
9B												
9C	88	63	597	160	908	0,17	**18,31**	3,07	2,66	0,10	0,07	1,40
9D	124	112	1082	280	1598	0,15	**16,06**	3,08	2,96	0,08	0,07	1,11
9E	262	168	1306	259	1995	0,22	**24,57**	3,67	2,83	0,13	0,08	1,56
9F												
9G	138	113	1221	297	1769	0,14	**15,37**	3,31	3,07	0,08	0,06	1,22
9H	168	159	1015	248	1590	0,21	**23,27**	2,91	2,82	0,11	0,10	1,06
10A												
10B												
10C	105	100	1112	301	1618	0,13	**13,59**	3,03	2,99	0,06	0,06	1,05
10D												
10E	121	119	1070	290	1600	0,15	**16,33**	2,91	2,89	0,08	0,07	1,02
10F												

10G	139	139	1094	314	1686	0,16	**18,13**	2,72	2,72	0,08	0,08	1,00
10H	144	157	1166	269	1736	0,17	**19,18**	3,08	3,20	0,08	0,09	0,92
11A	116	84	1079	292	1571	0,13	**13,66**	3,18	2,85	0,07	0,05	1,38
11B	152	152	1366	293	1963	0,15	**16,92**	3,41	3,41	0,08	0,08	1,00
11C												
11D	114	108	1047	303	1572	0,14	**15,29**	2,82	2,77	0,07	0,07	1,06
11E	164	96	1124	286	1670	0,16	**17,02**	3,37	2,71	0,10	0,06	1,71
11F	55	47	360	98	560	0,18	**20,27**	2,86	2,66	0,10	0,08	1,17
11G	107	104	1015	290	1516	0,14	**15,05**	2,85	2,82	0,07	0,07	1,03
11H												
12A	135	110	947	257	1449	0,17	**18,65**	2,95	2,70	0,09	0,08	1,23
12B												
12C												
12D	277	157	1217	241	1892	0,23	**26,43**	3,75	2,65	0,15	0,08	1,76
12E	58	55	538	160	811	0,14	**15,07**	2,77	2,72	0,07	0,07	1,05
12F	94	102	641	156	993	0,20	**22,20**	2,85	2,97	0,09	0,10	0,92
12G	154	155	1049	237	1595	0,19	**21,74**	3,07	3,08	0,10	0,10	0,99
12H	116	113	1048	304	1581	0,14	**15,72**	2,79	2,76	0,07	0,07	1,03
18A	119	137	1033	280	1569	0,16	**17,92**	2,76	2,93	0,08	0,09	0,87
18B	124	126	1076	260	1586	0,16	**17,25**	3,11	3,13	0,08	0,08	0,98
18C	99	120	1119	327	1665	0,13	**14,15**	2,72	2,91	0,06	0,07	0,83
18D	87	98	1064	315	1564	0,12	**12,63**	2,79	2,89	0,06	0,06	0,89
18E												
18F	91	98	1084	287	1560	0,12	**12,95**	3,05	3,13	0,06	0,06	0,93
18G	190	159	1146	307	1802	0,19	**21,73**	2,87	2,63	0,11	0,09	1,19
18H	82	79	1098	323	1582	0,10	**10,76**	2,94	2,91	0,05	0,05	1,04

Supplemental table 2. Recombination frequency measurements for CDM-0 x 420 1.1.3

Individual	Green	Red	Both	None	Total	None/Total	RF(%)	G/non G	R/non R	G/T	R/T	G/R
1A												
1B												
1C	143	117	1036	289	1585	0,16	**18,03**	2,90	2,67	0,09	0,07	1,22
1D	201	145	965	220	1531	0,23	**25,97**	3,19	2,64	0,13	0,09	1,39
1E	102	107	963	270	1442	0,14	**15,73**	2,82	2,88	0,07	0,07	0,95
1F	129	94	1135	316	1674	0,13	**14,35**	3,08	2,76	0,08	0,06	1,37
1G	161	158	1078	244	1641	0,19	**21,82**	3,08	3,05	0,10	0,10	1,02
1H	87	100	941	269	1397	0,13	**14,43**	2,79	2,92	0,06	0,07	0,87
2A	175	123	1220	320	1838	0,16	**17,80**	3,15	2,71	0,10	0,07	1,42
2B												
2C	137	118	989	245	1489	0,17	**18,91**	3,10	2,90	0,09	0,08	1,16
2D	144	182	911	254	1491	0,22	**24,99**	2,42	2,75	0,10	0,12	0,79
2E	128	106	922	206	1362	0,17	**18,98**	3,37	3,08	0,09	0,08	1,21
2F	188	186	1139	268	1781	0,21	**23,84**	2,92	2,91	0,11	0,10	1,01
2G	200	207	1296	338	2041	0,20	**22,46**	2,74	2,79	0,10	0,10	0,97

2H												
3A	186	152	1241	312	1891	0,18	**19,84**	3,08	2,80	0,10	0,08	1,22
3B												
3C	97	86	989	304	1476	0,12	**13,28**	2,78	2,68	0,07	0,06	1,13
3D	195	169	1221	291	1876	0,19	**21,77**	3,08	2,86	0,10	0,09	1,15
3E	392	7	1161	26	1586	0,25	**29,51**	47,06	2,79	0,25	0,00	56,00
3F	124	139	1308	363	1934	0,14	**14,68**	2,85	2,97	0,06	0,07	0,89
3G	128	133	918	252	1431	0,18	**20,30**	2,72	2,77	0,09	0,09	0,96
3H	143	105	1024	204	1476	0,17	**18,52**	3,78	3,25	0,10	0,07	1,36
4A	179	121	946	224	1470	0,20	**23,07**	3,26	2,65	0,12	0,08	1,48
4B	217	178	1189	301	1885	0,21	**23,78**	2,94	2,64	0,12	0,09	1,22
4C												
4D												
4E	209	277	1172	251	1909	0,25	**29,94**	2,62	3,15	0,11	0,15	0,75
4F	128	118	786	157	1189	0,21	**23,44**	3,32	3,17	0,11	0,10	1,08
4G												
4H												
5A	118	113	998	276	1505	0,15	**16,75**	2,87	2,82	0,08	0,08	1,04
5B	412	404	923	12	1751	0,47	**73,93**	3,21	3,13	0,24	0,23	1,02
5C	232	133	1365	274	2004	0,18	**20,27**	3,92	2,96	0,12	0,07	1,74
5D	139	137	1128	305	1709	0,16	**17,72**	2,87	2,85	0,08	0,08	1,01
5E												
5F	140	129	996	207	1472	0,18	**20,34**	3,38	3,24	0,10	0,09	1,09
5G	137	185	801	170	1293	0,25	**29,15**	2,64	3,21	0,11	0,14	0,74
5H	210	199	1245	299	1953	0,21	**23,77**	2,92	2,84	0,11	0,10	1,06
6A	203	164	1104	257	1728	0,21	**24,16**	3,10	2,76	0,12	0,09	1,24
6B	133	148	976	238	1495	0,19	**21,00**	2,87	3,03	0,09	0,10	0,90
6C	115	124	1211	312	1762	0,14	**14,64**	3,04	3,13	0,07	0,07	0,93
6D	139	141	947	262	1489	0,19	**21,01**	2,69	2,71	0,09	0,09	0,99
6E	114	137	894	245	1390	0,18	**20,07**	2,64	2,87	0,08	0,10	0,83
6F	137	109	986	278	1510	0,16	**17,89**	2,90	2,64	0,09	0,07	1,26
6G	129	109	1065	263	1566	0,15	**16,57**	3,21	2,99	0,08	0,07	1,18
6H	185	108	1115	284	1692	0,17	**19,15**	3,32	2,61	0,11	0,06	1,71
7A	433	364	847	18	1662	0,48	**79,77**	3,35	2,69	0,26	0,22	1,19
7B	400	8	1255	27	1690	0,24	**28,09**	47,29	2,96	0,24	0,00	50,00
7C	168	117	1351	343	1979	0,14	**15,62**	3,30	2,87	0,08	0,06	1,44
7D	104	131	900	208	1343	0,17	**19,38**	2,96	3,30	0,08	0,10	0,79
7E												
7F	226	242	1282	273	2023	0,23	**26,70**	2,93	3,05	0,11	0,12	0,93
7G	124	163	1085	246	1618	0,18	**19,67**	2,96	3,37	0,08	0,10	0,76
7H	324	53	1142	123	1642	0,23	**26,46**	8,33	2,67	0,20	0,03	6,11
8A	109	104	966	294	1473	0,14	**15,69**	2,70	2,66	0,07	0,07	1,05
8B	118	124	1047	295	1584	0,15	**16,67**	2,78	2,84	0,07	0,08	0,95
8C	197	153	1138	257	1745	0,20	**22,61**	3,26	2,84	0,11	0,09	1,29
8D	121	110	1010	262	1503	0,15	**16,78**	3,04	2,92	0,08	0,07	1,10

8E	226	190	1220	259	1895	0,22	**25,10**	3,22	2,91	0,12	0,10	1,19
8F	174	141	1067	254	1636	0,19	**21,58**	3,14	2,82	0,11	0,09	1,23
8G	118	102	1078	278	1576	0,14	**15,10**	3,15	2,98	0,07	0,06	1,16
8H												
9A	110	119	948	289	1466	0,16	**17,08**	2,59	2,67	0,08	0,08	0,92
9B	136	93	984	261	1474	0,16	**16,98**	3,16	2,71	0,09	0,06	1,46
9C	219	185	1218	285	1907	0,21	**24,09**	3,06	2,78	0,11	0,10	1,18
9D	149	156	902	227	1434	0,21	**24,20**	2,74	2,81	0,10	0,11	0,96
9E	169	128	1243	313	1853	0,16	**17,57**	3,20	2,84	0,09	0,07	1,32
9F	205	207	1231	258	1901	0,22	**24,73**	3,09	3,11	0,11	0,11	0,99
9G	213	146	1188	286	1833	0,20	**22,01**	3,24	2,67	0,12	0,08	1,46
9H	235	66	1164	192	1657	0,18	**20,21**	5,42	2,88	0,14	0,04	3,56
10A	165	143	1179	291	1778	0,17	**19,16**	3,10	2,90	0,09	0,08	1,15
10B												
10C	388	7	1165	40	1600	0,25	**28,85**	33,04	2,74	0,24	0,00	55,43
10D	133	100	804	214	1251	0,19	**20,79**	2,98	2,61	0,11	0,08	1,33
10E	223	210	1139	258	1830	0,24	**27,42**	2,91	2,80	0,12	0,11	1,06
10F	132	126	1030	276	1564	0,16	**18,14**	2,89	2,83	0,08	0,08	1,05
10G												
10H	154	144	1141	304	1743	0,17	**18,88**	2,89	2,81	0,09	0,08	1,07
11A	207	193	1163	294	1857	0,22	**24,55**	2,81	2,71	0,11	0,10	1,07
11B	193	204	1219	291	1907	0,21	**23,60**	2,85	2,94	0,10	0,11	0,95
11C	200	177	1172	286	1835	0,21	**23,25**	2,96	2,78	0,11	0,10	1,13
11D	245	91	1255	246	1837	0,18	**20,36**	4,45	2,74	0,13	0,05	2,69
11E	269	114	1137	246	1766	0,22	**24,75**	3,91	2,43	0,15	0,06	2,36
11F	223	195	1157	289	1864	0,22	**25,74**	2,85	2,64	0,12	0,10	1,14
11G	158	168	1142	306	1774	0,18	**20,47**	2,74	2,82	0,09	0,09	0,94
11H												

Supplemental table 3. Recombination frequency measurements for Co-1 x 420 1.1.2

Individual	Green	Red	Both	None	Total	None/Total	RF(%)	G/non G	R/non R	G/T	R/T	G/R
1A	133	100	1114	284	1631	0,14	**15,48**	3,25	2,91	0,08	0,06	1,33
1B	149	150	1021	285	1605	0,19	**20,79**	2,69	2,70	0,09	0,09	0,99
1C	155	121	1037	275	1588	0,17	**19,23**	3,01	2,69	0,10	0,08	1,28
1D	126	134	1020	244	1524	0,17	**18,83**	3,03	3,12	0,08	0,09	0,94
1E												
1F												
1G												
1H	118	112	1111	317	1658	0,14	**15,00**	2,86	2,81	0,07	0,07	1,05
2A	164	120	1073	288	1645	0,17	**19,09**	3,03	2,64	0,10	0,07	1,37
2B	133	139	1017	251	1540	0,18	**19,58**	2,95	3,01	0,09	0,09	0,96
2C	177	133	1015	233	1558	0,20	**22,41**	3,26	2,80	0,11	0,09	1,33
2D	145	113	1068	269	1595	0,16	**17,75**	3,18	2,85	0,09	0,07	1,28
2E	184	134	1233	288	1839	0,17	**19,12**	3,36	2,90	0,10	0,07	1,37

2F	152	113	987	230	1482	0,18	**19,85**	3,32	2,88	0,10	0,08	1,35
2G	139	126	1124	265	1654	0,16	**17,56**	3,23	3,09	0,08	0,08	1,10
2H	124	77	1098	313	1612	0,12	**13,36**	3,13	2,69	0,08	0,05	1,61
3A	169	136	1116	256	1677	0,18	**20,23**	3,28	2,95	0,10	0,08	1,24
3B	175	111	1060	270	1616	0,18	**19,62**	3,24	2,63	0,11	0,07	1,58
3C	149	139	1122	267	1677	0,17	**18,97**	3,13	3,03	0,09	0,08	1,07
3D	118	131	1069	305	1623	0,15	**16,74**	2,72	2,84	0,07	0,08	0,90
3E	149	113	1103	269	1634	0,16	**17,58**	3,28	2,91	0,09	0,07	1,32
3F	89	88	991	262	1430	0,12	**13,26**	3,09	3,07	0,06	0,06	1,01
3G	155	127	991	269	1542	0,18	**20,36**	2,89	2,64	0,10	0,08	1,22
3H	107	125	984	238	1454	0,16	**17,48**	3,01	3,21	0,07	0,09	0,86
4A	145	99	948	233	1425	0,17	**18,91**	3,29	2,77	0,10	0,07	1,46
4B	145	146	1068	246	1605	0,18	**20,16**	3,09	3,10	0,09	0,09	0,99
4C	184	134	1032	227	1577	0,20	**22,75**	3,37	2,84	0,12	0,08	1,37
4D	160	156	1073	266	1655	0,19	**21,38**	2,92	2,88	0,10	0,09	1,03
4E												
4F	133	119	1010	293	1555	0,16	**17,79**	2,77	2,65	0,09	0,08	1,12
4G	162	130	964	226	1482	0,20	**22,16**	3,16	2,82	0,11	0,09	1,25
4H	95	81	980	245	1401	0,13	**13,47**	3,30	3,12	0,07	0,06	1,17
5A												
5B	143	147	1125	278	1693	0,17	**18,92**	2,98	3,02	0,08	0,09	0,97
5C												
5D												
5E	165	175	1079	264	1683	0,20	**22,80**	2,83	2,92	0,10	0,10	0,94
5F	150	117	1117	277	1661	0,16	**17,63**	3,22	2,89	0,09	0,07	1,28
5G	142	108	1101	274	1625	0,15	**16,79**	3,25	2,91	0,09	0,07	1,31
5H	144	119	1090	267	1620	0,16	**17,82**	3,20	2,94	0,09	0,07	1,21
6A	195	157	1142	237	1731	0,20	**22,97**	3,39	3,01	0,11	0,09	1,24
6B	137	144	1112	269	1662	0,17	**18,65**	3,02	3,09	0,08	0,09	0,95
6C	194	145	1139	269	1747	0,19	**21,78**	3,22	2,77	0,11	0,08	1,34
6D												
6E	76	75	1014	316	1481	0,10	**10,78**	2,79	2,78	0,05	0,05	1,01
6F	103	89	1037	255	1484	0,13	**13,90**	3,31	3,15	0,07	0,06	1,16
6G	149	137	1115	272	1673	0,17	**18,88**	3,09	2,97	0,09	0,08	1,09
6H	90	80	1128	301	1599	0,11	**11,27**	3,20	3,09	0,06	0,05	1,13
1E	2	87	749	5	843	0,11	**11,18**	8,16	119,43	0,00	0,10	0,02
1G	2	81	836	4	923	0,09	**9,44**	9,86	152,83	0,00	0,09	0,02
7A	168	182	1151	289	1790	0,20	**21,97**	2,80	2,92	0,09	0,10	0,92
7B	100	108	1070	298	1576	0,13	**14,21**	2,88	2,96	0,06	0,07	0,93
7C	123	155	1112	292	1682	0,17	**18,18**	2,76	3,05	0,07	0,09	0,79
7D	167	161	1130	274	1732	0,19	**21,18**	2,98	2,93	0,10	0,09	1,04
7E	145	126	1117	302	1690	0,16	**17,58**	2,95	2,78	0,09	0,07	1,15
7F	148	149	896	239	1432	0,21	**23,50**	2,69	2,70	0,10	0,10	0,99
7G	133	132	1111	266	1642	0,16	**17,71**	3,13	3,12	0,08	0,08	1,01

7H	119	144	1095	288	1646	0,16	**17,51**	2,81	3,04	0,07	0,09	0,83
8A	127	117	853	288	1385	0,18	**19,52**	2,42	2,34	0,09	0,08	1,09
8B	203	122	1064	276	1665	0,20	**21,92**	3,18	2,48	0,12	0,07	1,66
8C	146	97	909	354	1506	0,16	**17,70**	2,34	2,01	0,10	0,06	1,51
8D	125	132	1075	244	1576	0,16	**17,91**	3,19	3,27	0,08	0,08	0,95
8E	169	146	1155	264	1734	0,18	**20,21**	3,23	3,00	0,10	0,08	1,16
8F	0	73	761	8	842	0,09	**9,08**	9,40	104,25	0,00	0,09	0,00
8G	136	139	1036	255	1566	0,18	**19,45**	2,97	3,01	0,09	0,09	0,98
8H	117	118	1076	270	1581	0,15	**16,17**	3,07	3,09	0,07	0,07	0,99
9A	0	89	692	2	783	0,11	**12,10**	7,60	390,50	0,00	0,11	0,00
9B	203	163	1059	215	1640	0,22	**25,59**	3,34	2,92	0,12	0,10	1,25
9C	152	137	1132	283	1704	0,17	**18,71**	3,06	2,92	0,09	0,08	1,11
9D	145	179	958	236	1518	0,21	**24,30**	2,66	2,98	0,10	0,12	0,81
9E	154	125	1020	258	1557	0,18	**19,90**	3,07	2,78	0,10	0,08	1,23
9F	153	160	1196	306	1815	0,17	**19,06**	2,89	2,95	0,08	0,09	0,96
9G	121	80	1077	304	1582	0,13	**13,64**	3,12	2,72	0,08	0,05	1,51
9H	127	110	1087	281	1605	0,15	**16,06**	3,10	2,93	0,08	0,07	1,15
10A	117	121	1006	280	1524	0,16	**17,07**	2,80	2,84	0,08	0,08	0,97
10B	116	97	1104	295	1612	0,13	**14,23**	3,11	2,92	0,07	0,06	1,20
10C	110	117	1078	266	1571	0,14	**15,68**	3,10	3,18	0,07	0,07	0,94
10D	134	88	1012	261	1495	0,15	**16,15**	3,28	2,78	0,09	0,06	1,52
10E	131	138	999	275	1543	0,17	**19,30**	2,74	2,80	0,08	0,09	0,95
10F	108	89	986	286	1469	0,13	**14,46**	2,92	2,73	0,07	0,06	1,21
10G	137	86	1128	299	1650	0,14	**14,58**	3,29	2,78	0,08	0,05	1,59
10H	116	126	947	220	1409	0,17	**18,98**	3,07	3,19	0,08	0,09	0,92
11A	145	164	1197	284	1790	0,17	**19,08**	3,00	3,17	0,08	0,09	0,88
11B	113	86	997	254	1450	0,14	**14,82**	3,26	2,95	0,08	0,06	1,31
11C	122	116	1178	302	1718	0,14	**14,97**	3,11	3,05	0,07	0,07	1,05
11D	154	104	1068	275	1601	0,16	**17,68**	3,22	2,73	0,10	0,06	1,48
11E	113	132	1114	279	1638	0,15	**16,28**	2,99	3,18	0,07	0,08	0,86
11F	124	117	1066	312	1619	0,15	**16,20**	2,77	2,71	0,08	0,07	1,06
11G	164	115	1048	277	1604	0,17	**19,25**	3,09	2,64	0,10	0,07	1,43
11H	129	111	1072	307	1619	0,15	**16,12**	2,87	2,71	0,08	0,07	1,16
12A												
12B	159	154	1181	276	1770	0,18	**19,61**	3,12	3,07	0,09	0,09	1,03
12C	126	129	1041	298	1594	0,16	**17,53**	2,73	2,76	0,08	0,08	0,98
12D	150	118	1119	256	1643	0,16	**17,92**	3,39	3,05	0,09	0,07	1,27
12E	148	118	1072	268	1606	0,17	**18,22**	3,16	2,86	0,09	0,07	1,25
12F	98	115	1104	324	1641	0,13	**13,95**	2,74	2,89	0,06	0,07	0,85
12G	160	139	1175	258	1732	0,17	**19,08**	3,36	3,14	0,09	0,08	1,15
12H	155	132	1111	277	1675	0,17	**18,93**	3,10	2,88	0,09	0,08	1,17

Supplemental table 4. Recombination frequency measurements for Co-1 x 420 1.1.3

Individual	Green	Red	Both	None	Total	None/Total	RF(%)	G/non G	R/non R	G/T	R/T	G/R

1A												
1B												
1C	143	97	1158	316	1714	0,14	**15,15**	3,15	2,73	0,08	0,06	1,47
1D	114	101	880	211	1306	0,16	**18,10**	3,19	3,02	0,09	0,08	1,13
1E	90	109	955	225	1379	0,14	**15,66**	3,13	3,38	0,07	0,08	0,83
1F	163	142	1164	320	1789	0,17	**18,82**	2,87	2,70	0,09	0,08	1,15
1G	125	118	997	251	1491	0,16	**17,90**	3,04	2,97	0,08	0,08	1,06
1H												
2A	149	170	1149	268	1736	0,18	**20,47**	2,96	3,16	0,09	0,10	0,88
2B	112	154	1088	248	1602	0,17	**18,27**	2,99	3,45	0,07	0,10	0,73
2C												
2D	113	80	704	165	1062	0,18	**20,22**	3,33	2,82	0,11	0,08	1,41
2E												
2F	73	89	998	252	1412	0,11	**12,22**	3,14	3,34	0,05	0,06	0,82
2G	212	187	1173	228	1800	0,22	**25,39**	3,34	3,09	0,12	0,10	1,13
2H	120	111	907	256	1394	0,17	**18,23**	2,80	2,71	0,09	0,08	1,08
3A	84	74	1012	305	1475	0,11	**11,36**	2,89	2,79	0,06	0,05	1,14
3B	142	158	1087	265	1652	0,18	**20,20**	2,91	3,06	0,09	0,10	0,90
3C	139	169	1084	274	1666	0,18	**20,61**	2,76	3,03	0,08	0,10	0,82
3D	106	106	1085	265	1562	0,14	**14,64**	3,21	3,21	0,07	0,07	1,00
3E	142	139	918	298	1497	0,19	**20,97**	2,43	2,40	0,09	0,09	1,02
3F	123	121	1057	273	1574	0,16	**16,94**	2,99	2,97	0,08	0,08	1,02
3G	141	129	1037	277	1584	0,17	**18,82**	2,90	2,79	0,09	0,08	1,09
3H	175	137	1066	229	1607	0,19	**21,79**	3,39	2,98	0,11	0,09	1,28
4A	116	127	1191	314	1748	0,14	**15,03**	2,96	3,07	0,07	0,07	0,91
4B	124	136	1014	251	1525	0,17	**18,82**	2,94	3,07	0,08	0,09	0,91
4C	99	83	1016	265	1463	0,12	**13,33**	3,20	3,02	0,07	0,06	1,19
4D	94	89	964	255	1402	0,13	**14,04**	3,08	3,02	0,07	0,06	1,06
4E	111	166	1007	258	1542	0,18	**19,95**	2,64	3,18	0,07	0,11	0,67
4F	124	115	1046	285	1570	0,15	**16,60**	2,93	2,84	0,08	0,07	1,08
4G	92	87	923	266	1368	0,13	**14,08**	2,88	2,82	0,07	0,06	1,06
4H	157	203	1060	248	1668	0,22	**24,61**	2,70	3,12	0,09	0,12	0,77
5A	102	76	1012	309	1499	0,12	**12,68**	2,89	2,65	0,07	0,05	1,34
5B	108	97	959	235	1399	0,15	**15,92**	3,21	3,08	0,08	0,07	1,11
5C	166	181	1103	228	1678	0,21	**23,42**	3,10	3,26	0,10	0,11	0,92
5D	107	107	1038	287	1539	0,14	**15,04**	2,91	2,91	0,07	0,07	1,00
5E	124	131	1186	303	1744	0,15	**15,88**	3,02	3,08	0,07	0,08	0,95
5F	134	116	1008	232	1490	0,17	**18,49**	3,28	3,07	0,09	0,08	1,16
5G	149	167	976	340	1632	0,19	**21,72**	2,22	2,34	0,09	0,10	0,89
5H	120	133	1023	235	1511	0,17	**18,44**	3,11	3,26	0,08	0,09	0,90
6A	114	114	1000	285	1513	0,15	**16,42**	2,79	2,79	0,08	0,08	1,00
6B	75	102	940	254	1371	0,13	**13,87**	2,85	3,17	0,05	0,07	0,74
6C	79	113	975	274	1441	0,13	**14,35**	2,72	3,08	0,05	0,08	0,70
6D	127	133	1046	318	1624	0,16	**17,55**	2,60	2,65	0,08	0,08	0,95
6E	121	135	1108	254	1618	0,16	**17,32**	3,16	3,31	0,07	0,08	0,90

6F	103	126	1079	301	1609	0,14	**15,42**	2,77	2,98	0,06	0,08	0,82
6G	137	130	961	234	1462	0,18	**20,33**	3,02	2,94	0,09	0,09	1,05
6H	135	120	1046	264	1565	0,16	**17,90**	3,08	2,92	0,09	0,08	1,13
7A	160	141	965	244	1510	0,20	**22,45**	2,92	2,74	0,11	0,09	1,13
7B	158	136	981	254	1529	0,19	**21,55**	2,92	2,71	0,10	0,09	1,16
7C	108	132	1091	277	1608	0,15	**16,24**	2,93	3,18	0,07	0,08	0,82
7D	120	114	1015	261	1510	0,15	**16,93**	3,03	2,96	0,08	0,08	1,05
7E	84	85	818	234	1221	0,14	**14,96**	2,83	2,84	0,07	0,07	0,99
7F	87	115	1000	286	1488	0,14	**14,65**	2,71	2,99	0,06	0,08	0,76
7G	91	107	1050	287	1535	0,13	**13,86**	2,90	3,06	0,06	0,07	0,85
7H	183	178	1111	269	1741	0,21	**23,50**	2,89	2,85	0,11	0,10	1,03
8A	126	138	1001	217	1482	0,18	**19,77**	3,17	3,32	0,09	0,09	0,91
8B	81	92	1047	306	1526	0,11	**12,06**	2,83	2,94	0,05	0,06	0,88
8C	109	126	1132	312	1679	0,14	**15,14**	2,83	2,99	0,06	0,08	0,87
8D	137	122	1130	285	1674	0,15	**16,90**	3,11	2,97	0,08	0,07	1,12
8E	118	117	1067	306	1608	0,15	**15,87**	2,80	2,79	0,07	0,07	1,01
8F	123	156	1055	263	1597	0,17	**19,34**	2,81	3,14	0,08	0,10	0,79
8G	190	150	1120	249	1709	0,20	**22,40**	3,28	2,89	0,11	0,09	1,27
8H	122	136	1042	269	1569	0,16	**18,08**	2,87	3,01	0,08	0,09	0,90
9A	111	124	997	259	1491	0,16	**17,25**	2,89	3,03	0,07	0,08	0,90
9B	107	115	1048	266	1536	0,14	**15,68**	3,03	3,12	0,07	0,07	0,93
9C	121	133	1115	284	1653	0,15	**16,77**	2,96	3,08	0,07	0,08	0,91
9D	155	173	1097	257	1682	0,20	**21,90**	2,91	3,08	0,09	0,10	0,90
9E	112	117	1050	235	1514	0,15	**16,48**	3,30	3,36	0,07	0,08	0,96
9F	126	124	1130	324	1704	0,15	**15,94**	2,80	2,79	0,07	0,07	1,02
9G	137	139	992	288	1556	0,18	**19,67**	2,64	2,66	0,09	0,09	0,99
9H	361	380	758	10	1509	0,49	**86,62**	2,87	3,07	0,24	0,25	0,95
10A												
10B	156	164	1088	275	1683	0,19	**21,28**	2,83	2,90	0,09	0,10	0,95
10C	97	113	1043	271	1524	0,14	**14,89**	2,97	3,14	0,06	0,07	0,86
10D	143	149	1116	285	1693	0,17	**19,06**	2,90	2,96	0,08	0,09	0,96
10E	122	175	1029	233	1559	0,19	**21,32**	2,82	3,39	0,08	0,11	0,70
10F	100	107	915	277	1399	0,15	**16,09**	2,64	2,71	0,07	0,08	0,93
10G	144	107	972	270	1493	0,17	**18,53**	2,96	2,61	0,10	0,07	1,35
10H	134	143	1111	259	1647	0,17	**18,54**	3,10	3,19	0,08	0,09	0,94
11A	82	99	884	271	1336	0,14	**14,62**	2,61	2,78	0,06	0,07	0,83
11B	143	154	1036	227	1560	0,19	**21,31**	3,09	3,22	0,09	0,10	0,93
11C	81	101	1136	323	1641	0,11	**11,79**	2,87	3,06	0,05	0,06	0,80
11D	113	131	1075	287	1606	0,15	**16,57**	2,84	3,02	0,07	0,08	0,86
11E	384	366	776	15	1541	0,49	**83,69**	3,04	2,86	0,25	0,24	1,05
11F	104	125	988	260	1477	0,16	**16,94**	2,84	3,06	0,07	0,08	0,83
11G	106	105	1008	273	1492	0,14	**15,31**	2,95	2,94	0,07	0,07	1,01
11H	144	165	1011	212	1532	0,20	**22,76**	3,06	3,30	0,09	0,11	0,87
1	113	93	1276	413	1895	0,11	**11,54**	2,75	2,60	0,06	0,05	1,22

2	115	119	1053	328	1615	0,14	**15,73**	2,61	2,65	0,07	0,07	0,97
3	86	105	1158	356	1705	0,11	**11,91**	2,70	2,86	0,05	0,06	0,82
4	133	144	1158	321	1756	0,16	**17,26**	2,78	2,87	0,08	0,08	0,92
5	156	159	1054	269	1638	0,19	**21,55**	2,83	2,85	0,10	0,10	0,98
6	136	112	1114	267	1629	0,15	**16,60**	3,30	3,04	0,08	0,07	1,21
7	128	142	1066	252	1588	0,17	**18,76**	3,03	3,18	0,08	0,09	0,90
8	132	140	1162	291	1725	0,16	**17,26**	3,00	3,08	0,08	0,08	0,94
9	104	127	1140	327	1698	0,14	**14,68**	2,74	2,94	0,06	0,07	0,82

Supplemental table 5. Recombination frequency measurements for Neo-6 x 420 1.2.2

Individual	Green	Red	Both	None	Total	None/Total	RF(%)	G/non G	R/non R	G/T	R/T	G/R
1A	151	135	1120	327	1733	0,17	**18,15**	2,75	2,63	0,09	0,08	1,12
1B	63	50	592	173	878	0,13	**13,83**	2,94	2,72	0,07	0,06	1,26
1C	101	124	1087	312	1624	0,14	**14,98**	2,72	2,93	0,06	0,08	0,81
1D	113	112	1028	263	1516	0,15	**16,14**	3,04	3,03	0,07	0,07	1,01
1E	76	232	660	274	1242	0,25	**29,01**	1,45	2,55	0,06	0,19	0,33
1F	73	95	909	279	1356	0,12	**13,27**	2,63	2,85	0,05	0,07	0,77
1G	116	129	891	235	1371	0,18	**19,84**	2,77	2,91	0,08	0,09	0,90
1H	106	115	989	262	1472	0,15	**16,35**	2,90	3,00	0,07	0,08	0,92
2A	125	151	1130	304	1710	0,16	**17,71**	2,76	2,99	0,07	0,09	0,83
2B	123	120	972	256	1471	0,17	**18,17**	2,91	2,88	0,08	0,08	1,03
2C	139	152	985	244	1520	0,19	**21,44**	2,84	2,97	0,09	0,10	0,91
2D	110	107	1039	271	1527	0,14	**15,40**	3,04	3,01	0,07	0,07	1,03
2E	125	172	1139	279	1715	0,17	**19,15**	2,80	3,25	0,07	0,10	0,73
2F	158	127	984	236	1505	0,19	**21,18**	3,15	2,82	0,10	0,08	1,24
2G	125	138	835	218	1316	0,20	**22,52**	2,70	2,84	0,09	0,10	0,91
2H	111	97	953	238	1399	0,15	**16,18**	3,18	3,01	0,08	0,07	1,14
3A	97	113	973	267	1450	0,14	**15,72**	2,82	2,98	0,07	0,08	0,86
3B	128	149	1137	309	1723	0,16	**17,63**	2,76	2,94	0,07	0,09	0,86
3C	110	111	1182	317	1720	0,13	**13,80**	3,02	3,03	0,06	0,06	0,99
3D												
3E	88	109	932	278	1407	0,14	**15,15**	2,64	2,84	0,06	0,08	0,81
3F	173	145	1158	215	1691	0,19	**21,01**	3,70	3,36	0,10	0,09	1,19
3G	116	123	921	226	1386	0,17	**19,06**	2,97	3,05	0,08	0,09	0,94
3H	145	145	1086	260	1636	0,18	**19,66**	3,04	3,04	0,09	0,09	1,00
4A	92	142	1126	313	1673	0,14	**15,13**	2,68	3,13	0,05	0,08	0,65
4B	112	153	1034	281	1580	0,17	**18,48**	2,64	3,02	0,07	0,10	0,73
4C	142	146	1068	263	1619	0,18	**19,74**	2,96	3,00	0,09	0,09	0,97
4D												
4E	142	138	916	236	1432	0,20	**21,97**	2,83	2,79	0,10	0,10	1,03
4F	125	165	1077	277	1644	0,18	**19,55**	2,72	3,09	0,08	0,10	0,76
4G	223	182	1262	267	1934	0,21	**23,76**	3,31	2,95	0,12	0,09	1,23
4H	85	95	990	288	1458	0,12	**13,22**	2,81	2,91	0,06	0,07	0,89
5A	172	157	1038	222	1589	0,21	**23,46**	3,19	3,03	0,11	0,10	1,10

5B	86	102	949	278	1415	0,13	**14,31**	2,72	2,89	0,06	0,07	0,84
5C												
5D	83	101	954	268	1406	0,13	**14,08**	2,81	3,01	0,06	0,07	0,82
5E	145	131	914	250	1440	0,19	**21,47**	2,78	2,65	0,10	0,09	1,11
5F												
5G	124	142	1045	276	1587	0,17	**18,47**	2,80	2,97	0,08	0,09	0,87
5H	152	130	1060	273	1615	0,17	**19,33**	3,01	2,80	0,09	0,08	1,17
6A	105	95	1001	259	1460	0,14	**14,79**	3,12	3,01	0,07	0,07	1,11
6B	160	150	990	236	1536	0,20	**22,78**	2,98	2,88	0,10	0,10	1,07
6C	137	136	1075	298	1646	0,17	**18,25**	2,79	2,78	0,08	0,08	1,01
6D	92	96	1109	349	1646	0,11	**12,16**	2,70	2,73	0,06	0,06	0,96
6E	124	161	1128	281	1694	0,17	**18,54**	2,83	3,18	0,07	0,10	0,77
6F												
6G	92	111	1076	307	1586	0,13	**13,74**	2,79	2,97	0,06	0,07	0,83
6H	86	62	478	115	741	0,20	**22,51**	3,19	2,69	0,12	0,08	1,39
7A	92	86	1057	292	1527	0,12	**12,43**	3,04	2,98	0,06	0,06	1,07
7B												
7C	157	157	1134	337	1785	0,18	**19,49**	2,61	2,61	0,09	0,09	1,00
7D	127	160	1106	286	1679	0,17	**18,87**	2,76	3,07	0,08	0,10	0,79
7E	91	89	647	177	1004	0,18	**19,91**	2,77	2,75	0,09	0,09	1,02
7F	118	178	1197	307	1800	0,16	**18,08**	2,71	3,24	0,07	0,10	0,66
7G	128	130	1025	258	1541	0,17	**18,44**	2,97	2,99	0,08	0,08	0,98
7H	114	149	1003	256	1522	0,17	**19,10**	2,76	3,11	0,07	0,10	0,77
8A	113	119	958	259	1449	0,16	**17,55**	2,83	2,90	0,08	0,08	0,95
8B												
8C	95	138	1004	281	1518	0,15	**16,75**	2,62	3,04	0,06	0,09	0,69
8D	158	125	1045	246	1574	0,18	**19,97**	3,24	2,90	0,10	0,08	1,26
8E	135	140	992	208	1475	0,19	**20,81**	3,24	3,30	0,09	0,09	0,96
8F	126	110	1015	238	1489	0,16	**17,36**	3,28	3,09	0,08	0,07	1,15
8G	117	112	1058	293	1580	0,14	**15,73**	2,90	2,85	0,07	0,07	1,04
8H	153	145	1075	270	1643	0,18	**20,17**	2,96	2,88	0,09	0,09	1,06
9A	148	151	1079	250	1628	0,18	**20,46**	3,06	3,09	0,09	0,09	0,98
9B	139	148	1119	248	1654	0,17	**19,19**	3,18	3,27	0,08	0,09	0,94
9C	131	136	1009	279	1555	0,17	**18,97**	2,75	2,79	0,08	0,09	0,96
9D	211	212	790	152	1365	0,31	**38,34**	2,75	2,76	0,15	0,16	1,00
9E	215	130	1205	298	1848	0,19	**20,84**	3,32	2,60	0,12	0,07	1,65
9F	115	127	812	168	1222	0,20	**22,29**	3,14	3,32	0,09	0,10	0,91
9G	160	153	902	243	1458	0,21	**24,46**	2,68	2,62	0,11	0,10	1,05
9H	118	103	1072	267	1560	0,14	**15,34**	3,22	3,05	0,08	0,07	1,15
10A	96	109	960	242	1407	0,15	**15,82**	3,01	3,16	0,07	0,08	0,88
10B	121	119	1008	241	1489	0,16	**17,68**	3,14	3,11	0,08	0,08	1,02
10C	139	94	803	203	1239	0,19	**21,01**	3,17	2,62	0,11	0,08	1,48
10D	107	117	910	257	1391	0,16	**17,66**	2,72	2,82	0,08	0,08	0,91
10E	137	128	1073	251	1589	0,17	**18,36**	3,19	3,10	0,09	0,08	1,07

10F	97	112	1038	308	1555	0,13	**14,49**	2,70	2,84	0,06	0,07	0,87
10G	112	128	984	266	1490	0,16	**17,67**	2,78	2,94	0,08	0,09	0,88
10H	162	136	1065	283	1646	0,18	**20,13**	2,93	2,70	0,10	0,08	1,19
11A	161	213	1149	291	1814	0,21	**23,34**	2,60	3,01	0,09	0,12	0,76
11B	114	97	1071	334	1616	0,13	**14,04**	2,75	2,61	0,07	0,06	1,18
11C	172	123	1083	249	1627	0,18	**20,16**	3,37	2,86	0,11	0,08	1,40
11D	224	132	1330	276	1962	0,18	**20,18**	3,81	2,92	0,11	0,07	1,70
11E	321	43	1177	104	1645	0,22	**25,34**	10,19	2,87	0,20	0,03	7,47
11F	128	123	1127	308	1686	0,15	**16,20**	2,91	2,87	0,08	0,07	1,04
11G												
11H	186	96	1048	242	1572	0,18	**19,92**	3,65	2,67	0,12	0,06	1,94
12A	145	139	1073	262	1619	0,18	**19,43**	3,04	2,98	0,09	0,09	1,04
12B	101	113	1111	291	1616	0,13	**14,26**	3,00	3,12	0,06	0,07	0,89
12C	134	128	1060	308	1630	0,16	**17,63**	2,74	2,69	0,08	0,08	1,05
12D	122	157	767	185	1231	0,23	**26,06**	2,60	3,01	0,10	0,13	0,78
12E												
12F	151	148	1056	245	1600	0,19	**20,86**	3,07	3,04	0,09	0,09	1,02
12G												
12H	96	120	1023	288	1527	0,14	**15,32**	2,74	2,98	0,06	0,08	0,80
13F	151	123	1152	280	1706	0,16	**17,61**	3,23	2,96	0,09	0,07	1,23
13G	137	130	1083	292	1642	0,16	**17,85**	2,89	2,83	0,08	0,08	1,05
13H	128	120	982	288	1518	0,16	**17,95**	2,72	2,65	0,08	0,08	1,07
14B	120	112	1099	261	1592	0,15	**15,83**	3,27	3,18	0,08	0,07	1,07
1	127	125	763	206	1221	0,21	**23,37**	2,69	2,67	0,10	0,10	1,02
2	159	160	1172	306	1797	0,18	**19,69**	2,86	2,86	0,09	0,09	0,99
3	171	166	1151	334	1822	0,18	**20,62**	2,64	2,61	0,09	0,09	1,03
4	144	136	874	179	1333	0,21	**23,85**	3,23	3,13	0,11	0,10	1,06
5	133	120	849	213	1315	0,19	**21,56**	2,95	2,80	0,10	0,09	1,11

Supplemental table 6. Recombination frequency measurements for Neo-6 x 420 3.1.3

Individual	Green	Red	Both	None	Total	None/Total	RF(%)	G/non G	R/non R	G/T	R/T	G/R
1A	132	104	983	257	1476	0,16	**17,52**	3,09	2,79	0,09	0,07	1,27
1B												
1C	156	157	1038	285	1636	0,19	**21,43**	2,70	2,71	0,10	0,10	0,99
1D	132	81	949	253	1415	0,15	**16,40**	3,24	2,68	0,09	0,06	1,63
1E	94	322	879	385	1680	0,25	**28,95**	1,38	2,51	0,06	0,19	0,29
1F	120	101	1061	271	1553	0,14	**15,42**	3,17	2,97	0,08	0,07	1,19
1G	130	119	940	215	1404	0,18	**19,67**	3,20	3,07	0,09	0,08	1,09
1H	143	130	976	244	1493	0,18	**20,36**	2,99	2,86	0,10	0,09	1,10
2A	351	409	799	67	1626	0,47	**74,47**	2,42	2,89	0,22	0,25	0,86
2B	154	126	983	216	1479	0,19	**21,17**	3,32	3,00	0,10	0,09	1,22
2C	174	147	992	225	1538	0,21	**23,67**	3,13	2,85	0,11	0,10	1,18
2D	126	119	1063	322	1630	0,15	**16,37**	2,70	2,64	0,08	0,07	1,06
2E	134	132	1004	225	1495	0,18	**19,74**	3,19	3,16	0,09	0,09	1,02

2F	149	125	1147	285	1706	0,16	**17,61**	3,16	2,93	0,09	0,07	1,19
2G	165	143	1120	278	1706	0,18	**20,07**	3,05	2,85	0,10	0,08	1,15
2H	54	55	404	106	619	0,18	**19,51**	2,84	2,87	0,09	0,09	0,98
3A												
3B	427	401	799	16	1643	0,50	**#NUM!**	2,94	2,71	0,26	0,24	1,06
3C	149	132	1042	265	1588	0,18	**19,62**	3,00	2,84	0,09	0,08	1,13
3D												
3E	126	123	1052	261	1562	0,16	**17,47**	3,07	3,04	0,08	0,08	1,02
3F	131	144	1086	296	1657	0,17	**18,26**	2,77	2,88	0,08	0,09	0,91
3G	160	126	1027	242	1555	0,18	**20,49**	3,23	2,87	0,10	0,08	1,27
3H	146	125	1045	288	1604	0,17	**18,63**	2,88	2,70	0,09	0,08	1,17
4A	124	111	986	251	1472	0,16	**17,50**	3,07	2,93	0,08	0,08	1,12
4B	132	133	990	279	1534	0,17	**19,10**	2,72	2,73	0,09	0,09	0,99
4C	138	162	1054	287	1641	0,18	**20,35**	2,65	2,86	0,08	0,10	0,85
4D	153	157	1014	250	1574	0,20	**22,15**	2,87	2,91	0,10	0,10	0,97
4E	162	126	998	240	1526	0,19	**21,10**	3,17	2,80	0,11	0,08	1,29
4F	143	107	1005	272	1527	0,16	**17,99**	3,03	2,68	0,09	0,07	1,34
4G	140	91	1019	279	1529	0,15	**16,46**	3,13	2,65	0,09	0,06	1,54
4H	134	116	1013	260	1523	0,16	**18,04**	3,05	2,87	0,09	0,08	1,16
5A	120	120	1082	254	1576	0,15	**16,61**	3,21	3,21	0,08	0,08	1,00
5B												
5C	138	103	1087	270	1598	0,15	**16,43**	3,28	2,92	0,09	0,06	1,34
5D	193	160	1220	322	1895	0,19	**20,79**	2,93	2,68	0,10	0,08	1,21
5E	247	193	1310	268	2018	0,22	**24,91**	3,38	2,92	0,12	0,10	1,28
5F	141	147	1033	297	1618	0,18	**19,75**	2,64	2,69	0,09	0,09	0,96
5G	126	124	1101	300	1651	0,15	**16,50**	2,89	2,88	0,08	0,08	1,02
5H												
6A	121	136	1106	271	1634	0,16	**17,21**	3,01	3,17	0,07	0,08	0,89
6B	120	115	996	243	1474	0,16	**17,47**	3,12	3,06	0,08	0,08	1,04
6C	150	124	1029	286	1589	0,17	**19,06**	2,88	2,64	0,09	0,08	1,21
6D	113	127	1075	287	1602	0,15	**16,31**	2,87	3,01	0,07	0,08	0,89
6E												
6F	315	14	988	73	1390	0,24	**27,43**	14,98	2,58	0,23	0,01	22,50
6G	131	136	1087	235	1589	0,17	**18,52**	3,28	3,34	0,08	0,09	0,96
6H	183	124	1009	227	1543	0,20	**22,41**	3,40	2,76	0,12	0,08	1,48
7A	139	139	981	225	1484	0,19	**20,92**	3,08	3,08	0,09	0,09	1,00
7B	129	141	939	243	1452	0,19	**20,75**	2,78	2,90	0,09	0,10	0,91
7C	103	123	1137	304	1667	0,14	**14,63**	2,90	3,10	0,06	0,07	0,84
7D	267	19	902	36	1224	0,23	**27,02**	21,25	3,04	0,22	0,02	14,05
7E	176	129	1034	243	1582	0,19	**21,62**	3,25	2,78	0,11	0,08	1,36
7F	90	108	1135	290	1623	0,12	**13,05**	3,08	3,27	0,06	0,07	0,83
7G	121	121	1102	244	1588	0,15	**16,62**	3,35	3,35	0,08	0,08	1,00
7H	99	107	980	292	1478	0,14	**15,07**	2,70	2,78	0,07	0,07	0,93
8A	115	109	1017	272	1513	0,15	**16,10**	2,97	2,91	0,08	0,07	1,06

8B	235	60	1127	169	1591	0,19	**20,68**	5,95	2,94	0,15	0,04	3,92
8C	218	185	1133	288	1824	0,22	**25,29**	2,86	2,60	0,12	0,10	1,18
8D	132	133	974	276	1515	0,17	**19,37**	2,70	2,71	0,09	0,09	0,99
8E												
8F	97	119	886	249	1351	0,16	**17,52**	2,67	2,90	0,07	0,09	0,82
8G	145	115	993	277	1530	0,17	**18,75**	2,90	2,63	0,09	0,08	1,26
8H	128	115	811	241	1295	0,19	**20,96**	2,64	2,51	0,10	0,09	1,11
9A	127	96	1071	262	1556	0,14	**15,54**	3,35	3,00	0,08	0,06	1,32
9B	114	106	1031	282	1533	0,14	**15,56**	2,95	2,87	0,07	0,07	1,08
9C	127	112	1046	295	1580	0,15	**16,49**	2,88	2,74	0,08	0,07	1,13
9D	111	114	1044	312	1581	0,14	**15,42**	2,71	2,74	0,07	0,07	0,97
9E	41	72	548	146	807	0,14	**15,15**	2,70	3,32	0,05	0,09	0,57
9F	284	41	1061	141	1527	0,21	**24,22**	7,39	2,59	0,19	0,03	6,93
9G	128	103	1001	269	1501	0,15	**16,80**	3,03	2,78	0,09	0,07	1,24
9H	193	142	1093	249	1677	0,20	**22,51**	3,29	2,79	0,12	0,08	1,36
10A	226	67	1223	202	1718	0,17	**18,83**	5,39	3,01	0,13	0,04	3,37
10B	174	191	1157	229	1751	0,21	**23,64**	3,17	3,34	0,10	0,11	0,91
10C	130	116	1027	284	1557	0,16	**17,30**	2,89	2,76	0,08	0,07	1,12
10D	135	122	1116	261	1634	0,16	**17,21**	3,27	3,13	0,08	0,07	1,11
10E	139	133	986	260	1518	0,18	**19,90**	2,86	2,80	0,09	0,09	1,05
10F	105	134	1031	296	1566	0,15	**16,65**	2,64	2,91	0,07	0,09	0,78
10G	125	135	1032	258	1550	0,17	**18,48**	2,94	3,05	0,08	0,09	0,93
10H	136	167	1089	270	1662	0,18	**20,29**	2,80	3,09	0,08	0,10	0,81
11A	143	141	1084	321	1689	0,17	**18,53**	2,66	2,64	0,08	0,08	1,01
11B	132	122	1111	266	1631	0,16	**17,02**	3,20	3,10	0,08	0,07	1,08
11C	153	127	1093	259	1632	0,17	**18,95**	3,23	2,96	0,09	0,08	1,20
11D	365	318	737	14	1434	0,48	**21,78**	3,32	2,78	0,25	0,22	1,15
11E	30	27	259	71	387	0,15	**16,01**	2,95	2,83	0,08	0,07	1,11
11F												
11G	0	98	610	1	709	0,14	**14,94**	6,16	708,00	0,00	0,14	0,00
11H	154	136	1041	246	1577	0,18	**20,49**	3,13	2,94	0,10	0,09	1,13
12A	113	101	1103	288	1605	0,13	**14,37**	3,13	3,00	0,07	0,06	1,12
12B	141	97	840	193	1271	0,19	**20,91**	3,38	2,81	0,11	0,08	1,45
12C	136	130	919	241	1426	0,19	**20,82**	2,84	2,78	0,10	0,09	1,05
12D	127	117	1038	300	1582	0,15	**16,84**	2,79	2,70	0,08	0,07	1,09
12E	185	161	911	198	1455	0,24	**27,58**	3,05	2,80	0,13	0,11	1,15
12F	142	103	1106	265	1616	0,15	**16,53**	3,39	2,97	0,09	0,06	1,38
12G	128	114	1053	275	1570	0,15	**16,83**	3,04	2,90	0,08	0,07	1,12
12H	120	112	1066	284	1582	0,15	**15,93**	2,99	2,92	0,08	0,07	1,07
13A	127	124	926	270	1447	0,17	**19,19**	2,67	2,64	0,09	0,09	1,02
13B	103	109	1088	281	1581	0,13	**14,45**	3,05	3,12	0,07	0,07	0,94
13C	132	117	1126	276	1651	0,15	**16,43**	3,20	3,05	0,08	0,07	1,13
13D	100	115	1122	274	1611	0,13	**14,38**	3,14	3,31	0,06	0,07	0,87

Supplemental table 7. Recombination frequency measurements for Oy-0 x 420 2.3.2

Individual	Green	Red	Both	None	Total	None/Total	RF(%)	G/non G	R/non R	G/T	R/T	G/R
1A	121	136	1069	301	1627	0,16	**17,29**	2,72	2,86	0,07	0,08	0,89
1B	106	124	1024	245	1499	0,15	**16,75**	3,06	3,27	0,07	0,08	0,85
1C	136	147	1009	228	1520	0,19	**20,78**	3,05	3,18	0,09	0,10	0,93
1D	136	134	1017	251	1538	0,18	**19,45**	2,99	2,97	0,09	0,09	1,01
1E	109	147	1076	258	1590	0,16	**17,66**	2,93	3,33	0,07	0,09	0,74
1F	130	140	1072	283	1625	0,17	**18,29**	2,84	2,93	0,08	0,09	0,93
1G	128	174	993	254	1549	0,19	**21,89**	2,62	3,05	0,08	0,11	0,74
1H												
2A												
2B	143	134	1093	255	1625	0,17	**18,82**	3,18	3,08	0,09	0,08	1,07
2C	148	141	1108	280	1677	0,17	**19,05**	2,98	2,92	0,09	0,08	1,05
2D	390	436	853	13	1692	0,49	**84,62**	2,77	3,20	0,23	0,26	0,89
2E	136	185	1049	231	1601	0,20	**22,60**	2,85	3,36	0,08	0,12	0,74
2F	136	161	996	261	1554	0,19	**21,40**	2,68	2,91	0,09	0,10	0,84
2G	129	156	1139	263	1687	0,17	**18,63**	3,03	3,30	0,08	0,09	0,83
2H	164	196	1027	211	1598	0,23	**25,88**	2,93	3,26	0,10	0,12	0,84
3A	128	150	1036	281	1595	0,17	**19,29**	2,70	2,90	0,08	0,09	0,85
3B	130	165	1004	255	1554	0,19	**21,24**	2,70	3,04	0,08	0,11	0,79
3C	172	171	978	209	1530	0,22	**25,73**	3,03	3,02	0,11	0,11	1,01
3D	150	178	1037	249	1614	0,20	**22,96**	2,78	3,05	0,09	0,11	0,84
3E	154	133	1024	228	1539	0,19	**20,81**	3,26	3,03	0,10	0,09	1,16
3F	147	170	1045	235	1597	0,20	**22,35**	2,94	3,18	0,09	0,11	0,86
3G	127	142	1023	248	1540	0,17	**19,34**	2,95	3,11	0,08	0,09	0,89
3H												
4A	160	141	1036	247	1584	0,19	**21,26**	3,08	2,89	0,10	0,09	1,13
4B	101	155	984	262	1502	0,17	**18,81**	2,60	3,14	0,07	0,10	0,65
4C	109	126	1032	285	1552	0,15	**16,50**	2,78	2,94	0,07	0,08	0,87
4D	122	128	1037	238	1525	0,16	**18,02**	3,17	3,24	0,08	0,08	0,95
4E	143	166	1065	224	1598	0,19	**21,69**	3,10	3,35	0,09	0,10	0,86
4F	118	163	979	252	1512	0,19	**20,73**	2,64	3,09	0,08	0,11	0,72
4G	372	417	752	9	1550	0,51	**#NUM!**	2,64	3,07	0,24	0,27	0,89
4H	107	164	1036	248	1555	0,17	**19,29**	2,77	3,38	0,07	0,11	0,65
5A	131	140	1075	253	1599	0,17	**18,70**	3,07	3,16	0,08	0,09	0,94
5B	134	143	970	256	1503	0,18	**20,54**	2,77	2,85	0,09	0,10	0,94
5C	134	147	971	254	1506	0,19	**20,83**	2,76	2,88	0,09	0,10	0,91
5D	152	173	980	217	1522	0,21	**24,31**	2,90	3,12	0,10	0,11	0,88
5E	111	120	1049	266	1546	0,15	**16,26**	3,01	3,10	0,07	0,08	0,93
5F	157	186	1007	208	1558	0,22	**25,19**	2,95	3,27	0,10	0,12	0,84
5G	128	166	999	230	1523	0,19	**21,65**	2,85	3,25	0,08	0,11	0,77
5H												
6A	107	122	973	261	1463	0,16	**17,12**	2,82	2,98	0,07	0,08	0,88
6B	117	121	973	266	1477	0,16	**17,68**	2,82	2,86	0,08	0,08	0,97
6C	152	171	1016	243	1582	0,20	**23,08**	2,82	3,01	0,10	0,11	0,89

6D	135	173	982	238	1528	0,20	**22,74**	2,72	3,10	0,09	0,11	0,78
6E	135	131	992	231	1489	0,18	**19,83**	3,11	3,07	0,09	0,09	1,03
6F	147	180	966	214	1507	0,22	**24,77**	2,82	3,17	0,10	0,12	0,82
6G	222	101	588	398	1309	0,25	**28,83**	1,62	1,11	0,17	0,08	2,20
6H	162	175	968	210	1515	0,22	**25,49**	2,94	3,07	0,11	0,12	0,93
7A	167	134	951	218	1470	0,20	**23,16**	3,18	2,82	0,11	0,09	1,25
7B												
7C	146	148	992	227	1513	0,19	**21,81**	3,03	3,06	0,10	0,10	0,99
7D	136	131	982	227	1476	0,18	**20,11**	3,12	3,07	0,09	0,09	1,04
7E	166	165	1055	238	1624	0,20	**23,03**	3,03	3,02	0,10	0,10	1,01
7F	165	197	1151	243	1756	0,21	**23,34**	2,99	3,30	0,09	0,11	0,84
7G	153	137	978	215	1483	0,20	**21,97**	3,21	3,03	0,10	0,09	1,12
7H	136	122	993	235	1486	0,17	**19,21**	3,16	3,01	0,09	0,08	1,11
8A												
8B	123	141	1015	243	1522	0,17	**19,19**	2,96	3,16	0,08	0,09	0,87
8C	153	119	989	231	1492	0,18	**20,29**	3,26	2,89	0,10	0,08	1,29
8D	133	117	1029	237	1516	0,16	**18,14**	3,28	3,10	0,09	0,08	1,14
8E	192	168	945	212	1517	0,24	**27,52**	2,99	2,75	0,13	0,11	1,14
8F	188	170	927	189	1474	0,24	**28,29**	3,11	2,91	0,13	0,12	1,11
8G	161	166	999	203	1529	0,21	**24,35**	3,14	3,20	0,11	0,11	0,97
8H	178	165	1127	318	1788	0,19	**21,49**	2,70	2,60	0,10	0,09	1,08
9A	184	211	1191	231	1817	0,22	**24,82**	3,11	3,38	0,10	0,12	0,87
9B	155	152	963	205	1475	0,21	**23,60**	3,13	3,10	0,11	0,10	1,02
9C	141	164	1034	240	1579	0,19	**21,66**	2,91	3,14	0,09	0,10	0,86
9D	132	139	1077	239	1587	0,17	**18,85**	3,20	3,28	0,08	0,09	0,95
9E												
9F	124	143	1025	262	1554	0,17	**18,98**	2,84	3,03	0,08	0,09	0,87
9G	152	125	984	240	1501	0,18	**20,57**	3,11	2,83	0,10	0,08	1,22
9H	137	145	1033	223	1538	0,18	**20,42**	3,18	3,27	0,09	0,09	0,94
10A	151	172	1025	227	1575	0,21	**23,20**	2,95	3,17	0,10	0,11	0,88
10B	126	128	976	238	1468	0,17	**19,13**	3,01	3,03	0,09	0,09	0,98
10C	181	188	969	212	1550	0,24	**27,62**	2,88	2,94	0,12	0,12	0,96
10D	156	154	1031	216	1557	0,20	**22,42**	3,21	3,19	0,10	0,10	1,01
10E	118	134	973	241	1466	0,17	**18,99**	2,91	3,08	0,08	0,09	0,88
10F	126	133	1010	223	1492	0,17	**19,20**	3,19	3,28	0,08	0,09	0,95
10G	119	129	1092	252	1592	0,16	**17,03**	3,18	3,29	0,07	0,08	0,92
10H	137	136	1004	213	1490	0,18	**20,40**	3,27	3,26	0,09	0,09	1,01
11A	129	110	1048	251	1538	0,16	**16,98**	3,26	3,05	0,08	0,07	1,17
11B	147	185	975	198	1505	0,22	**25,25**	2,93	3,36	0,10	0,12	0,79
11C	170	169	930	217	1486	0,23	**26,26**	2,85	2,84	0,11	0,11	1,01
11D												
11E	156	172	1008	209	1545	0,21	**24,14**	3,06	3,23	0,10	0,11	0,91
11F	172	213	1274	306	1965	0,20	**22,02**	2,79	3,11	0,09	0,11	0,81
11G												

11H												
12A	183	201	1228	244	1856	0,21	**23,44**	3,17	3,35	0,10	0,11	0,91
12B	168	131	1022	227	1548	0,19	**21,66**	3,32	2,92	0,11	0,08	1,28
12C	191	166	1049	203	1609	0,22	**25,42**	3,36	3,08	0,12	0,10	1,15
12D	191	187	1226	269	1873	0,20	**22,78**	3,11	3,07	0,10	0,10	1,02
12E												
12F	140	131	1094	252	1617	0,17	**18,46**	3,22	3,13	0,09	0,08	1,07
12G	129	169	966	243	1507	0,20	**22,25**	2,66	3,05	0,09	0,11	0,76
12H	149	149	1023	279	1600	0,19	**20,79**	2,74	2,74	0,09	0,09	1,00
15A	74	64	936	87	1161	0,12	**12,69**	6,69	6,21	0,06	0,06	1,16
15B	177	173	1087	232	1669	0,21	**23,80**	3,12	3,08	0,11	0,10	1,02
15C	168	185	1324	295	1972	0,18	**19,88**	3,11	3,26	0,09	0,09	0,91
15D	158	130	1248	291	1827	0,16	**17,25**	3,34	3,07	0,09	0,07	1,22
15E	177	207	1351	329	2064	0,19	**20,76**	2,85	3,08	0,09	0,10	0,86
15F	156	193	1346	320	2015	0,17	**19,15**	2,93	3,23	0,08	0,10	0,81
15G	200	245	1202	278	1925	0,23	**26,67**	2,68	3,03	0,10	0,13	0,82
15H	152	208	1316	332	2008	0,18	**19,91**	2,72	3,15	0,08	0,10	0,73

Supplemental table 8. Recombination frequency measurements for Oy-0 x 420 3.1.3

Individual	Green	Red	Both	None	Total	None/Total	RF(%)	G/non G	R/non R	G/T	R/T	G/R
1A	168	183	1010	252	1613	0,22	**24,85**	2,71	2,84	0,10	0,11	0,92
1B	148	166	1037	261	1612	0,19	**21,87**	2,78	2,94	0,09	0,10	0,89
1C	159	191	920	222	1492	0,23	**27,14**	2,61	2,92	0,11	0,13	0,83
1D	124	160	1011	261	1556	0,18	**20,32**	2,70	3,04	0,08	0,10	0,78
1E												
1F	118	118	1028	229	1493	0,16	**17,30**	3,30	3,30	0,08	0,08	1,00
1G	154	120	977	230	1481	0,19	**20,63**	3,23	2,86	0,10	0,08	1,28
1H	117	110	1000	257	1484	0,15	**16,69**	3,04	2,97	0,08	0,07	1,06
2A	146	162	982	250	1540	0,20	**22,54**	2,74	2,89	0,09	0,11	0,90
2B	160	175	1025	239	1599	0,21	**23,78**	2,86	3,01	0,10	0,11	0,91
2C	113	136	1038	243	1530	0,16	**17,87**	3,04	3,30	0,07	0,09	0,83
2D	113	136	1001	246	1496	0,17	**18,32**	2,92	3,17	0,08	0,09	0,83
2E	187	199	1057	195	1638	0,24	**27,29**	3,16	3,29	0,11	0,12	0,94
2F	161	179	986	222	1548	0,22	**25,12**	2,86	3,04	0,10	0,12	0,90
2G	163	169	1018	212	1562	0,21	**24,18**	3,10	3,17	0,10	0,11	0,96
2H	95	119	1083	264	1561	0,14	**14,81**	3,08	3,35	0,06	0,08	0,80
3A												
3B	143	179	1044	221	1587	0,20	**22,92**	2,97	3,36	0,09	0,11	0,80
3C	93	120	1051	279	1543	0,14	**14,92**	2,87	3,15	0,06	0,08	0,78
3D	111	154	1031	250	1546	0,17	**18,93**	2,83	3,28	0,07	0,10	0,72
3E	137	166	1026	256	1585	0,19	**21,41**	2,76	3,03	0,09	0,10	0,83
3F												
3G	91	105	1029	290	1515	0,13	**13,90**	2,84	2,98	0,06	0,07	0,87
3H	126	179	1003	224	1532	0,20	**22,42**	2,80	3,38	0,08	0,12	0,70

4A	143	169	1067	225	1604	0,19	**21,84**	3,07	3,36	0,09	0,11	0,85
4B												
4C	147	148	1014	244	1553	0,19	**21,25**	2,96	2,97	0,09	0,10	0,99
4D	121	121	1012	228	1482	0,16	**17,94**	3,25	3,25	0,08	0,08	1,00
4E	151	157	1022	200	1530	0,20	**22,71**	3,29	3,36	0,10	0,10	0,96
4F	147	148	975	208	1478	0,20	**22,49**	3,15	3,16	0,10	0,10	0,99
4G	147	134	937	208	1426	0,20	**22,16**	3,17	3,02	0,10	0,09	1,10
4H	170	156	906	190	1422	0,23	**26,41**	3,11	2,95	0,12	0,11	1,09
5A	143	139	922	254	1458	0,19	**21,69**	2,71	2,67	0,10	0,10	1,03
5B	121	147	994	232	1494	0,18	**19,92**	2,94	3,23	0,08	0,10	0,82
5C	123	138	1098	248	1607	0,16	**17,83**	3,16	3,33	0,08	0,09	0,89
5D	198	166	967	218	1549	0,23	**27,20**	3,03	2,72	0,13	0,11	1,19
5E	129	120	1047	245	1541	0,16	**17,73**	3,22	3,12	0,08	0,08	1,08
5F	178	159	968	219	1524	0,22	**25,32**	3,03	2,84	0,12	0,10	1,12
5G	158	225	1106	257	1746	0,22	**25,08**	2,62	3,21	0,09	0,13	0,70
5H	145	129	1022	230	1526	0,18	**19,94**	3,25	3,07	0,10	0,08	1,12
6A	126	119	1055	272	1572	0,16	**17,04**	3,02	2,95	0,08	0,08	1,06
6B	166	153	1008	229	1556	0,21	**23,19**	3,07	2,94	0,11	0,10	1,08
6C	130	141	1051	255	1577	0,17	**18,99**	2,98	3,10	0,08	0,09	0,92
6D	153	153	1053	244	1603	0,19	**21,37**	3,04	3,04	0,10	0,10	1,00
6E	139	158	1013	244	1554	0,19	**21,40**	2,87	3,06	0,09	0,10	0,88
6F	160	167	1282	261	1870	0,17	**19,36**	3,37	3,44	0,09	0,09	0,96
6G	127	126	962	245	1460	0,17	**19,17**	2,94	2,92	0,09	0,09	1,01
6H	136	131	984	237	1488	0,18	**19,93**	3,04	2,99	0,09	0,09	1,04
7A	209	179	1079	209	1676	0,23	**26,72**	3,32	3,01	0,12	0,11	1,17
7B	153	208	1234	272	1867	0,19	**21,69**	2,89	3,39	0,08	0,11	0,74
7C	120	138	1040	252	1550	0,17	**18,32**	2,97	3,17	0,08	0,09	0,87
7D	197	171	1201	275	1844	0,20	**22,48**	3,13	2,91	0,11	0,09	1,15
7E	124	124	1016	216	1480	0,17	**18,46**	3,35	3,35	0,08	0,08	1,00
7F	156	167	947	224	1494	0,22	**24,66**	2,82	2,93	0,10	0,11	0,93
7G	183	187	1008	195	1573	0,24	**27,23**	3,12	3,16	0,12	0,12	0,98
7H	122	171	1296	297	1886	0,16	**16,98**	3,03	3,50	0,06	0,09	0,71
8A												
8B	207	176	1329	334	2046	0,19	**20,90**	3,01	2,78	0,10	0,09	1,18
8C	182	170	1118	237	1707	0,21	**23,35**	3,19	3,07	0,11	0,10	1,07
8D	190	187	1293	269	1939	0,19	**21,82**	3,25	3,22	0,10	0,10	1,02
8E	148	128	1066	232	1574	0,18	**19,42**	3,37	3,14	0,09	0,08	1,16
8F												
8G	111	125	1095	259	1590	0,15	**16,15**	3,14	3,30	0,07	0,08	0,89
8H	153	149	1321	290	1913	0,16	**17,28**	3,36	3,32	0,08	0,08	1,03
9A	181	157	1068	266	1672	0,20	**22,82**	2,95	2,74	0,11	0,09	1,15
9B	187	191	1391	294	2063	0,18	**20,40**	3,25	3,29	0,09	0,09	0,98
9C	155	111	1134	287	1687	0,16	**17,26**	3,24	2,82	0,09	0,07	1,40
9D	133	135	1076	265	1609	0,17	**18,34**	3,02	3,04	0,08	0,08	0,99

9E												
9F	209	206	961	212	1588	0,26	**30,91**	2,80	2,77	0,13	0,13	1,01
9G	155	138	999	269	1561	0,19	**20,97**	2,84	2,68	0,10	0,09	1,12
9H	133	131	1003	264	1531	0,17	**19,06**	2,88	2,86	0,09	0,09	1,02
10A	161	150	979	228	1518	0,20	**23,17**	3,02	2,90	0,11	0,10	1,07
10B	143	136	1045	271	1595	0,17	**19,37**	2,92	2,85	0,09	0,09	1,05
10C												
10D	145	140	1007	245	1537	0,19	**20,68**	2,99	2,94	0,09	0,09	1,04
10E												
10F	139	166	1008	248	1561	0,20	**21,95**	2,77	3,03	0,09	0,11	0,84
10G	152	146	1026	224	1548	0,19	**21,58**	3,18	3,12	0,10	0,09	1,04
10H												
11A	137	144	1048	242	1571	0,18	**19,86**	3,07	3,15	0,09	0,09	0,95
11B	144	109	1040	240	1533	0,17	**18,15**	3,39	2,99	0,09	0,07	1,32
11C	144	126	993	264	1527	0,18	**19,60**	2,92	2,74	0,09	0,08	1,14
11D	142	132	1052	261	1587	0,17	**19,09**	3,04	2,94	0,09	0,08	1,08
11E	174	178	986	208	1546	0,23	**26,20**	3,01	3,05	0,11	0,12	0,98
11F	165	212	1189	261	1827	0,21	**23,36**	2,86	3,29	0,09	0,12	0,78
11G	139	148	1039	218	1544	0,19	**20,74**	3,22	3,32	0,09	0,10	0,94
11H	172	143	1045	222	1582	0,20	**22,43**	3,33	3,02	0,11	0,09	1,20
12A	153	170	1046	211	1580	0,20	**23,11**	3,15	3,34	0,10	0,11	0,90
12B	250	227	1130	245	1852	0,26	**30,37**	2,92	2,74	0,13	0,12	1,10
12C	133	115	963	251	1462	0,17	**18,71**	2,99	2,81	0,09	0,08	1,16
12D												
12E	195	191	1031	203	1620	0,24	**27,65**	3,11	3,07	0,12	0,12	1,02
12F	170	138	1022	230	1560	0,20	**22,21**	3,24	2,90	0,11	0,09	1,23
12G	130	143	1050	228	1551	0,18	**19,50**	3,18	3,33	0,08	0,09	0,91
12H	138	136	1006	274	1554	0,18	**19,54**	2,79	2,77	0,09	0,09	1,01
14A												
14B												
14C												
14D												
14E	123	147	1014	256	1540	0,18	**19,42**	2,82	3,06	0,08	0,10	0,84
14F	137	147	1048	223	1555	0,18	**20,33**	3,20	3,32	0,09	0,09	0,93
14G	165	159	1036	211	1571	0,21	**23,35**	3,25	3,18	0,11	0,10	1,04
14H												

Supplemental table 9. Recombination frequency measurements for Per-1 x 420 1.3.3

Individual	Green	Red	Both	None	Total	None/Total	RF(%)	G/non G	R/non R	G/T	R/T	G/R
1A	122	170	1027	267	1586	0,18	**20,52**	2,63	3,08	0,08	0,11	0,72
1B	143	169	986	256	1554	0,20	**22,64**	2,66	2,89	0,09	0,11	0,85
1C	154	171	1183	325	1833	0,18	**19,66**	2,70	2,83	0,08	0,09	0,90
1D	104	113	1204	318	1739	0,12	**13,37**	3,03	3,12	0,06	0,06	0,92
1E												

1F	116	137	1246	345	1844	0,14	**14,82**	2,83	3,00	0,06	0,07	0,85
1G	111	139	1169	325	1744	0,14	**15,54**	2,76	3,00	0,06	0,08	0,80
1H	120	164	1067	283	1634	0,17	**19,23**	2,66	3,05	0,07	0,10	0,73
2A	116	145	1185	329	1775	0,15	**15,98**	2,74	2,99	0,07	0,08	0,80
2B	154	185	1193	297	1829	0,19	**20,67**	2,79	3,06	0,08	0,10	0,83
2C	137	154	1068	278	1637	0,18	**19,72**	2,79	2,94	0,08	0,09	0,89
2D	148	158	1265	294	1865	0,16	**18,03**	3,13	3,22	0,08	0,08	0,94
2E	115	143	1107	286	1651	0,16	**17,09**	2,85	3,12	0,07	0,09	0,80
2F	158	155	1156	278	1747	0,18	**19,90**	3,03	3,01	0,09	0,09	1,02
2G	147	143	1063	286	1639	0,18	**19,62**	2,82	2,79	0,09	0,09	1,03
2H	136	136	1040	301	1613	0,17	**18,59**	2,69	2,69	0,08	0,08	1,00
3A	132	142	999	267	1540	0,18	**19,74**	2,77	2,86	0,09	0,09	0,93
3B	151	192	1066	275	1684	0,20	**23,02**	2,61	2,95	0,09	0,11	0,79
3C												
3D	95	101	1159	312	1667	0,12	**12,54**	3,04	3,10	0,06	0,06	0,94
3E	117	120	1063	265	1565	0,15	**16,51**	3,06	3,10	0,07	0,08	0,98
3F	107	117	1179	361	1764	0,13	**13,63**	2,69	2,77	0,06	0,07	0,91
3G	131	147	1127	302	1707	0,16	**17,89**	2,80	2,94	0,08	0,09	0,89
3H	115	182	1106	287	1690	0,18	**19,47**	2,60	3,20	0,07	0,11	0,63
4A	94	129	1121	324	1668	0,13	**14,41**	2,68	2,99	0,06	0,08	0,73
4B	167	145	1002	254	1568	0,20	**22,41**	2,93	2,72	0,11	0,09	1,15
4C	103	124	1013	254	1494	0,15	**16,57**	2,95	3,18	0,07	0,08	0,83
4D	121	129	1173	366	1789	0,14	**15,12**	2,61	2,67	0,07	0,07	0,94
4E	129	134	1131	310	1704	0,15	**16,85**	2,84	2,88	0,08	0,08	0,96
4F	127	151	1090	286	1654	0,17	**18,52**	2,78	3,00	0,08	0,09	0,84
4G												
4H	48	53	441	124	666	0,15	**16,53**	2,76	2,87	0,07	0,08	0,91
5A	152	156	1097	268	1673	0,18	**20,51**	2,95	2,98	0,09	0,09	0,97
5B	119	152	1044	255	1570	0,17	**19,08**	2,86	3,20	0,08	0,10	0,78
5C	73	97	623	151	944	0,18	**20,01**	2,81	3,21	0,08	0,10	0,75
5D	122	128	1075	311	1636	0,15	**16,67**	2,73	2,78	0,07	0,08	0,95
5E	112	152	1094	300	1658	0,16	**17,44**	2,67	3,02	0,07	0,09	0,74
5F												
5G	125	131	1168	291	1715	0,15	**16,25**	3,06	3,12	0,07	0,08	0,95
5H	153	160	1259	315	1887	0,17	**18,25**	2,97	3,03	0,08	0,08	0,96
6A	112	124	1082	335	1653	0,14	**15,47**	2,60	2,70	0,07	0,08	0,90
6B	137	124	1146	322	1729	0,15	**16,45**	2,88	2,77	0,08	0,07	1,10
6C	161	137	1143	311	1752	0,17	**18,77**	2,91	2,71	0,09	0,08	1,18
6D	106	113	1059	281	1559	0,14	**15,20**	2,96	3,03	0,07	0,07	0,94
6E	76	82	778	237	1173	0,13	**14,52**	2,68	2,75	0,06	0,07	0,93
6F	147	135	1110	321	1713	0,16	**18,10**	2,76	2,66	0,09	0,08	1,09
6G	171	167	950	239	1527	0,22	**25,35**	2,76	2,72	0,11	0,11	1,02
6H	155	163	1064	286	1668	0,19	**21,34**	2,71	2,78	0,09	0,10	0,95
7A												

7B	127	129	1073	291	1620	0,16	**17,30**	2,86	2,88	0,08	0,08	0,98
7C	129	149	1025	259	1562	0,18	**19,75**	2,83	3,03	0,08	0,10	0,87
7D	129	138	968	283	1518	0,18	**19,49**	2,61	2,68	0,08	0,09	0,93
7E	132	132	1133	314	1711	0,15	**16,85**	2,84	2,84	0,08	0,08	1,00
7F	108	112	1098	303	1621	0,14	**14,64**	2,91	2,94	0,07	0,07	0,96
7G	131	135	1076	310	1652	0,16	**17,66**	2,71	2,75	0,08	0,08	0,97
7H	102	129	1087	317	1635	0,14	**15,30**	2,67	2,90	0,06	0,08	0,79
8A	123	155	1116	306	1700	0,16	**17,97**	2,69	2,96	0,07	0,09	0,79
8B	177	164	1038	231	1610	0,21	**24,08**	3,08	2,95	0,11	0,10	1,08
8C	141	113	1153	328	1735	0,15	**15,90**	2,93	2,70	0,08	0,07	1,25
8D	116	145	1118	315	1694	0,15	**16,82**	2,68	2,93	0,07	0,09	0,80
8E	108	100	1148	270	1626	0,13	**13,74**	3,39	3,30	0,07	0,06	1,08
8F	159	132	1198	287	1776	0,16	**18,01**	3,24	2,98	0,09	0,07	1,20
8G	109	121	1188	314	1732	0,13	**14,30**	2,98	3,09	0,06	0,07	0,90
8H	139	130	1169	300	1738	0,15	**16,91**	3,04	2,96	0,08	0,07	1,07
9A	154	129	1199	324	1806	0,16	**17,14**	2,99	2,78	0,09	0,07	1,19
9B												
9C												
9D	94	78	914	252	1338	0,13	**13,81**	3,05	2,87	0,07	0,06	1,21
9E	67	62	679	181	989	0,13	**14,03**	3,07	2,99	0,07	0,06	1,08
9F	81	84	630	143	938	0,18	**19,49**	3,13	3,19	0,09	0,09	0,96
9G	130	133	1176	309	1748	0,15	**16,39**	2,95	2,98	0,07	0,08	0,98
9H												
10A	165	136	1168	263	1732	0,17	**19,23**	3,34	3,05	0,10	0,08	1,21
10B	148	145	1138	245	1676	0,17	**19,36**	3,30	3,26	0,09	0,09	1,02
10C	185	148	1353	341	2027	0,16	**18,06**	3,15	2,85	0,09	0,07	1,25
10D	139	162	1155	268	1724	0,17	**19,33**	3,01	3,24	0,08	0,09	0,86
10E	172	165	1168	273	1778	0,19	**21,20**	3,06	3,00	0,10	0,09	1,04
10F	107	89	1129	306	1631	0,12	**12,84**	3,13	2,95	0,07	0,05	1,20
10G	98	115	1157	323	1693	0,13	**13,49**	2,87	3,02	0,06	0,07	0,85
10H												
11A	126	130	1194	293	1743	0,15	**15,96**	3,12	3,16	0,07	0,07	0,97
11B	109	93	1128	291	1621	0,12	**13,35**	3,22	3,05	0,07	0,06	1,17
11C	171	143	1261	335	1910	0,16	**18,07**	3,00	2,77	0,09	0,07	1,20
11D	122	111	1116	261	1610	0,14	**15,71**	3,33	3,20	0,08	0,07	1,10
11E	105	109	1237	307	1758	0,12	**13,02**	3,23	3,27	0,06	0,06	0,96
11F	142	127	1156	273	1698	0,16	**17,35**	3,25	3,09	0,08	0,07	1,12
11G	144	156	1006	232	1538	0,20	**21,90**	2,96	3,09	0,09	0,10	0,92
11H	187	198	1333	293	2011	0,19	**21,44**	3,10	3,19	0,09	0,10	0,94
12A	151	155	1159	260	1725	0,18	**19,67**	3,16	3,20	0,09	0,09	0,97
12B	125	118	1020	281	1544	0,16	**17,22**	2,87	2,80	0,08	0,08	1,06
12C	156	118	1226	301	1801	0,15	**16,59**	3,30	2,94	0,09	0,07	1,32
12D												
12E	148	209	1332	316	2005	0,18	**19,76**	2,82	3,32	0,07	0,10	0,71

12F	122	128	1125	283	1658	0,15	**16,43**	3,03	3,09	0,07	0,08	0,95
12G												
12H	134	136	1248	315	1833	0,15	**16,01**	3,06	3,08	0,07	0,07	0,99
17C	142	116	1119	267	1644	0,16	**17,17**	3,29	3,02	0,09	0,07	1,22
17D	127	119	1142	298	1686	0,15	**15,85**	3,04	2,97	0,08	0,07	1,07
17E	160	141	1095	274	1670	0,18	**20,03**	3,02	2,85	0,10	0,08	1,13
17F	158	210	1360	368	2096	0,18	**19,45**	2,63	2,98	0,08	0,10	0,75
17G	119	108	1013	278	1518	0,15	**16,28**	2,93	2,82	0,08	0,07	1,10
17H	144	166	1374	323	2007	0,15	**16,87**	3,10	3,30	0,07	0,08	0,87
1	119	103	936	245	1403	0,16	**17,32**	3,03	2,85	0,08	0,07	1,16
2	122	122	971	203	1418	0,17	**19,02**	3,36	3,36	0,09	0,09	1,00
3	146	109	980	247	1482	0,17	**19,01**	3,16	2,77	0,10	0,07	1,34
4	149	138	963	210	1460	0,20	**22,10**	3,20	3,07	0,10	0,09	1,08
5	87	141	993	253	1474	0,15	**16,90**	2,74	3,34	0,06	0,10	0,62
6	112	106	904	254	1376	0,16	**17,35**	2,82	2,76	0,08	0,08	1,06
7	84	64	708	201	1057	0,14	**15,15**	2,99	2,71	0,08	0,06	1,31
8	110	141	989	237	1477	0,17	**18,75**	2,91	3,26	0,07	0,10	0,78

Supplemental table 10. Recombination frequency measurements for Per-1 x 420 3.2.3

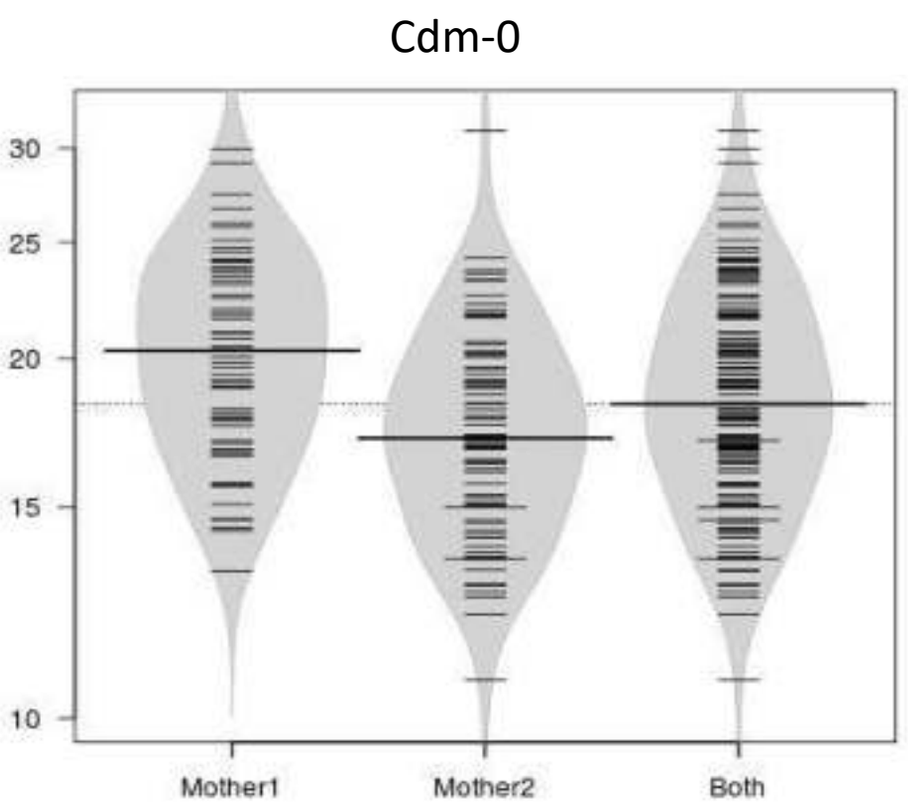
Individual	Green	Red	Both	None	Total	None/Total	RF(%)	G/non G	R/non R	G/T	R/T	G/R
1A	133	138	1192	331	1794	0,15	**16,46**	2,83	2,87	0,07	0,08	0,96
1B												
1C	158	138	1150	333	1779	0,17	**18,32**	2,78	2,62	0,09	0,08	1,14
1D	132	94	1010	272	1508	0,15	**16,32**	3,12	2,73	0,09	0,06	1,40
1E												
1F	138	111	1114	324	1687	0,15	**16,05**	2,88	2,65	0,08	0,07	1,24
1G	111	111	948	282	1452	0,15	**16,68**	2,69	2,69	0,08	0,08	1,00
1H	157	168	1068	220	1613	0,20	**22,73**	3,16	3,28	0,10	0,10	0,93
2A	155	148	1166	300	1769	0,17	**18,92**	2,95	2,89	0,09	0,08	1,05
2B	125	119	1057	275	1576	0,15	**16,91**	3,00	2,94	0,08	0,08	1,05
2C	143	150	1148	280	1721	0,17	**18,79**	3,00	3,07	0,08	0,09	0,95
2D	126	144	1095	266	1631	0,17	**18,21**	2,98	3,16	0,08	0,09	0,88
2E	179	103	1237	334	1853	0,15	**16,60**	3,24	2,61	0,10	0,06	1,74
2F												
2G	140	102	1174	292	1708	0,14	**15,35**	3,34	2,95	0,08	0,06	1,37
2H	132	149	1180	332	1793	0,16	**17,14**	2,73	2,86	0,07	0,08	0,89
3A	129	116	1176	324	1745	0,14	**15,19**	2,97	2,85	0,07	0,07	1,11
3B	128	110	1176	310	1724	0,14	**14,92**	3,10	2,94	0,07	0,06	1,16
3C	159	117	1250	350	1876	0,15	**15,99**	3,02	2,69	0,08	0,06	1,36
3D	138	129	1092	300	1659	0,16	**17,65**	2,87	2,79	0,08	0,08	1,07
3E	162	128	1094	268	1652	0,18	**19,45**	3,17	2,84	0,10	0,08	1,27
3F	106	112	1041	290	1549	0,14	**15,23**	2,85	2,91	0,07	0,07	0,95
3G	112	111	1047	298	1568	0,14	**15,41**	2,83	2,82	0,07	0,07	1,01
3H	123	137	1181	301	1742	0,15	**16,24**	2,98	3,11	0,07	0,08	0,90

4A	82	84	837	243	1246	0,13	**14,35**	2,81	2,83	0,07	0,07	0,98
4B	150	150	1124	296	1720	0,17	**19,31**	2,86	2,86	0,09	0,09	1,00
4C	137	131	1107	286	1661	0,16	**17,70**	2,98	2,93	0,08	0,08	1,05
4D	156	178	1108	226	1668	0,20	**22,57**	3,13	3,37	0,09	0,11	0,88
4E												
4F	166	145	1121	297	1729	0,18	**19,98**	2,91	2,73	0,10	0,08	1,14
4G	149	120	1072	275	1616	0,17	**18,33**	3,09	2,81	0,09	0,07	1,24
4H	34	54	504	143	735	0,12	**12,79**	2,73	3,15	0,05	0,07	0,63
5A	145	151	987	249	1532	0,19	**21,67**	2,83	2,89	0,09	0,10	0,96
5B	140	133	1133	307	1713	0,16	**17,46**	2,89	2,83	0,08	0,08	1,05
5C	157	135	1134	306	1732	0,17	**18,59**	2,93	2,74	0,09	0,08	1,16
5D	127	116	1100	297	1640	0,15	**16,12**	2,97	2,87	0,08	0,07	1,09
5E												
5F	141	162	1261	289	1853	0,16	**17,97**	3,11	3,31	0,08	0,09	0,87
5G	132	99	1158	285	1674	0,14	**14,91**	3,36	3,01	0,08	0,06	1,33
5H												
6A	122	93	1186	296	1697	0,13	**13,59**	3,36	3,06	0,07	0,05	1,31
6B	140	146	992	262	1540	0,19	**20,72**	2,77	2,83	0,09	0,09	0,96
6C	103	116	1073	303	1595	0,14	**14,83**	2,81	2,93	0,06	0,07	0,89
6D	130	88	1151	320	1689	0,13	**13,87**	3,14	2,75	0,08	0,05	1,48
6E	119	137	1217	319	1792	0,14	**15,48**	2,93	3,09	0,07	0,08	0,87
6F	139	128	1172	325	1764	0,15	**16,50**	2,89	2,80	0,08	0,07	1,09
6G	140	126	1112	317	1695	0,16	**17,17**	2,83	2,71	0,08	0,07	1,11
6H	138	119	1229	317	1803	0,14	**15,45**	3,14	2,96	0,08	0,07	1,16
7A												
7B												
7C	134	127	1141	302	1704	0,15	**16,71**	2,97	2,91	0,08	0,07	1,06
7D	152	100	1167	298	1717	0,15	**15,95**	3,31	2,82	0,09	0,06	1,52
7E	122	123	1115	303	1663	0,15	**16,01**	2,90	2,91	0,07	0,07	0,99
7F	133	119	1232	293	1777	0,14	**15,36**	3,31	3,17	0,07	0,07	1,12
7G	141	110	1102	274	1627	0,15	**16,85**	3,24	2,92	0,09	0,07	1,28
7H												
8A												
8B	191	176	1172	238	1777	0,21	**23,39**	3,29	3,14	0,11	0,10	1,09
8C	141	96	1160	321	1718	0,14	**14,91**	3,12	2,72	0,08	0,06	1,47
8D	111	122	1246	349	1828	0,13	**13,68**	2,88	2,97	0,06	0,07	0,91
8E	165	164	1111	259	1699	0,19	**21,72**	3,02	3,01	0,10	0,10	1,01
8F												
8G	130	101	1238	313	1782	0,13	**13,93**	3,30	3,02	0,07	0,06	1,29
8H	147	120	1194	285	1746	0,15	**16,68**	3,31	3,04	0,08	0,07	1,23
9A	117	110	1154	304	1685	0,13	**14,53**	3,07	3,00	0,07	0,07	1,06
9B	154	148	1191	311	1804	0,17	**18,44**	2,93	2,88	0,09	0,08	1,04
9C	133	113	1172	320	1738	0,14	**15,33**	3,01	2,84	0,08	0,07	1,18
9D												

9E	150	144	1260	389	1943	0,15	**16,49**	2,65	2,60	0,08	0,07	1,04
9F	135	126	1216	293	1770	0,15	**16,03**	3,22	3,14	0,08	0,07	1,07
9G	155	146	1206	301	1808	0,17	**18,33**	3,04	2,96	0,09	0,08	1,06
9H												
10A	127	142	1136	328	1733	0,16	**16,96**	2,69	2,81	0,07	0,08	0,89
10B	141	131	1107	302	1681	0,16	**17,76**	2,88	2,79	0,08	0,08	1,08
10C	132	150	1082	292	1656	0,17	**18,80**	2,75	2,91	0,08	0,09	0,88
10D	156	133	1121	284	1694	0,17	**18,83**	3,06	2,85	0,09	0,08	1,17
10E	144	99	1075	275	1593	0,15	**16,64**	3,26	2,80	0,09	0,06	1,45
10F	119	120	1125	270	1634	0,15	**15,89**	3,19	3,20	0,07	0,07	0,99
10G	127	119	1068	324	1638	0,15	**16,36**	2,70	2,63	0,08	0,07	1,07
10H	131	134	1219	311	1795	0,15	**16,05**	3,03	3,06	0,07	0,07	0,98
11A	120	110	1031	286	1547	0,15	**16,18**	2,91	2,81	0,08	0,07	1,09
11B												
11C												
11D	121	118	1063	294	1596	0,15	**16,30**	2,87	2,85	0,08	0,07	1,03
11E												
11F	129	157	1176	307	1769	0,16	**17,74**	2,81	3,06	0,07	0,09	0,82
11G	104	101	1010	297	1512	0,14	**14,63**	2,80	2,77	0,07	0,07	1,03
11H	142	131	1152	281	1706	0,16	**17,54**	3,14	3,03	0,08	0,08	1,08
12A	109	118	1062	284	1573	0,14	**15,66**	2,91	3,00	0,07	0,08	0,92
12B	177	151	1160	250	1738	0,19	**21,10**	3,33	3,07	0,10	0,09	1,17
12C												
12D	122	100	1232	323	1777	0,12	**13,39**	3,20	2,99	0,07	0,06	1,22
12E	166	131	1223	308	1828	0,16	**17,84**	3,16	2,86	0,09	0,07	1,27
12F	128	171	1201	278	1778	0,17	**18,53**	2,96	3,38	0,07	0,10	0,75
12G												
12H	162	154	1318	374	2008	0,16	**17,22**	2,80	2,75	0,08	0,08	1,05
16A	145	131	1264	332	1872	0,15	**16,03**	3,04	2,92	0,08	0,07	1,11
16B												
16C												
16D												
16E	160	146	1094	281	1681	0,18	**20,25**	2,94	2,81	0,10	0,09	1,10
16F	65	126	844	223	1258	0,15	**16,55**	2,60	3,37	0,05	0,10	0,52
16G												
16H	119	123	1075	280	1597	0,15	**16,52**	2,96	3,00	0,07	0,08	0,97

6.2 Plotted recombination frequency measurements in 420 for the five F2 segregating populations

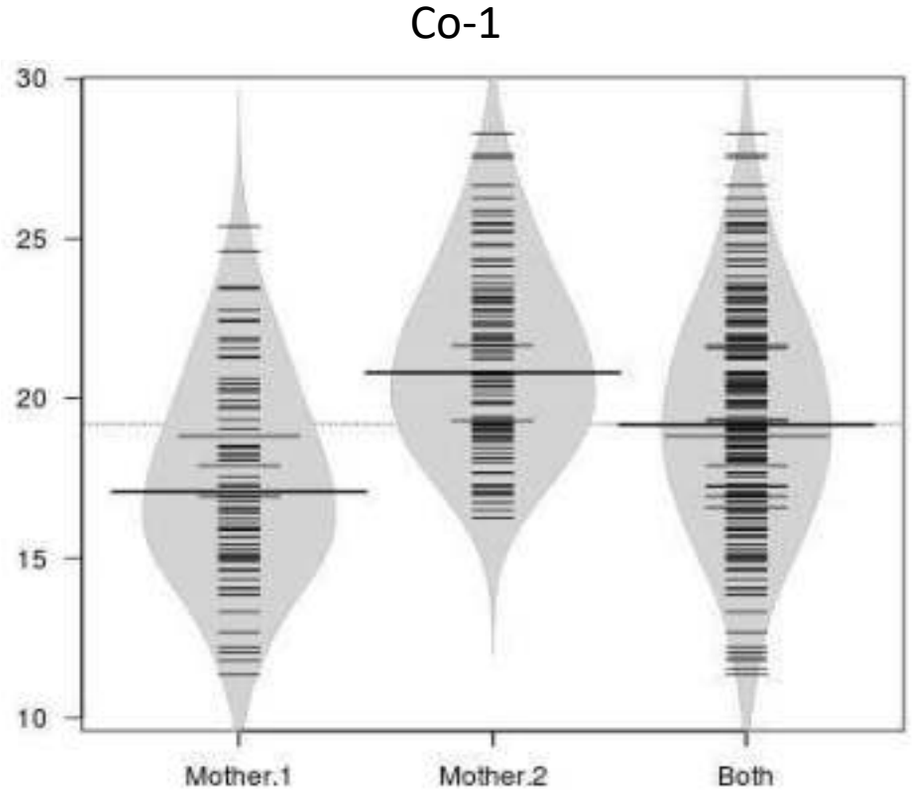
Each F2 segregating population originates from two F1 mothers. The data from supplemental tables 1 to 10 were plotted in the following figures. Short statistical analysis of the data is associated to each plot.



Box plot statistics

	Mother1	Mother2	Both
Upper whisker	29,94	24,28	29,94
3rd quartile	23,6	19,63	21,73
Median	20,3	17,15	18,31
1st quartile	17,57	15	15,73
Lower whisker	13,28	10,76	10,76
Nr. of data points	61	80	141

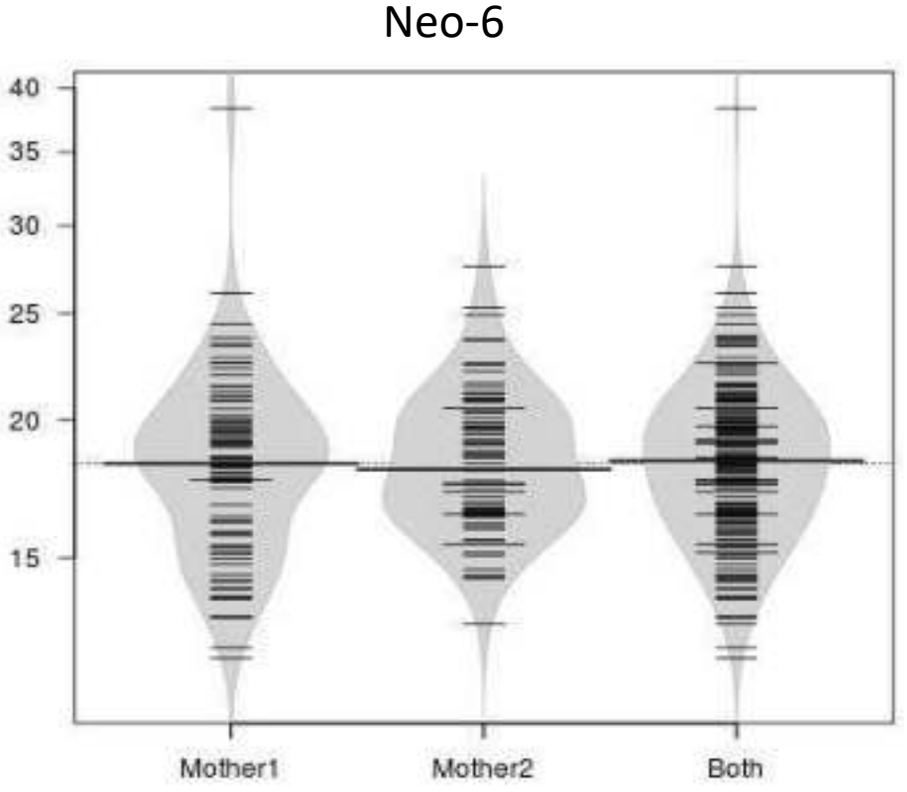
Supplemental figure 1. Violin plot of the Cdm-0 x 420 F2 population with each mother separately and the whole population. Data from Supplemental table 1 Supplemental table 2.



Box plot statistics

	Mother 1	Mother 2	Both
Upper whisker	25,39	28,29	28,29
3rd quartile	19,51	23,08	21,74
Median	16,94	20,81	19,17
1st quartile	15,14	19,13	16,91
Lower whisker	11,36	16,26	11,36
Nr. of data points	87	89	176

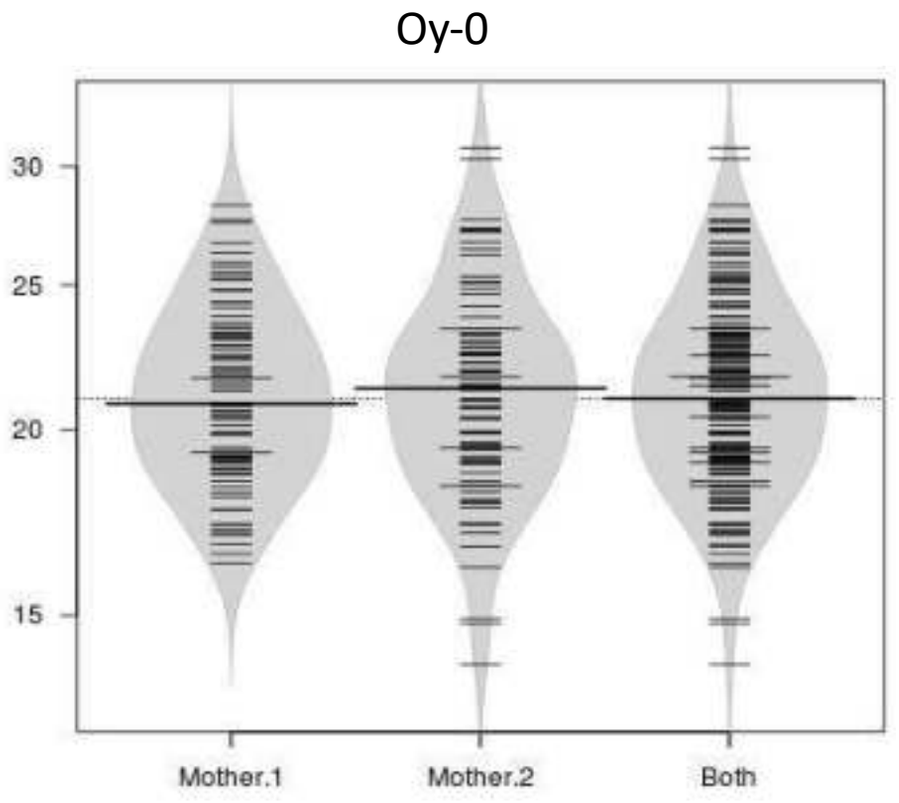
Supplemental figure 2. Violin plot of the Co-1 x 420 F2 population with each mother separately and the whole population. Data from Supplemental table 3Supplemental table 4.



Box plot statistics

	Mother1	Mother2	Both
Upper whisker	26,06	25,29	26,06
3rd quartile	20,13	20,35	20,36
Median	18,25	18,04	18,36
1st quartile	15,82	16,49	16,38
Lower whisker	12,16	13,05	12,16
Nr. of data points	85	81	171

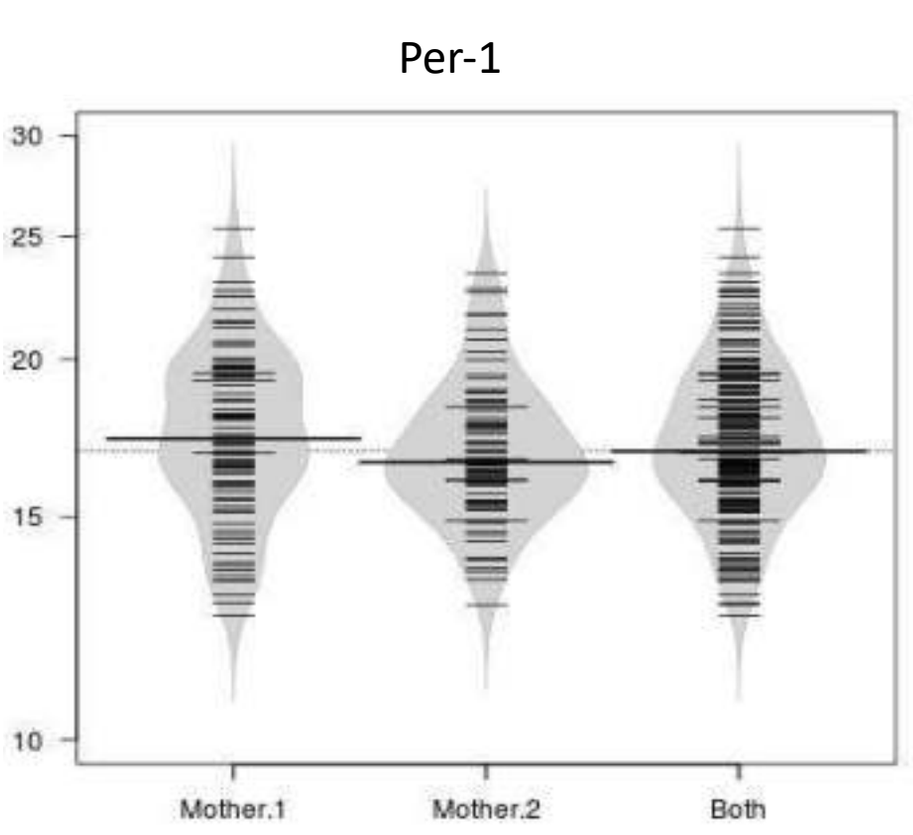
Supplemental figure 3. Violin plot of the Neo-6 x 420 F2 population with each mother separately and the whole population. Data from Supplemental table 5Supplemental table 6.



Box plot statistics

	Mother 1	Mother 2	Both
Upper whisker	28,29	27,65	28,29
3rd quartile	23,08	23,17	23,13
Median	20,81	21,31	20,97
1st quartile	19,13	19,06	19,07
Lower whisker	16,26	13,9	13,9
Nr. of data points	89	86	175

Supplemental figure 4. Violin plot of the Oy-0 x 420 F2 population with each mother separately and the whole population. Data from Supplemental table 7Supplemental table 8.



Box plot statistics

	Mother 1	Mother 2	Both
Upper whisker	24,08	22,57	23,39
3rd quartile	19,49	18,32	18,92
Median	17,3	16,58	16,91
1st quartile	15,93	15,48	15,89
Lower whisker	12,54	12,79	12,54
Nr. of data points	91	82	181

Supplemental figure 5. Violin plot of the Per-1 x 420 F2 population with each mother separately and the whole population. Data from Supplemental table 9Supplemental table 10.

6.3 Variance calculation of the recombination frequency of the five F2 segregation populations

Variance of a sample population and whiskers difference were calculated for the five populations according to the following equations:

Equation 3: **Variance = (x'-x)2/(n-1)**

Equation 4: **Whiskers difference = upper whisker - lower whisker**

These values are used as qualitative estimations of potential existence of quantitative trait loci. Interestingly, the two populations with the highest variance and whiskers differences showed potential QTLs, whereas the three others did not.

F2 population	Measurements	Variance	Whiskers difference
CDM-0	141	15,43	19.18 cM
Co-1	176	12,49	16.93 cM
Neo-6	171	10,6	13.9 cM
Oy-0	175	9,54	14.39 cM
Per-1	173	5,93	10.85 cM

6.4 Library indexes

Supplemental table 11. List of the indexes used for generating DNA libraries

oligo.name	sequence
N7-01	CAAGCAGAAGACGGCATACGAGATTCGCCTTAGTCTCGTGGGCTCGG
N7-02	CAAGCAGAAGACGGCATACGAGATCTAGTACGGTCTCGTGGGCTCGG
N7-03	CAAGCAGAAGACGGCATACGAGATTTCTGCCTGTCTCGTGGGCTCGG
N7-04	CAAGCAGAAGACGGCATACGAGATGCTCAGGAGTCTCGTGGGCTCGG
N7-05	CAAGCAGAAGACGGCATACGAGATAGGAGTCCGTCTCGTGGGCTCGG
N7-06	CAAGCAGAAGACGGCATACGAGATCATGCCTAGTCTCGTGGGCTCGG
N7-07	CAAGCAGAAGACGGCATACGAGATGTAGAGAGGTCTCGTGGGCTCGG
N7-08	CAAGCAGAAGACGGCATACGAGATCCTCTCTGGTCTCGTGGGCTCGG
N7-09	CAAGCAGAAGACGGCATACGAGATAGCGTAGCGTCTCGTGGGCTCGG
N7-10	CAAGCAGAAGACGGCATACGAGATCAGCCTCGGTCTCGTGGGCTCGG
N7-11	CAAGCAGAAGACGGCATACGAGATTGCCTCTTGTCTCGTGGGCTCGG
N7-12	CAAGCAGAAGACGGCATACGAGATTCCTCTACGTCTCGTGGGCTCGG
N7-14	CAAGCAGAAGACGGCATACGAGATTCATGAGCGTCTCGTGGGCTCGG
N7-15	CAAGCAGAAGACGGCATACGAGATCCTGAGATGTCTCGTGGGCTCGG
N7-16	CAAGCAGAAGACGGCATACGAGATTAGCGAGTGTCTCGTGGGCTCGG
N7-18	CAAGCAGAAGACGGCATACGAGATGTAGCTCCGTCTCGTGGGCTCGG
N7-19	CAAGCAGAAGACGGCATACGAGATTACTACGCGTCTCGTGGGCTCGG
N7-20	CAAGCAGAAGACGGCATACGAGATAGGCTCCGGTCTCGTGGGCTCGG
N7-21	CAAGCAGAAGACGGCATACGAGATGCAGCGTAGTCTCGTGGGCTCGG
N7-22	CAAGCAGAAGACGGCATACGAGATCTGCGCATGTCTCGTGGGCTCGG
N7-23	CAAGCAGAAGACGGCATACGAGATGAGCGCTAGTCTCGTGGGCTCGG
N7-24	CAAGCAGAAGACGGCATACGAGATCGCTCAGTGTCTCGTGGGCTCGG
N7-26	CAAGCAGAAGACGGCATACGAGATGTCTTAGGGTCTCGTGGGCTCGG
N7-27	CAAGCAGAAGACGGCATACGAGATACTGATCGGTCTCGTGGGCTCGG
S5-01	AATGATACGGCGACCACCGAGATCTACACTAGATCGCTCGTCGGCAGCGTC
S5-02	AATGATACGGCGACCACCGAGATCTACACCTCTCTATTCGTCGGCAGCGTC
S5-03	AATGATACGGCGACCACCGAGATCTACACTATCCTCTTCGTCGGCAGCGTC
S5-04	AATGATACGGCGACCACCGAGATCTACACAGAGTAGATCGTCGGCAGCGTC
S5-05	AATGATACGGCGACCACCGAGATCTACACGTAAGGAGTCGTCGGCAGCGTC
S5-06	AATGATACGGCGACCACCGAGATCTACACACTGCATATCGTCGGCAGCGTC
S5-07	AATGATACGGCGACCACCGAGATCTACACAAGGAGTATCGTCGGCAGCGTC
S5-08	AATGATACGGCGACCACCGAGATCTACACCTAAGCCT TCGTCGGCAGCGTC
S5-09	AATGATACGGCGACCACCGAGATCTACACGGCTACTCTCGTCGGCAGCGTC
S5-10	AATGATACGGCGACCACCGAGATCTACACCCTCAGACTCGTCGGCAGCGTC
S5-11	AATGATACGGCGACCACCGAGATCTACACTCCTTACGTCGTCGGCAGCGTC
S5-12	AATGATACGGCGACCACCGAGATCTACACACGCGTGGTCGTCGGCAGCGTC

S5-13	AATGATACGGCGACCACCGAGATCTACACGGAACTCCTCGTCGGCAGCGTC
S5-14	AATGATACGGCGACCACCGAGATCTACACTGGCCATGTCGTCGGCAGCGTC
S5-15	AATGATACGGCGACCACCGAGATCTACACGAGAGATTTCGTCGGCAGCGTC
S5-16	AATGATACGGCGACCACCGAGATCTACACCGCGGTTATCGTCGGCAGCGTC
S5-17B	AATGATACGGCGACCACCGAGATCTACACGACCGCCATCGTCGGCAGCGTC
S5-18	AATGATACGGCGACCACCGAGATCTACACTAAGATGGTCGTCGGCAGCGTC
S5-19	AATGATACGGCGACCACCGAGATCTACACATTGACATTCGTCGGCAGCGTC
S5-20	AATGATACGGCGACCACCGAGATCTACACAGCCAACTTCGTCGGCAGCGTC
S5-21	AATGATACGGCGACCACCGAGATCTACACTACTAGGTTCGTCGGCAGCGTC
S5-22	AATGATACGGCGACCACCGAGATCTACACTCACGGTTTCGTCGGCAGCGTC
S5-23	AATGATACGGCGACCACCGAGATCTACACTGTAATGATCGTCGGCAGCGTC
S5-24	AATGATACGGCGACCACCGAGATCTACACCACGTCAGTCGTCGGCAGCGTC

Chapter 3:

Investigation of the effect of *MutL* genes expression levels on meiotic crossover recombination

This work was funded by NCN grants Sonata Bis 6 (2016/22/E/NZ2/00455) to PAZ. Preludium 18 (2019/35/N/NZ2/02933) to NK and FNP grant Team POIR.04.04.00-00-5C0F/17 to PAZ.

Chapter 3:

Investigation of the effect of *MutL* genes expression levels on meiotic crossover recombination

List of abbreviations

-/-	Null mutant
+/-	Heterozygous
+/+	Wild type
%	Percent
BASTA	Glufosinate ammonium
bp	Base pair
C	Celsius
Cas	CRISPR-associated protein
cDNA	Coding desoxyribonucleic acid
cm	Centimenter
cM	CentiMorgan
CO	Crossover
Col-0	Columbia-0
CRISPR	Clustered regularly interspaced short palindromic repeats
CTL	Columbia traffic line
dbl Oe	Double overexpressor
dHJ	Double Holliday junction
DMC1	Disrupted meiotic coding DNA
DNA	Desoxyribonucleic acid
dNTP	Deoxynucleoside triphosphate
DSB	Double strand break
dsRED	Discosoma red fluorescent protein
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
eGFP	Enhanced green fluorescent protein
EMS	Ethyl methanesulfonate
EXO1	Exonuclease 1
F1	Fillial generation 1
FANCM	Fanconi anemia complementation M
FTL	Fluorescent tagged line
GK	GABI-Kat
gRNA	Guide RNA
h	hour
HEI10	Homolog of human enhancer of cell invasion 10
JM	Joint molecule
LB	Lysate Broth
L*er-0*	Lansberg *erecta*-0
LTL	Lansberg erecta traffic line

Mbp	Mega base pair
min	Minute
mL	Milliliter
MLH	Mutator S homolog
MMR	Mismatch repair
MSH	Mutator L homolog
MUS81	MMS and UV sensitive 81
MutH	Mutator H
MutL	Mutator L
MutS	Mutator S
N.A.	Not applicable
N.S.	Not significant
NASC	The European Arabidopsis Stock Centre
ng	Nanogram
Oe#	Overexpressor number
P	Probability
PAM	Protospacer Adjacent Motif
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
pMLH1	MLH1 promoter
PMS1	Post meiotic segregation 1
r^2	Coefficient of determination
RF	Recombination frequency
RNA	Ribonucleic acid
RPA	Replication Protein A
rpm	Rotation per minute
SALK	Produced by Salk Institute Genomic Analysis Laboratory
sec	Second
SK	Saskatoon
ssDNA	Single stranded DNA
T-DNA	Transfer DNA
T1	Tranformant generation 1
TAE	Tris Acetate EDTA
TAIR	The Arabidopsis Information Resource
ug	Microgram
uL	Microliter
um	Micrimeter
V	Volume
W	Weight
Ws	Wassilewskija ecotype

WT	Wild type
Zip	Zing transporter precursor
ZMM	Zip1, Zip2, Zip3, Zip4, Msh4-Msh5, Mer3, and Spo16
RFC	Replication factor C
indel	Insertion/deletion

List of figures:

List of tables:

List of supplemental figures:

List of supplemental tables:

1 Introduction

Meiotic cell divisions are initiated with the replication of the genetic material. Akin to mitotic divisions, the mismatch repair (hereafter MMR) system scans the DNA for mismatches and corrects them (Iyer et al., 2006; Li, 2008; Larrea et al., 2010; Jiricny, 2013; Fishel, 2015; Han et al., 2022). The bacterial MMR system is composed of three proteins that operate in homodimers: MutS, MutL, and MutH. The plant MMR system presents multiple homologs for MutS, and MutL (Table 1). As of today, no Eukaryote showed the existence of any homologs for the MutH factor (Culligan et al., 2000; Lin et al., 2007a; Fukui, 2010; Jiricny, 2013; Fishel, 2015; Reyes et al., 2015). Additionally to their canonical roles, MMR proteins are also involved in meiotic crossover distribution and formation. Some of them, MSH4, MSH5, and MLH3, evolved to be specifically involved in meiotic crossover formation (Aguilera and Rothstein, 2007; Hunter, 2007; Larrea et al., 2010; Hunter, 2015; Mercier et al., 2015; Lambing et al., 2017; Dluzewska et al., 2018).

Table 1. Mismatch repair genes in *E. coli* and their homologs in *S. cerevisiae* and *A. thaliana*.

E. coli	*S. cerevisiae*	*A. thaliana*
MutS	*MSH1*	*MSH1*
	MSH2	*MSH2*
	MSH3	*MSH3*
	MSH4	*MSH4*
	MSH5	*MSH5*
	MSH6	*MSH6*
		MSH7
MutH	---	---
MutL	*MLH1*	*MLH1*
	MLH2	
	MLH3	*MLH3*
	PMS1	*PMS1*

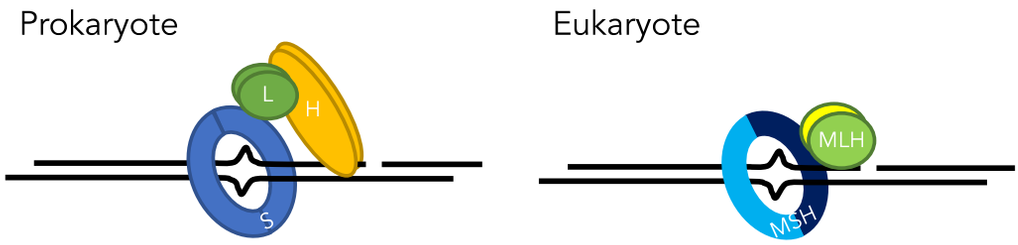


Figure 1. Simplified operating of the mismatch repair system in prokaryotes and eukaryotes. In prokaryotes, the MutS homodimer sliding clamp scans the replicated DNA. When MutS recognizes a mismatch, it recruits MutL homodimer which recruits the MutH homodimer. MutH has an endonuclease activity and nicks the neo-synthesized DNA. It is then resected and repaired. In eukaryotes, the MutS complex is a heterodimer formed from homologs of MutS (MSH). It also operates as a sliding clamp that recognizes mismatches. Eukaryotes do not have homologs for MutH. MSH heterodimers recruit MutL homologs (MLH), which also form a heterodimer. The MLH heterodimers hold the needed endonuclease activity. The nicked DNA is then processed and repaired. Figure adapted from Lin et al., 2007b.

1.1 Mismatch repair system

The mismatch repair (MMR) system is very highly conserved. It is found in the Prokaryote, Archaea, and Eukaryote branches. The overall *modus operandi* is similar (Iyer et al., 2006; Lin et al., 2007b; Larrea et al., 2010; Jiricny, 2013; Han et al., 2022). Replicated DNA is scanned for mismatches. When a mismatch is recognized, the machinery stops scanning and recruits additional molecules. The DNA is nicked and the neosynthesized strand is resected and repaired (Figure 1). The MMR activity is very important for maintaining the integrity and stability of the genetic material. The estimated efficiency of MMR can improve the fidelity of the replicated DNA up to 1000 folds (Modrich and Lahue, 1996; Umar and Kunkel, 1996; Harfe and Jinks-Robertson, 2000; Hegan et al., 2006; Lin et al., 2007b).

1.1.1 Mismatch recognition

Cell divisions are gargantuan series of steps and stages that evolved to ensure the prosperity of the organisms undertaking the endeavor. They are always initiated by the replication of the whole genetic material. Replication is a high-fidelity process that can nevertheless make errors. Substitutions, insertions, and deletions can occur for numerous reasons such as proofreading deficiency, polymerase slippage,

and environmental mutagenic or stressful conditions (Kunkel and Bebenek, 2000; Bębenek and Ziuzia-Graczyk, 2018). Additionally to the proofreading abilities of the DNA polymerases, the MMR system evolved to improve fidelity.

Table 2. MutS and MutL complexes and their functions.

Complex		Proteins	Function
MutS	**α**	MSH2/MSH6	Recognizes substitutions and small indels (less than 3 nucleotides) and recruits MutL **α**
	β	MSH2/MSH3	Recognizes longer indels and recruits MutL **α**
	γ	MSH4/MSH5	Recognizes and stabilizes double Holiday junctions and recruits MutL **γ**
	δ	MSH2/MSH7	Preferentially recognizes substitutions and is partially redundant with MutS**α** and recruits MutL **α**
MutL	**α**	MLH1/PMS1	Excises the mismatched nucleotides
	β	---	Does not exist in plants
	γ	MLH1/MLH3	Nicks double Holiday junctions in an oriented fashion to yield crossovers

The recognition of uncorrected mismatches is done by the MutS complexes. Arabidopsis has 6 homologs of the MutS protein: MSH2 – 7 (Table 1). Apart from MutSγ (MSH4/ 5), all Arabidopsis MutS complexes are formed of MSH2 in a heterodimer with one of the three remaining MSHs, MSH3, 6 or 7 (Table 2). The MSH2-dependent heterodimers form sliding clamps that scan DNA during replication (Sachadyn, 2010; Putnam, 2020; Han et al., 2022). The different complexes recognize different types of mismatches. MutSα, formed from MSH2 and MSH6, recognizes nucleotide substitutions and small indels of fewer than 3 nucleotides. MutSβ, formed from MSH2 and MSH3, recognizes larger indels (Tian et al., 2009). MutSδ, formed from MSH2 and MSH7, preferentially recognizes substitutions. It is partially redundant with MutSα, where it recognizes the same substitutions with lesser or greater efficiency (Wu, 2003; Tam et al., 2009). MutSδ is a plant-specific heterodimer as MSH7 is only found in plants (Lin et al., 2007b).

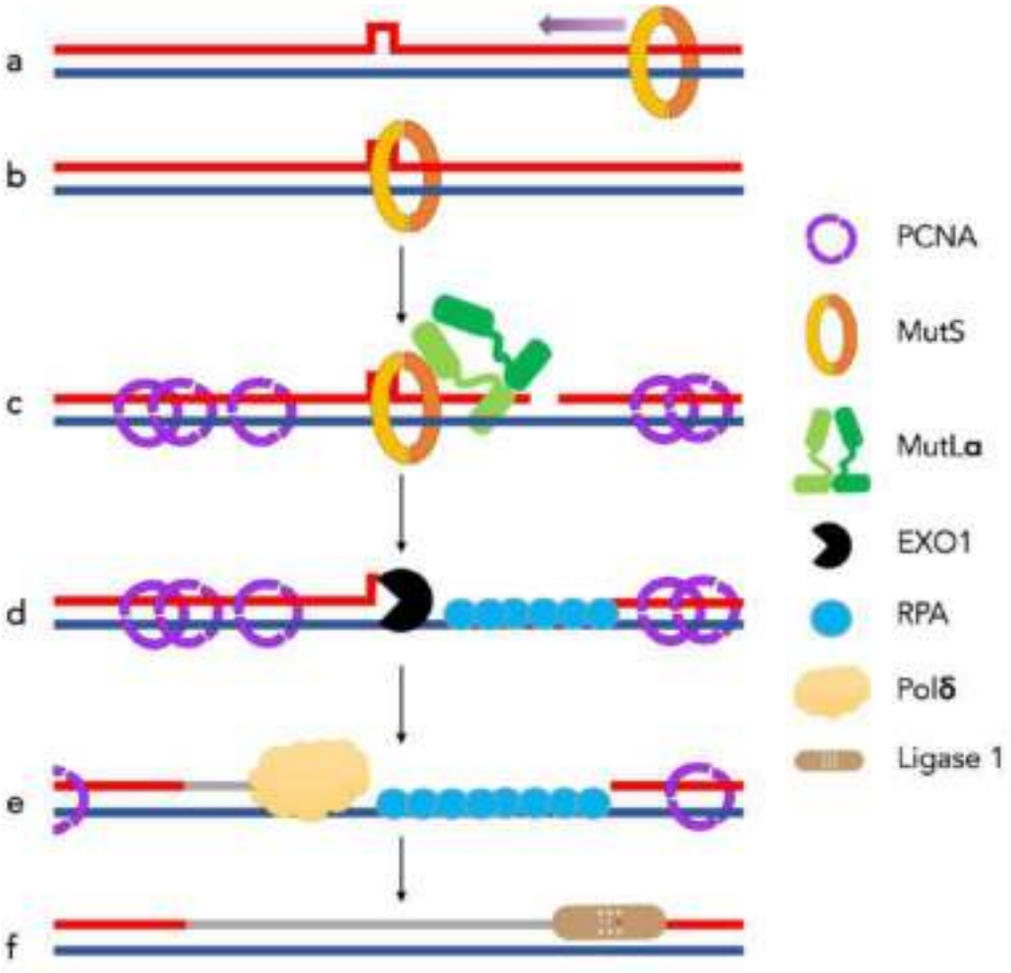


Figure 2. Model representation of mismatch repair. (a) Mismatch recognition, the MutS sliding clamp scans post-replication DNA **(b)** when it recognizes a mismatch it stops and initiates MMR. **(c)** Excision, PCNA and RFC (not represented) stabilize the stretch of DNA. A MutL complex is recruited to cut the newly synthesized strand. **(d)** EXO1 resects the faulty DNA, RPA protects the single-stranded DNA. **(e)** Error-free gap-filling DNA polymerase Polδ resynthesizes the missing portion. **(f)** Ligation, DNA ligase 1 restores the integrity of the DNA molecule. Adapted from Yang and Hsieh, 2016.

1.1.2 Mismatch correction

In Arabidopsis, when a MutS sliding clamp recognizes a mismatch, it halts and recruits MutLα heterodimer (Jiang and Marszalek, 2011; Groothuizen et al., 2015; Qiu et al., 2015; Han et al., 2022). MutLα nicks the neosynthesized DNA molecule (Figure 2). The cutting site is stabilized by PCNA and RFC proteins. The strand with the mismatch is resected by EXO1 and the single-stranded DNA is protected by RPA. Polδ resynthesizes the missing DNA which is finally ligated to the rest of the molecule by DNA Ligase 1 (Iyer et al., 2006; Li, 2008; Jiricny, 2013; Fishel, 2015).

1.2 Mismatch repair proteins in meiosis

MMR proteins evolved to also control meiotic recombination. However, MSH4, MSH5, and MLH3 are only active during meiosis and do not intervene in plant MMR.

MSH2, the core subunits of all MMR MutS complexes that are responsible for mismatch recognition, has been identified as indispensable for the juxtaposition effect. The juxtaposition effect, or heterozygosity *in-cis* effect, is the phenomenon by which when a heterozygous region is juxtaposed to a homozygous region, the heterozygous region receives more crossovers at the expense of the homozygous region (Ziolkowski et al., 2015; Blackwell et al., 2020). Indeed, in the absence of Arabidopsis MSH2, crossovers are evenly distributed along chromosome arms independently from the level of heterozygosity (Blackwell et al., 2020). In yeast models, *msh2* null mutants also display a reduction in meiotic crossover recombination and a decrease in heteroduplex rejection (Schär et al., 1997; Sugawara et al., 2004). These two phenotypes were not observed in Arabidopsis and mouse models (Blackwell et al., 2020; Peterson et al., 2020).

MSH7, which operates in a heterodimer with MSH2, was shown to be involved in limiting homeologous recombination in wheat (Serra et al., 2021). Its loss of function also negatively affects seed set in barley (Lloyd et al., 2007) and Arabidopsis (Chirinos-Arias and Spampinato, 2020). Additionally, the *msh7* null mutant seems to induce an increase in crossover meiotic recombination, as tested in Arabidopsis chromosome 3 subtelomeric interval 420 (Lario et al., 2015).

MSH4 and MSH5 form the MutSγ heterodimer and are meiosis-specific proteins. They are part of the ZMM machinery, which is responsible for the majority of the meiotic crossover events in most eukaryotes, including *A. thaliana* (Hunter, 2007; Lynn et al., 2007; Hunter, 2015; Mercier et al., 2015; Wang and Copenhaver, 2018; Ziolkowski, 2022). Furthermore, MutSγ is subjected to post-translational regulation that could be involved in crossover designation (Figure 3). Indeed, a special peptide in MSH4 active pocket, called degron, can be phosphorylated by Cdc7-Dbf4, which

protects MSH4 from degradation by the proteasome. It also increases its stability and the halftime of its presence onto DNA (He et al., 2020). Additionally, MSH4 can be sumolated by the E2 ligase Ubc9 and this process is triggered by DSB formation. MSH4 sumolation fosters its interaction with MSH5 and facilitates crossing-over (He et al., 2021). Similarly to all ZMM factors, loss of function *msh4 -/-* and *msh5 -/-* mutants display severe recombination and fertility issues (Chelysheva et al., 2007; Lynn et al., 2007; Ward et al., 2007; Lu et al., 2008; Chelysheva et al., 2012; Wang et al., 2012b; Pyatnitskaya et al., 2019).

MLH1 and MLH3 form the MutLγ heterodimer, which is another meiosis-specific heterodimer. The MutLγ endonuclease is responsible for resolving the crossover intermediates designated by the ZMM machinery (Martín et al., 2014; Hunter, 2015; Mercier et al., 2015). Recent biochemical characterization of the MutLγ endonuclease activity shows that its interaction with MutSγ, PCNA, RFC, and its add-on subunit EXO1, stabilizes the heterodimer and improves its nicking activity (Figure 3, Cannavo et al., 2020; Kulkarni et al., 2020). Moreover, the phosphorylation of EXO1 by Cdc5 is shown to also improve the MutLγ endonuclease activity (Sanchez et al., 2020). Loss of function *mlh1 -/-* and *mlh3 -/-* is detrimental to meiotic crossover recombination and plant fertility (Lipkin et al., 2002; Jackson et al., 2006; Dion et al., 2007).

1.3 Aim and biological relevance

MutL complexes are at the hearts of MMR system and class I crossover formation. Expression patterns show that *MLH1* and *PMS1* are expressed ubiquitously throughout Arabidopsis tissues. On the other hand, *MLH3* is predominantly expressed in the early stages of flower bud development. Moreover, MutLγ is believed to be the main resolvase for the ZMM-designated crossover intermediates. However, MutLγ is not considered a *bona fide* part of the ZMM machinery. It is recruited by the MutSγ, which is part of ZMM.

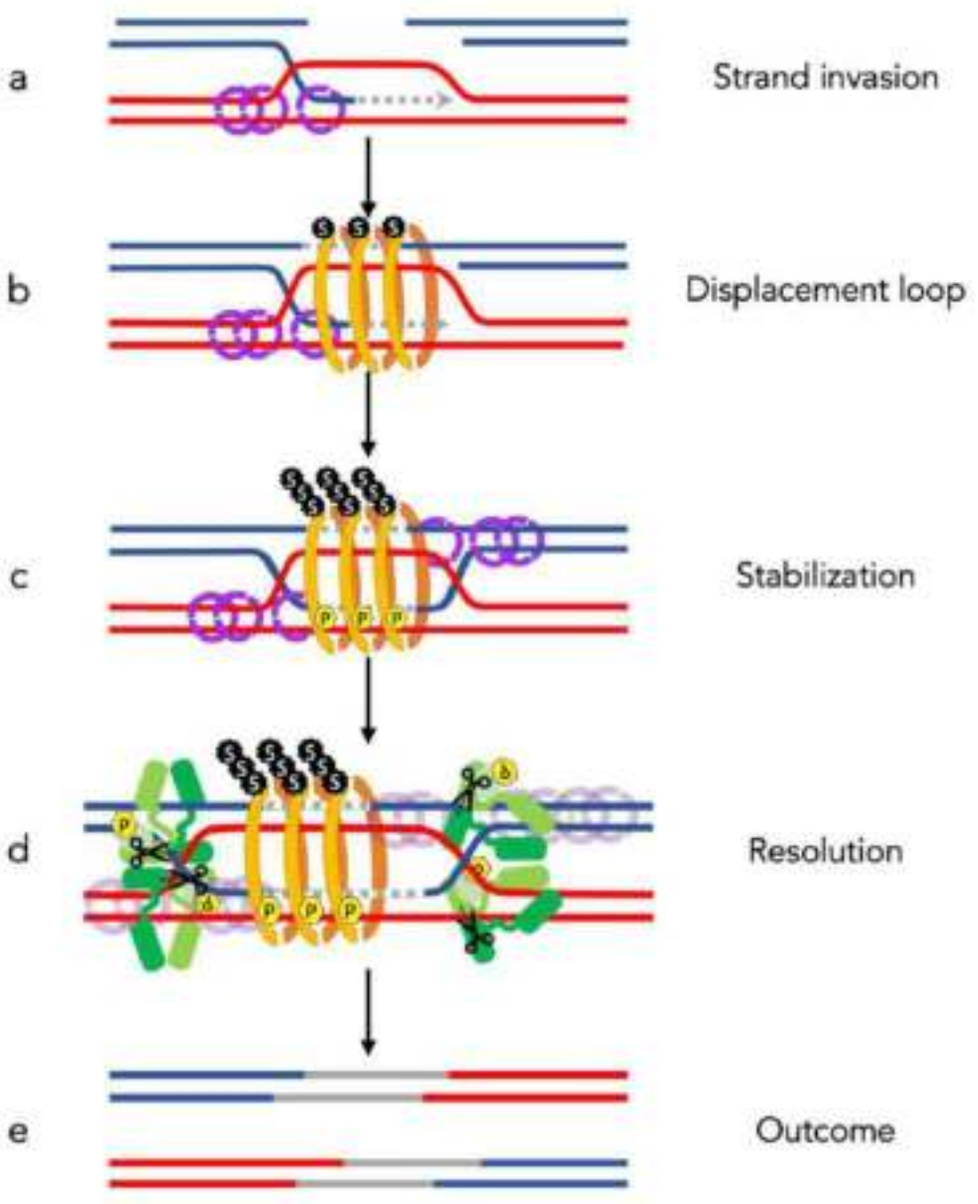


Figure 3. Representation of the recent understanding of MutSγ and MutLγ dynamics during meiotic recombination. (a) Following double-strand breaks (DSBs) during meiosis I early Prophase I, the DNA is resected and protected by RPA and RAD51 (not represented). Single-stranded DNA that harbors also DMC1, in addition to RAD51, can proceed into strand invasion. **(b)** Additionally, DSBs trigger the mono-sumolation of MSH4 by UBC9 (He et al., 2021). This is believed to promote dimerization with MSH5 and DNA loading which eventually fosters displacement loops (D-loop). **(c-e)** D-loops can further stabilized by polysumolation of MSH4 (He et al., 2021), the phosphorylation of the degron peptide present in MSH4 active pocket (He et al., 2020) and second end capture, forming a double holiday junction (dHJ). The stabilized dHJ is then bound by polymers of MutLγ (only cutting molecules represented for simplicity) which cause strand migration and finally resolution through DNA nicks oriented by the positioning of PCNA onto the dHJ (Cannavo et al., 2020; Kulkarni et al., 2020). MutLγ yields primarily Class I crossovers. Adapted from Kbiri and Ziolkowski, manuscript in preparation.

In this chapter, I explore the effect of different expression levels of the MutL genes, *MLH1, MLH3,* and *PMS1,* on meiotic crossover recombination. For this purpose, I used different mutant lines including T-DNA insertion mutants and CRISPR-cas9-mediated deletion mutants. I also used overexpression lines, at two different levels of expression. For the first set of overexpressors, I used the respective native promoters of the genes. For the second set, I used the meiosis-specific *DMC1* promoter. I assessed the crossover rate within specific intervals, *e.g.* 420 and 3.9. I also evaluated the effect of the different expression levels on *Arabidopsis thaliana* fertility.

2 Material and methods

2.1 Biological material

2.1.1 Plant material

Arabidopsis thaliana seeds for the accessions Col-0 (N1092) and L*er*-0 (NW20) were purchased from the Nottingham Arabidopsis Stock Centre (NASC). The fluorescent tagged line (FTL) Col-420 was generously shared by Professor Avraham Levy (Melamed-Bessudo et al., 2005). The other FTLs were obtained from Prof. Piotr Ziolkowski's collection.

Table 3. *Arabidopsis thaliana* mutant lines used in the study.

Gene	AGI code	Allele	Type	Line	NASC ID	Source
HEI10	AT1G53490	*hei10-2*	T-DNA insertion	Salk_014624	N514624	NASC
		hei10-4	CRISPR	-	-	Nadia Kbiri
MSH4	AT4G17380	*msh4-1*	T-DNA insertion	Salk_136296	N677967	NASC
MLH1	AT4G09140	*mlh1-1*	T-DNA insertion	-	-	Monica Pradillo
		mlh1-2	T-DNA insertion	GK-067E10	N2019332 - 48	NASC
		mlh1-3	T-DNA insertion	SK25975	N1008089	Raphael Mercier
		mlh1-4	CRISPR	-	-	Nadia Kbiri
		mlh1-5	CRISPR	-	-	Nadia Kbiri
MLH3	AT4G35520	*mlh3-1*	T-DNA insertion	SALK_015849C	N659001	NASC
		mlh3-2	T-DNA insertion	SALKseq_067953	N567953	NASC
		mlh3-3	T-DNA insertion	SALKseq_69853	N881476	NASC
		mlh3-4	CRISPR	-	-	Nadia Kbiri
PMS1	AT4G02460	*pms1-2*	T-DNA insertion	SALK_124014C	N684539	NASC
FANCM	AT1G35530	*fancm-1*	EMS	-	-	Raphael Mercier
		fancm-9	T-DNA insertion	SALK_120621	N620621	NASC
ZIP4	At5g48390	*zip4-2*	T-DNA insertion	SALK_068052	N568052	NASC
MUS81	AT4G30870	*mus81-1*	T-DNA insertion	GK-113F11		NASC
		mus81-2	T-DNA insertion	SALK_107515	N607515	NASC

2.1.2 Bacterial material

2.1.2.1 Escherichia coli

Heat-shock competent Dh5α and Top10 were purchased from Thermo-Fisher Scientific™. Genotype: F- Φ80lacZΔM15 Δ(lacZYA-argF) U169 recA1 endA1

hsdR17(rk-, mk+) phoA supE44 thi-1 gyrA96 relA1 λ. *E. coli* was used for intermediate cloning and plasmid amplification. The lines are maintained by the institute lab manager. Liquid cultures at the exponential stage are used to prepare 50 uL aliquots that are kept at -80°C till they are used for transformation.

2.1.2.2 *Agrobacterium tumefaciens*

Heat-shock competent GV3101 was used for binary cloning and plant transformation. It is resistant to rifampicin and gentamicin. The line is maintained by the institute lab manager. Liquid cultures at the exponential stage are used to prepare 50 uL aliquots that are kept at -80°C till they are used for transformation.

2.2 Chemical reagents

2.2.1 Plant culture

2.2.1.1 Fertilizers

Plants were watered three times a week. Once per week, fertilizers were added to the water (5mM KNO_3, 2mM Ca $(NO_3)_2$, 2,5mM KH_2PO_4, 2mM $MgSO_4$, 50uM Fe-EDTA, 70uM H_3BO_3, 14uM $MnCl_2$, 0,5uM $CuSO_4$, 1uM $ZuSO_4$, 0,2uM Na_2MoO_4, 10mM NaCl and 0,01uM $CaCl_2$).

2.2.1.2 Pesticide treatments

Once per month, or when needed, the plants were watered with the insecticide Substral Polysect 005 SL (Acetamiprid - 5 g/l, used at 1:100 dilution) and sprayed with the fungicide Syngenta Amistar OPTI 480 SC (32 % azoxystrobin, 0,5 % chlorothalonil, used at 1:200 dilution).

2.2.2 Polymerase chain reaction

2.2.2.1 Phire Plant Direct PCR Kit

Plants were sampled using the dilution buffer and the PCR was run according to Thermo-Fisher Scientific™ recommendations. Reagent proportions were maintained for a final reaction volume of 7, 10 or 20uL. The final volume depends on the follow-up experiments. This kit was used for genotyping for mutations and screening transformants (7uL). For amplicons that were followed by enzymatic restriction, 10uL final volume was used. For amplicons that were to be sequenced, 20uL final volume was used. Catalog number: F130WH.

2.2.2.2 DreamTaq DNA polymerase

The enzyme was purchased from Thermo-Fisher Scientific™. DreamTaq was used for colony PCR and cDNA quality testing. Recommended proportions were maintained for a final reaction volume of 10uL. Catalog number: EP0711.

2.2.2.3 CloneAmpTM HiFi PCR Premix

The enzyme was purchased from TaKaRa. It was used for high-fidelity cloning of gene ectopic expression and gRNA cloning for CRISPR-Cas9. Recommended proportions were maintained for final volumes of 10 or 20 uL. The smaller volume was favored, the higher volume was used for very long and high GC content amplicons. Catalog number: 639298.

2.2.2.4 SYBR™ Green PCR Master Mix

SYBR Green was used for real-time PCR to quantify gene expression after reverse transcription of RNA into cDNA. The PCRs were run in 384 plates and a 5uL final volume. Recommended proportions were maintained. Catalog number: 4309155.

2.2.2.5 Hiscript III 1 st Strand cDNA Synthesis Kit (+ gDNA wiper)

0,5 ug to 1ug of total RNA was used for the reverse transcription. The reaction mix was set according to the recommendations of Vazyme. Both oligo-dT and random primers were used for a 30 min reaction time. The obtained cDNA was diluted 10 times and aliquoted before being stored at -80°C. Catalog number: R312-01.

2.2.3 Kits

2.2.3.1 Nucleic acid extraction

2.2.3.1.1 GeneJET Plasmid Miniprep Kit

3 mL of *E. coli* liquid culture, grown overnight, was used for plasmid extraction. The purification was conducted according to Thermo-Fisher Scientific recommendations. A 30 to 50 uL elution volume was used according to the culture density. Catalog number: K0482.

2.2.3.1.2 GeneJET PCR Purification Kit

GeneJET PCR Purification Kit was used to purify PCR products for cloning and sequencing. DNA was purified according to Thermo-Fisher Scientific recommendations. Catalog number: K0702.

2.2.3.1.3 GeneJET Gel Extraction Kit

In-gel purification was to clean up DNA fragments of specific size after PCR or enzymatic restriction. The extraction was carried out according to ThermoFisher Scientific recommendations. Catalog number: K0692.

2.2.3.1.4 RNeasy Plant Mini Kit

Leaf or flower buds were collected in liquid nitrogen. They were ground using the Qiagen Tissue-Lyser II. RNA was extracted according to Qiagen recommendation. Total RNA was stored at -80°C. Catalog numbers: 74904 and 85300.

2.2.3.2 Cloning

2.2.3.2.1 CloneJET PCR Cloning Kit

pJET cloning was used for subcloning of PCR products for sequencing and amplification of inserts for subsequent binary cloning. Ligation of inserts was conducted according to the proportions recommended by ThermoFisher Scientific in a 10 uL final volume. Catalog number: K1232.

2.2.3.2.2 ClonExpress MultiS One Step Cloning Kit

The recombinase was used to clone single or multiple inserts into the pFGC binary vector. The cloning was conducted according to Vazyme recommendations. Catalog number: C113-02.

2.2.4 Restriction enzymes

Fast-Digest and regular restriction enzymes were purchased from ThermoFisher Scientific. These enzymes were used for dCAPS genotyping, restriction testing, and cloning. The reaction was run for 15 to 60 min when using Fast-Digest. The longer restriction time was used for cloning experiments to ensure complete digestion of the substrate. The regular enzyme reactions were run for 3 to 16h. Reaction mixes were prepared according to the recommendations of ThermoFisher Scientific. An inactivation cycle was used when possible.

2.2.5 Electrophoresis

2.2.5.1 50X Tris Acetate EDTA (TAE)

The 50X TAE stock solution of 50 mM EDTA, 2M Tris base, and 1M glacial acetic acid solution was periodically prepared by the laboratory manager. This solution was diluted 100 times for use as a buffer for electrophoresis.

2.2.5.2 Agarose

Powder agarose was purchased from ABO Sp. z o.o. It was dissolved in 0,5X TAE at concentrations from 1 to 2% according to the size of the nucleic acid to be resolved. Catalog number: BLE1.

2.2.5.3 Nucleic acid dye

SimpliSafe, the DNA stain, was purchased from EURX Sp. z o.o. It was used to visualize nucleic acid after resolution by electrophoresis and UV exposure. Catalog number: E4600-01

2.2.5.4 Nucleic acid molecular weight markers

GeneRuler DNA ladder collection from ThermoFisher was used to estimate the size of the nucleic acid run of gel. The used ladders were: 50 bp, 100bp plus, 1kb and 1kb plus. Catalog numbers in the same order: SM0371, SM0322, SM0311, SM1331.

2.3 Maps of the used vectors

2.3.1 pJet1.2

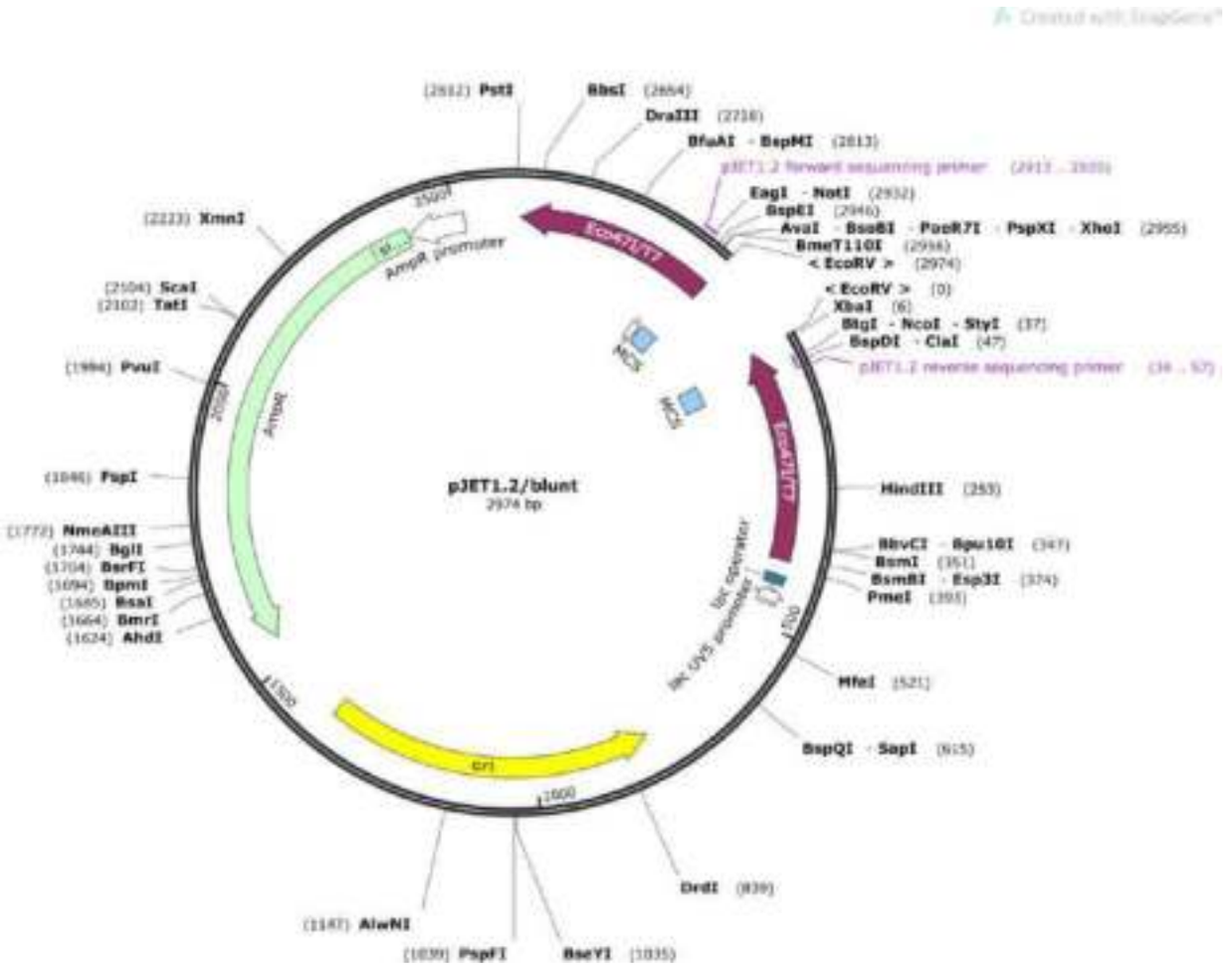


Figure 4. Blunt end open vector provided with the CloneJET PCR Cloning Kit. It confers ampicillin resistance to the transformant bacteria and activates a killer cassette if closed empty. The pJET1.2 forward and reverse primers are used for colony PCR, for checking the insert size, and for sequencing.

2.3.2 pJET1.2-U3

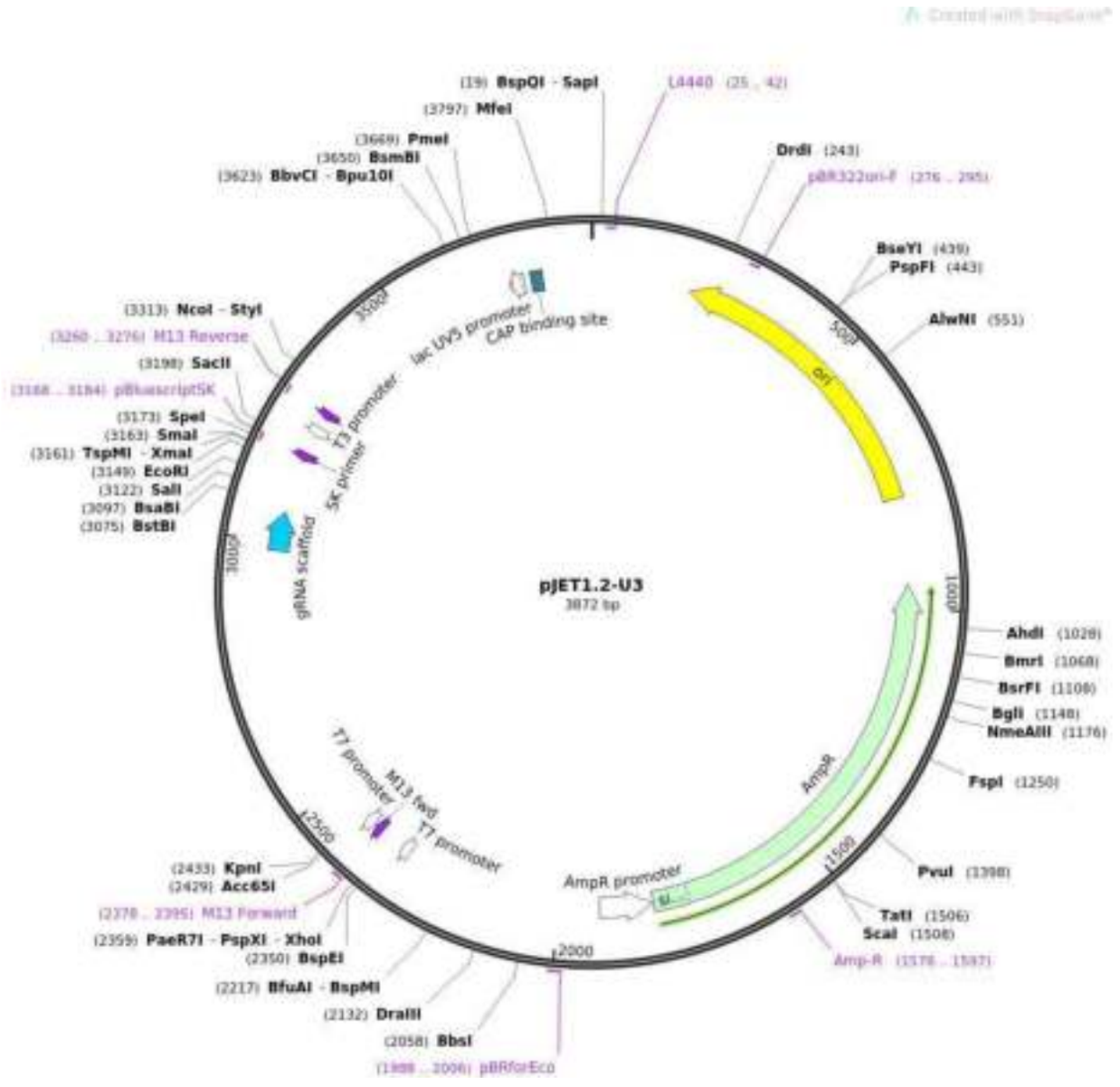


Figure 5. Modified pJet1.2 where the U3 promoter and binary cloning overhangs were added. It is used for cloning gRNAs under the control of U3 promoter. As the original pJET1.2, it confers ampicillin resistance to bacteria. The modified vector was developed by Dr. Tomasz Bieluszewski (Bieluszewski et al., 2022). Addgene catalog number: 173156.

2.3.3 pJET1.2-U6

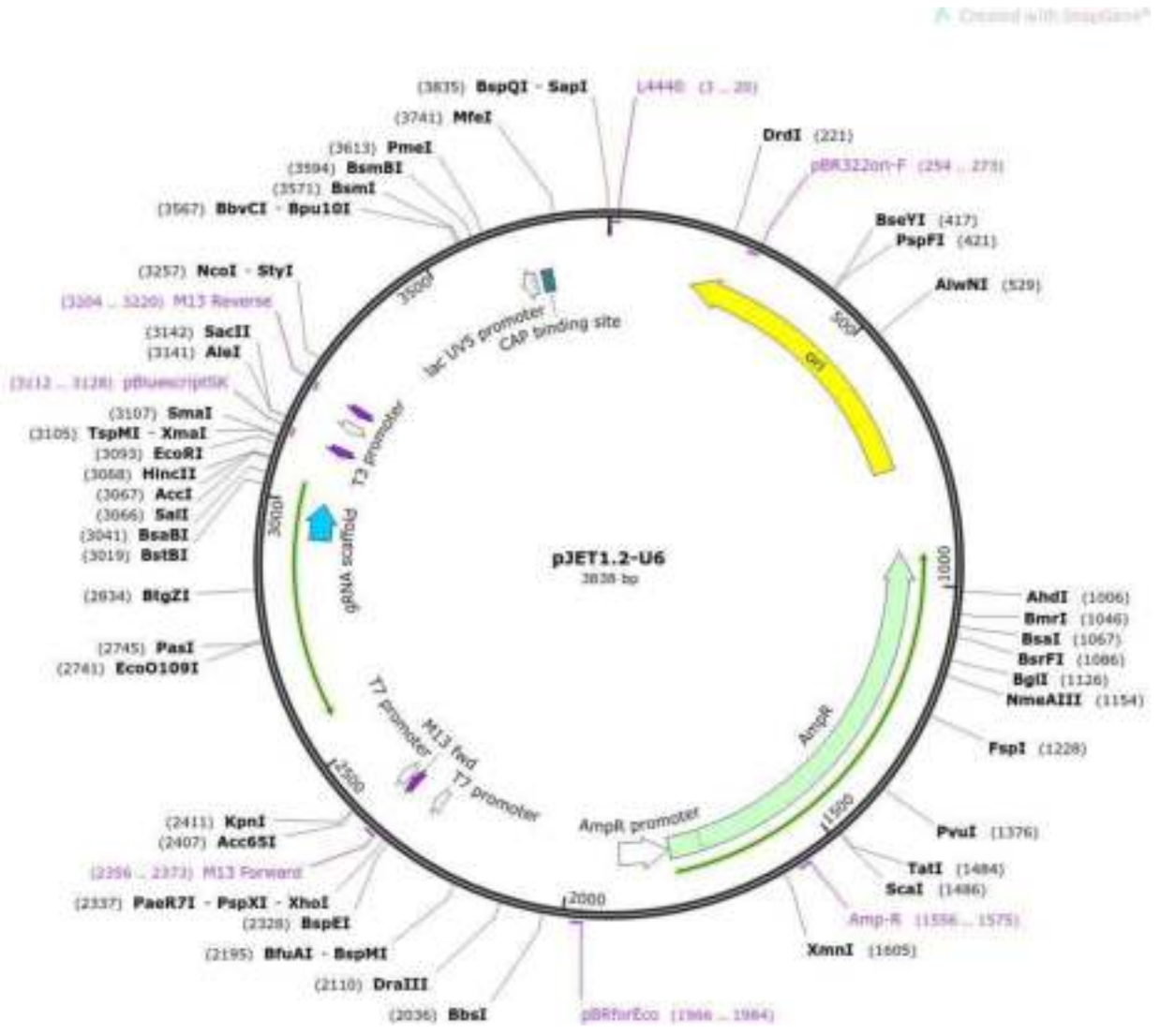


Figure 6. Modified pJet1.2 where the U6 promoter and binary cloning overhangs were added. It is used for cloning gRNAs under the control of U6 promoter. As the original pJET1.2, it confers ampicillin resistance to bacteria. The modified vector was developed by Dr. Tomasz Bieluszewski (Bieluszewski et al., 2022). Addgene catalog number: 173157.

2.3.4 pFGC-I2Cas9 vector

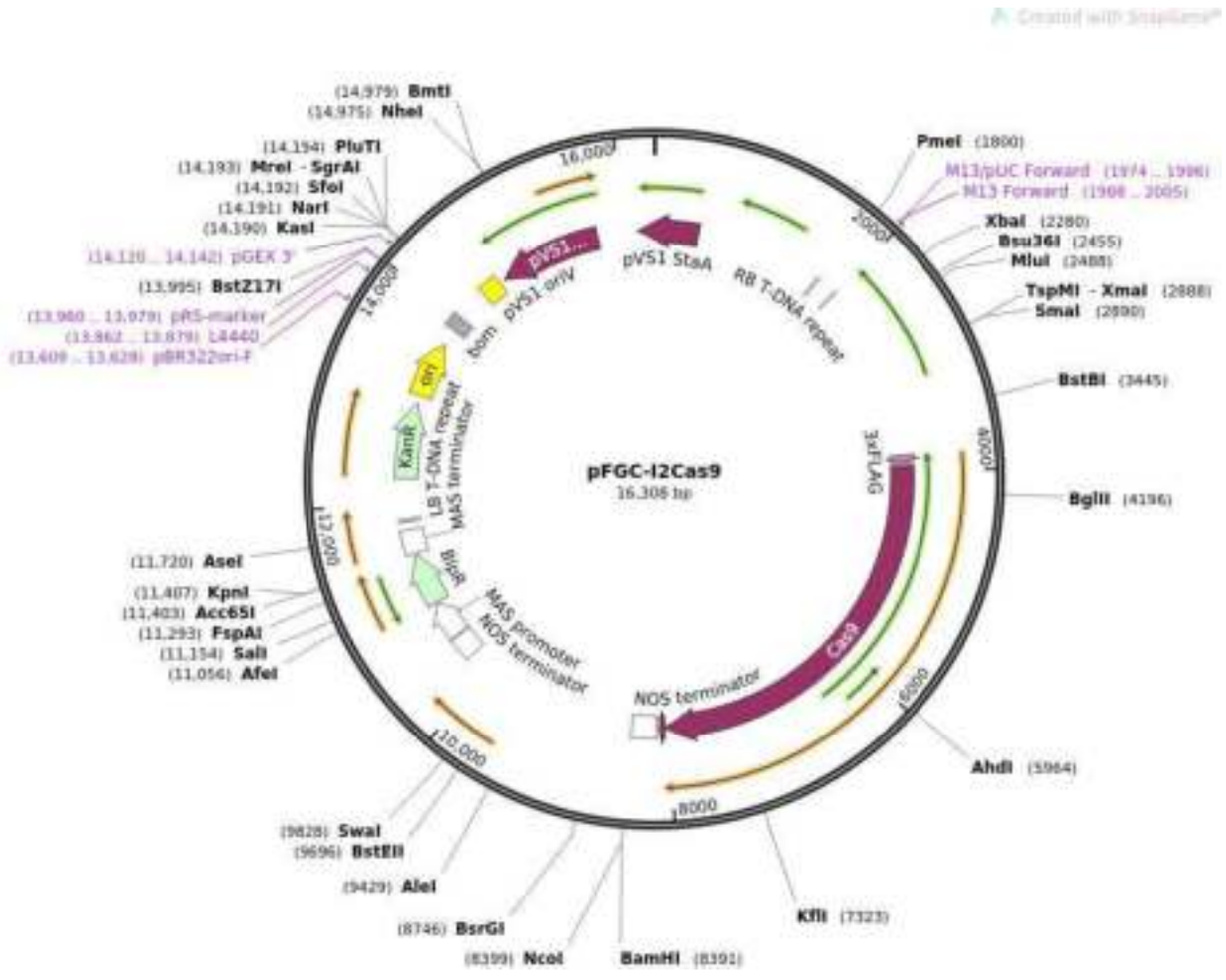


Figure 7. Modified pFGC vector where CRISPR and Cas9 coding sequences were introduced. The vector is opened using BamHI restriction enzyme to clone the gRNAs. It is also used for overexpression by opening the vector using HindIII in addition to BamHI. This cuts out the CRISPR-Cas9 cassettes. pFGC confers a Kanamycin resistance to the transformant bacteria and a BASTA resistance to the transformant plant subsequently. pFGC-I2Cas9 modified vector was developed by Dr. Tomasz Bieluszewski (Bieluszewski et al., 2022). Addgene catalog number: 173158.

2.4 Methods

2.4.1 Fertility assays:

2.4.1.1 Seed set

Five siliques starting from the seventh oldest silique of the main stem were collected and discolored in 96% ethanol for at least 3 days. The samples were pictured using the Zeiss Lumar V12 Fluorescence Stereomicroscope at the

magnification 6.4X. ImageJ was used to count the number of seeds per silique and the silique length in centimeters.

2.4.1.2 Alexander staining

Alexander stain was prepared following the Peterson et al., 2010 protocol. The working solution was 10% ethanol, 0.01% Malachite green, 25% glycerol, 0.05% Fuchsin acid, 0.005% Orange G and 4% glacial acetic acid diluted in sterile MilliQ water. The stain was then kept in an amber glass bottle and stored in the dark. Pollen viability and density were investigated as in Alexander, 1969 and Hord et al., 2008

2.4.1.3 Pollen viability

Five to ten stage 15 (open flower) were immersed in alexander staining to obtain a pollen suspension. This suspension was then mounted between a slide and cover slip and observed at 10X magnification under the Leica DM4 B using the bright field. Viable pollen grains were colored in magenta/red and perfectly round shaped. The dead pollen grains are green/brown and are scrunched and lost their round shape. 500 pollen grains from three replicates (1500 events in total) were processed for each genotype.

2.4.1.4 Pollen density

Stage 12 flower buds were collected from three different plants for each genotype. They were discolored using Carnoy fixative for 1h, then incubated for at least one week in Alexander staining at 4°C or for 7h at 55°C. Arabidopsis flowers have 6 anthers, 4 bigger anthers, and 2 smaller ones. Two stage 12 flower buds were dissected for each one of the triplicates, and three of the bigger anthers were mounted between a slide and cover slip. The sampled anthers were observed under the bright field using the Leica DM4 B at magnification 20X.

2.4.2 CRISPR-Cas9 mutagenesis

Guide RNAs (gRNAs) were designed according to the protocol designed by Bieluszewski et al., 2022. CRISPOR online software, http://crispor.tefor.net, was used to identify Protospacer Adjacent Motifs (PAM) and potential gRNAs. The target region was input into the "Step 1" window. "Arabidopsis thaliana – Thale-cress – Ensemblplants 76 (TAIR10)" was the selected genome for "Step 2", 20bp-NGG – Sp Cas9, SpCas9-HF1, eSpCas9 1.1 was selected for "Step 3" to identify the PAMs for Cas9. gRNAs were selected based on their position in the genome, potential off-targets, and predicted efficiency. gRNAs targeting Exons were favored. No more than 3 off-targets were tolerated. Off-targets were checked for their possible involvement in meiosis, recombination, and DNA integrity. Both Predicted efficiency, Doench '16 and Mor-Mateos, scores must be above 50. The gRNAs were cloned under the control of U3 or U6 promoter then introduced to a modified pFGC binary vector that carries CRISPR and Cas9. This vector gives kanamycin resistance at the bacteria level and BASTA resistance at the plant level. A variant without the fluorescent marker dsRed was used as the transformed plants had fluorescent markers from the FTLs.

2.4.2.1 *MLH1* mutagenesis

Three gRNAs were targeted to the region from the 4th intron to the 6th exon of *MLH1*, to make sure to target both splicing variants (Figure 8A). They were cloned in two combinations, gRNA1 & 2 and gRNA1 & 3. Deletion mutants were obtained from the first combination. Four independent mutants were selected and sequenced.

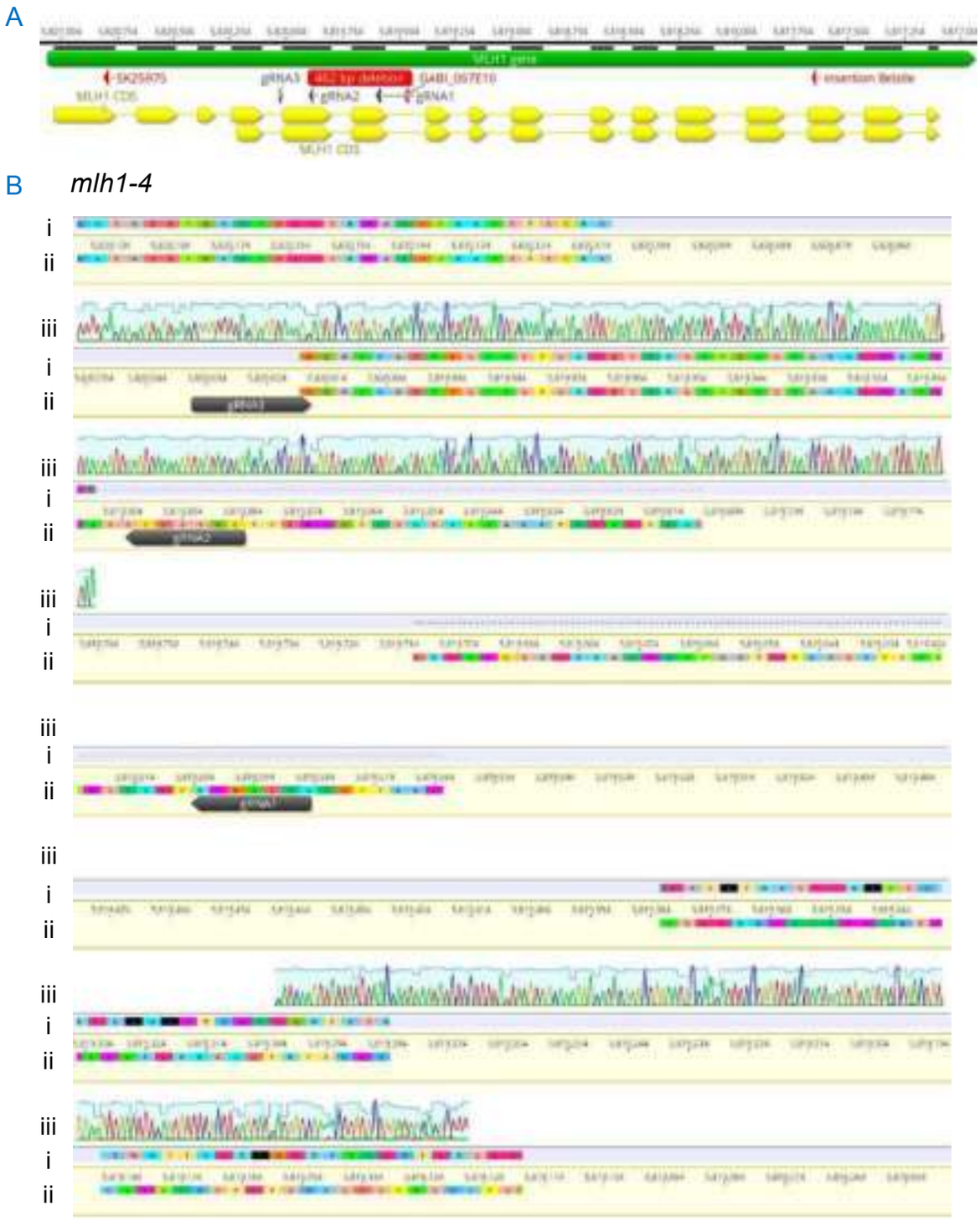


Figure 8. *mlh1-4* CRISPR-Cas9 mediated deletion mutant in Col-0. (A) Scheme showing the *MLH1* gene structure. The exons are represented with yellow arrows, the gRNAs used for the mutagenesis in grey arrowheads, the obtained 462 bp deletion in a red rectangle. The position of the T-DNA insertions of the other MLH1 mutants is represented with red arrowheads. (B) DNA sequence of the *mlh1-4* mutant aligned to the wildtype reference. **i.** Translation of *mlh1-4*. **ii.** Wildtype reference. **iii.** Sequencing of *mlh1-4*. The 462 bp in genomic and 250 bp in coding sequence deletion introduces a frameshift and multiple STOP codons. The STOP codons are represented with black rectangles. The scale in "A" and "ii. Wildtype reference" represents the genomic position, the reverse sequence is used for simplicity.

All tested individuals showed the same 462 bp genomic deletion and 250 bp coding sequence deletion. Later testing through RT-seq shows that the deletion at the transcript level is 298 bp big, introducing multiple stop codons and a frameshift. Two of these lines were used for recombination frequency scoring, fertility assays, and cytology.

Figure 9. *mlh3-4* CRISPR-Cas9 mediated deletion mutant in Col-0. (A) The structure of the *MLH3* gene. The exons are represented with yellow arrows, the gRNAs used for the mutagenesis in grey arrowheads, the obtained 234 bp deletion in a red rectangle. The position of the T-DNA insertions of the other MLH1 mutants is represented with red arrowheads. (B) *mlh3-4* alighed to the wildtype reference. **i.** Translation of *mlh3-4*. **ii.** Wildtype reference. **iii.** Sequencing of *mlh3-4*. The 234 bp in genomic/coding sequence deletion is inframe.

2.4.2.2 *MLH3* mutagenesis

Thirteen gRNAs were targeted to different regions of *MLH3*, making sure to target all/most predicted splicing variants. The gRNAs were cloned in pairs. Deletion mutants were only obtained from the gRNA11 & 13 combination.
Six independent mutants were selected and sequenced. They all showed the same 234 bp genomic /coding sequence deletion. The deletion did not induce a frameshift or STOP codons. This mutant cannot be used as it is more likely to produce defective proteins rather than a null mutant. The CRISPR-Cas9 construct was maintained for one more generation to induce novel mutations or sub-mutations. Therefore, I used *mlh3* insertional mutants for further experiments.

2.4.3 Sanger sequencing

Sanger sequencing was used to check cloning constructs, deletion positions, and exact sizes. It was entrusted to the Molecular Biology Techniques Laboratory at Adam Mickiewicz University.

2.4.4 Nucleic acid quantification

2.4.4.1 Nanodrop

Denovix DS-11+ nanodrop was calibrated using sterile milliQ water. 2ul of purified plasmid, PCR product, genomic DNA or coding DNA were quantified using the "double-stranded DNA" built-in standards. TE (Tris-EDTA) was used to blank. Total RNA was quantified using the "RNA" built-in standards.

2.4.4.2 Quibit

Qubit 4 fluorometer was used to quantify genomic DNA and whole genome tagmented libraries. The samples were prepared according to ThermoFisher 1X dsDNA HS (high sensitivity) assay kit.

2.4.5 Bacteria transformation and selection:

2.4.5.1 *Agrobacterium tumefaciens*:

10% V with 10 to 100 ng of binary vector were added to an aliquot of heat-shock competent GV3101 *Agrobacterium tumefaciens*. The bacteria were then incubated in ice for 5 min, then in liquid nitrogen for 5 min, and finally at 37°C for 5 min. 700 uL of sterile LB were added then the bacteria were placed in a thermos-mixer at 28°C | 850 rpm for 3 to 4h for recovery (Weigel and Glazebrook, 2002). After recovery, the bacteria were plated on a selective medium: Lysate Broth (LB), 50 ug/mL Kanamycin, 40 ug/mL gentamycin, 80 ug/mL rifampicin. GV3101 harbors gentamycin and rifampicin resistance. The kanamycin resistance is introduced by the binary vector. The bacteria are left to grow for 48h at 28°C.

2.4.5.2 *Escherichia coli*:

10% V containing 10ng of circular plasmid were added to heat-shock competent Dh5α *Escherichia coli*. The bacteria were then incubated for at least 20 min in ice, followed by a heat-shock consisting of 1min 30 sec incubation at 42°C and 2 min incubation in ice. 700 uL of sterile LB were added to each aliquot and the bacteria were put for recovery for 1h at 37°C with 850 rpm shaking. Finally, the bacteria were span down for 2min at 4500 rpm and plated on solid LB with the appropriate antibiotics. Transformant colonies were obtained after a 16h growth at 37°C.

2.4.6 Plant transformation and selection:

2.4.6.1 Floral dip:

Single *Agrobacterium tumefaciens* colonies with the vectors of interest were inoculated into 20 mL of LB with 50 ug/mL Kanamycin, 40 ug/mL gentamycin, and 80 ug/mL rifampicin and grown for 24h. 100 uL of the saturated culture were inoculated into 100 mL of fresh LB, 50 ug/mL Kanamycin, 40 ug/mL gentamycin

and 80 ug/mL rifampicin and grown for 16h. The cultures were span down at 4500 rpm and resuspended in 200 mL of 5% (W:V) sucrose and 0.005% (V:V) Silwet-11. 5 weeks old plants, about 10 cm long stems, were dipped into the bacteria suspension for 1 min then laid down overnight in a humid and dark container (Weigel and Glazebrook, 2002). The following day they were tied and put back into standard culture conditions. To increase transformation yield, plants were dipped twice with a one week interval. These dipped plants are the T0 generation.

2.4.6.2 BASTA selection:

One week old seedlings were sprayed three times with 60mg/L BASTA 150 SL (150 g/L glufosinate-ammonium) Bayer, over the course of one week. BASTA operates by inhibiting the glutamine synthase. This disrupts plants metabolisms at multiple levels which stops growth and leads to the death of the non-resistant plants. Alternatively, transformant resistant plants are able to grow. They appear as the only green healthy plants (Weigel and Glazebrook, 2002).

3 Results

3.1 MutL mutants do not show haploinsufficiency in Arabidopsis

Haploinsufficiency is the phenomenon by which the presence of only one functional copy of the two allelic copies of a gene in a diploid organism is not sufficient to maintain a wildtype phenotype (Veitia, 2002; Johnson et al., 2019; Morrill and Amon, 2019). This is for example true for the meiosis-specific E3 ligase HEI10. Indeed *hei10-2 +/-* Arabidopsis plants display a lower crossover recombination level than the wild type, and higher than the homozygous mutant. The homozygous mutant shows very low recombination frequency and segregation distortion. It is not scorable using the fluorescent tag system. *HEI10* also displays a dosage effect where the number of copies correlates positively with the recombination crossover rate (Figure 10) (Ziolkowski et al., 2017). Haploinsufficiency was observed for *MLH1* in mouse and human models (Wang et al., 2012a; Shrestha et al., 2020; Harada et al., 2021; Shrestha et al., 2021). Therefore, the starting point was to check if the haploinsufficiency and dosage effect phenotypes are observed in Arabidopsis for the *MutL* genes.

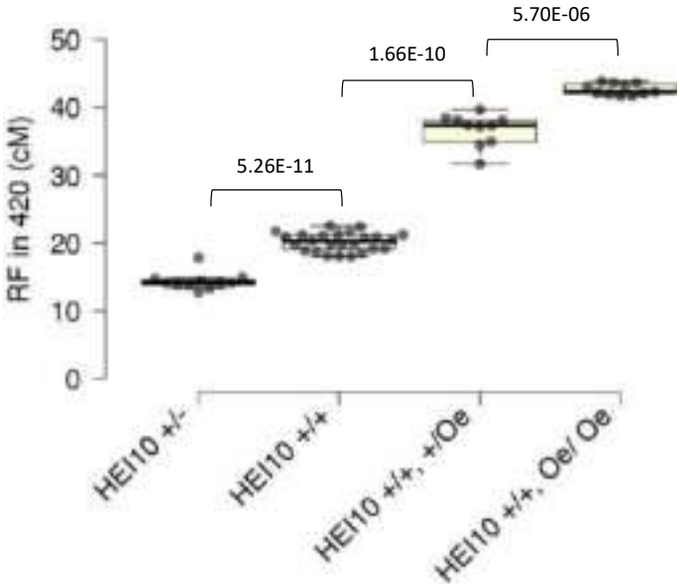


Figure 10. Haploinsufficiency and dosage effect of *HEI10* expression level in Arabidopsis meiotic crossover recombination. Dosage effect of *HEI10* expression in Arabidopsis on recombination frequency as measured (RF) in the 420 interval. *hei10 -/-* is not plotted because of segregation distortion making measurements unreliable. The heterozygous mutant *hei10 +/-* shows lower RF than *HEI10 +/+*, which has lower RF than *HEI10 +/+ Oe +* (overexpression construct in hemizygous state), which all are lower than *HEI10 +/+ Oe + +* (overexpression construct in homozygous state). One-tail T-test values, with a 5% accepted error, are represented between samples connected with brackets.

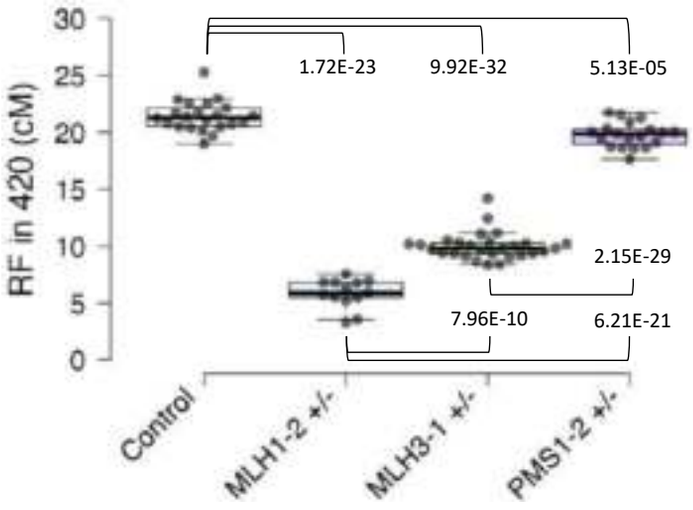


Figure 11. Decreased crossover recombination rate in the heterozygous mutants of *MutL* genes. All three, *mlh1-2 +/-*, *mlh3-1 +/-*, and *pms1-2 +/-* show a significant decrease in RF. One-tail T-test values, with a 5% accepted error, are represented between samples connected with brackets.

To investigate this hypothesis, I used T-DNA insertional mutants for the three MutL subunits, MLH1, MLH3, and PMS1. The T-DNA insertions for all three alleles, *mlh1-2*, *mlh3-1,* and *pms1-2,* are at the beginnings of the genes, intron 6/15, exon 10/24, and exon 2/11 respectively (Supplemental figure 18). All three yield null mutants in a homozygous state (tested with RT-PCR). *mlh1-2+/-*, *mlh3-1+/-,* and *pms1-2+/-* were crossed to Col-420 and scored at the heterozygous state (Figure 11). *mlh1-2 +/-* and *mlh3-1 +/-* initially showed a dramatic decrease in recombination frequency in the 420 chromosome 3 subtelomeric region. *pms1-2 +/-* showed a slight but significant decrease. This phenotype was very mild and overall looked wildtype-like (Figure 11).

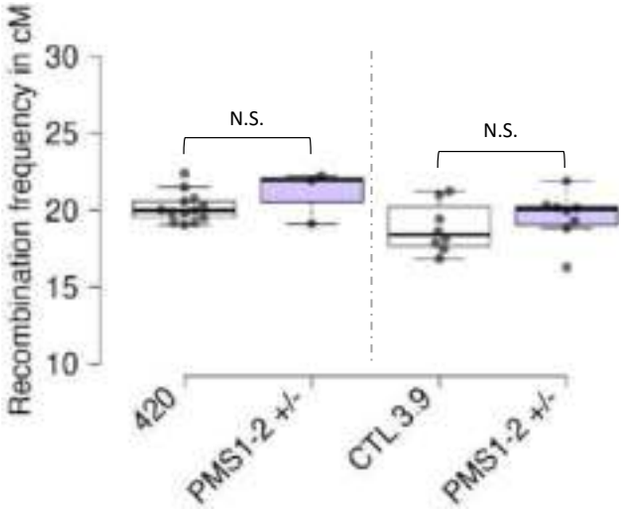


Figure 12. Recombination frequency for *pms1-2* heterozygous mutant in 420 and 3.9 intervals in F2 generation. One-tail T-test values, with 5% accepted error, are represented between samples connected with brackets. N.S.= Not Significant, p >0.05.

In filial generations, segregation issues were observed for *mlh1-2* and *mlh3-1* mutant lines. In the *mlh1-2* case, I observed that the recombination phenotype and

the genotype do not co-segregate. In the *mlh3-1* case, I could not obtain homozygous mutants that still carried the 420 fluorescent tags. The segregation of the insertion when seeds were preselected for the hemizygosity of the fluorescent tags showed a strong bias against the mutant allele. The *pms1-2* mutant did not show any segregation issues in filial generations for the insertion and the fluorescent tags. However, the slight decrease in recombination frequency for *pms1-2* was not confirmed in filial generations both in 420 and 3.9 intervals: recombination frequency was at the wildtype level (Figure 12). The small decrease observed initially is probably not due to the *pms1-2* mutation and was quickly lost in filial generations thanks to the homogenization of the genetic background.

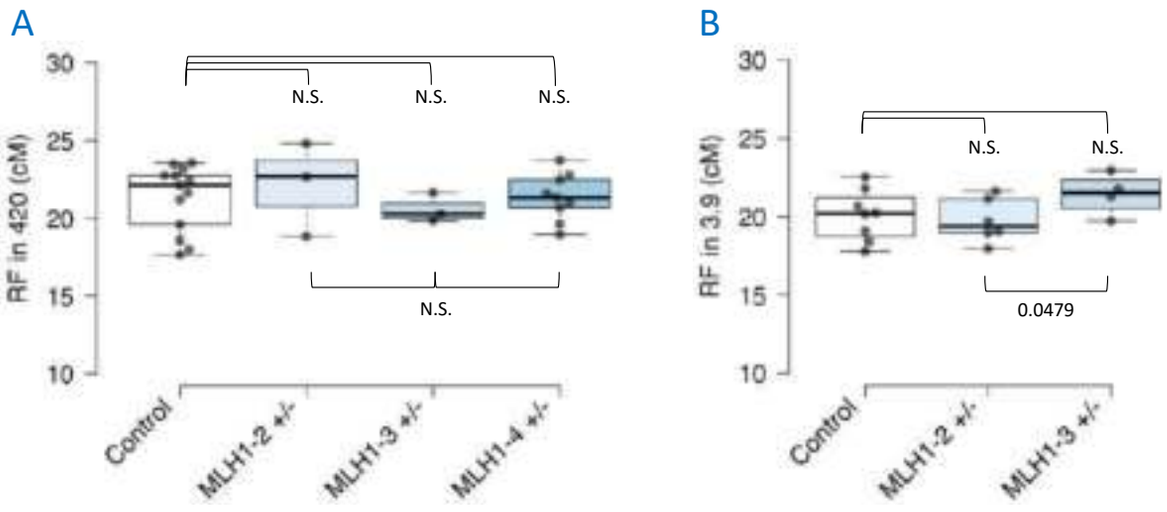


Figure 13. Recombination frequency for *MLH1* mutants in the heterozygous state. (A) RF in the subtelomeric 420 interval for *mlh1-2 +/-* BC, and two additional mutants, T-DNA insertion *mlh1-3 +/-* and CRISPR *mlh1-4 +/-*. (B) RF in the pericentromeric 3.9 interval for *mlh1-2 +/-* and *mlh1-3 +/-*. One-tail T-test values, with a 5% accepted error, are represented between samples connected with brackets. N.S.= Not Significant, $p > 0.05$.

mlh1-2 +/- was backcrossed twice to Col-0. The segregation of the fluorescent tags and the mutation were checked and followed mendelian segregation. However, the reduced recombination phenotype for the heterozygous mutant state was lost (Figure 13). The reduction in RF was only observed in the homozygous mutant state. In addition, I acquired and generated additional mutants for *MLH1*: *mlh1-3,* a T-DNA insertion mutant (1st exon), and *mlh1-4* and *mlh1-5* CRISPR-Cas9 deletion mutants, in Col and L*er* backgrounds respectively (Figure 8). The three new alleles

are knock-out mutants (tested with RT-PCR and RT-seq). *mlh1-3* and *mlh1-4* heterozygous mutants showed a wildtype-like phenotype in the 420 and 3.9 tested intervals (Figure 13).

For *mlh3-1*, I performed three backcrosses to Col-420 and Col-0. The resulting plants still exhibited a reduced 420 crossover frequency (Figure 14A). I also crossed *mlh3-1+/-* mutant to Col-3.9, which carries fluorescent tags spanning the pericentromeric region of chromosome 3. In contrast to the 420 interval, *mlh3-1* showed a significant increase in recombination frequency at the 3.9 interval (Figure 14B). To validate these results, I acquired additional insertional mutant alleles, *mlh3-2* and *mlh3-3*, which I crossed to Col-420 and Col-3.9 reporter lines. *mlh3-2* is null and *mlh3-3* yields a truncated transcript missing the endonuclease domain (Supplemental figure 18). Both alleles did not reveal any change in recombination frequency as compared to wild-type plants (Figure 14A, B). I also crossed *mlh3-1* +/- to other fluorescent tagged lines with intervals located on other chromosomes or genetic backgrounds (Col Tagged Lines (CTLs) 1.18, 5.1 and 5.2, and L*er* TLs (LTLs) 3.4 and 5.5). *mlh3-1* +/- did not display any reduction in crossover frequency as observed in 420 when compared to the wildtype in the novel intervals (Figure 14C). Therefore, I concluded that the observed changes in recombination frequencies for the 420 and 3.9 intervals are due to a local rearrangement in the *mlh3-1* mutant within the region corresponding to the 420 interval, e.g. an inversion, that results in crossover repression, and not due to the haploinsufficiency of the *MLH3* gene. The slight, yet significant, increase in 3.9 could be due to crossover assurance and compensation for the decreased recombination rate in the subtelomeric region of chromosome 3. However, many of the results presented in this thesis concern the *mlh3-1* mutant because they were performed before other mutant alleles of this gene were available and before it became clear that *mlh3-1*

was affected by a structural rearrangement. I tried to present them taking into account the possible impact of an additional mutation on the observed phenotype.

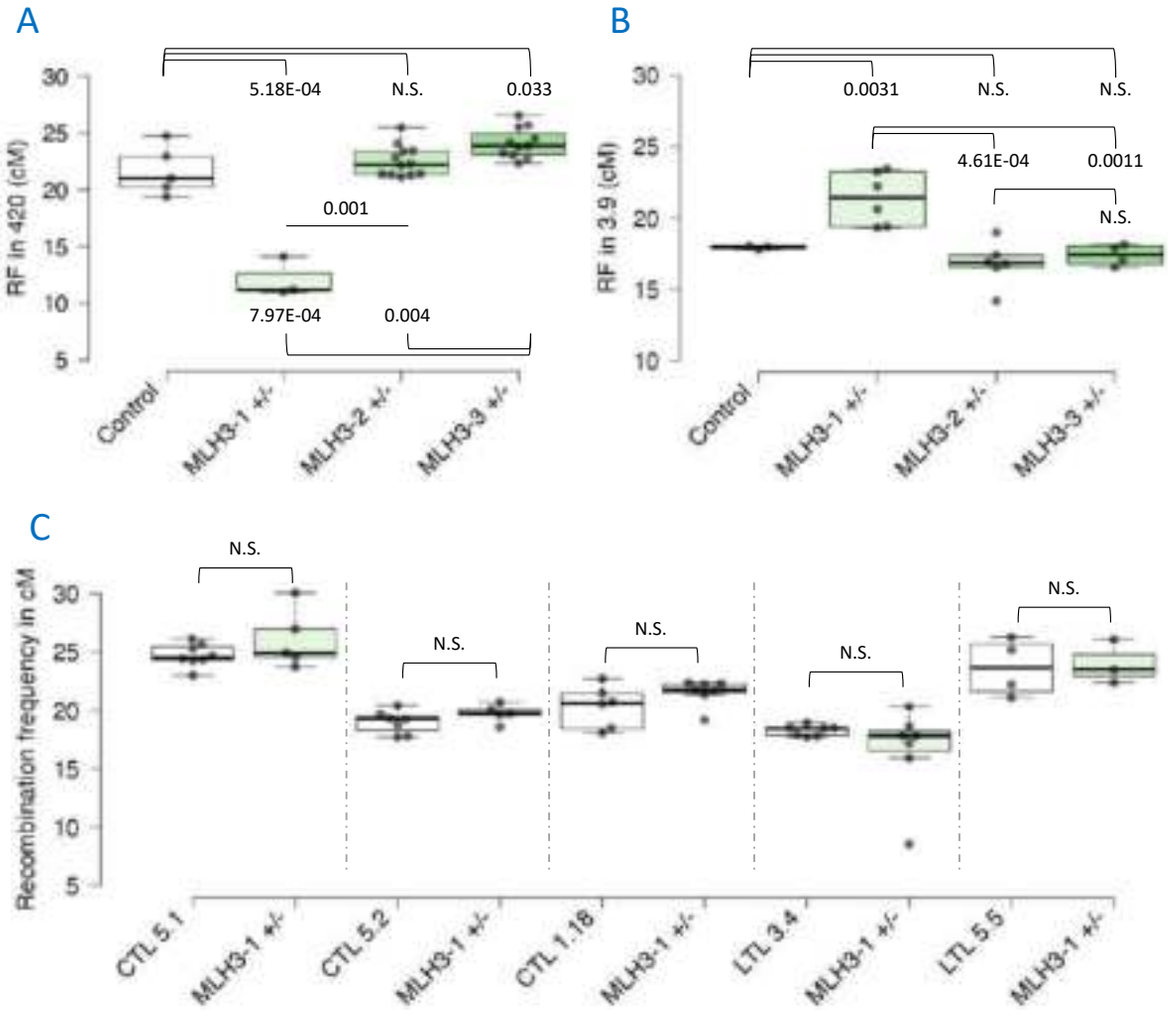


Figure 14. Recombination frequency for *MLH3* mutants in the heterozygous state. (A) RF in the subtelomeric 420 interval for *mlh1-3 +/-* BC, and two additional T-DNA insertion mutants, *mlh3-2 +/-* and *mlh3-3 +/-*. (B) RF in the pericentromeric 3.9 interval for the same lines. (C) RF in several intervals for *mlh1-3 +/-*. One-tail T-test values, with a 5% accepted error, are represented between samples connected with brackets. N.S.= Not Significant, $p > 0.05$.

3.2 Only MutLγ subunits affect meiotic crossover recombination frequency

MLH1 and MLH3 form the MutLγ heterodimer with an endonuclease activity which is responsible for the resolution of dHJs into class I crossovers (Hunter, 2015; Mercier et al., 2015; Ziolkowski, 2022). *PMS1* was investigated for its role in MMR but never in the context of meiotic crossover recombination in plants (Alou et al., 2004; Li et al., 2009). To confirm the effect of these proteins on crossover formation

in Arabidopsis, I used mutants of *MLH1, MLH3* and *PMS1* genes in homozygous state to score crossover recombination frequencies in different fluorescent tagged intervals. The homozygous *pms1-2* mutant does not show any effect on crossover recombination frequency in the two tested intervals 420 and 3.9, located in subtelomeric and pericentromeric regions of chromosome 3, respectively (Figure 15).

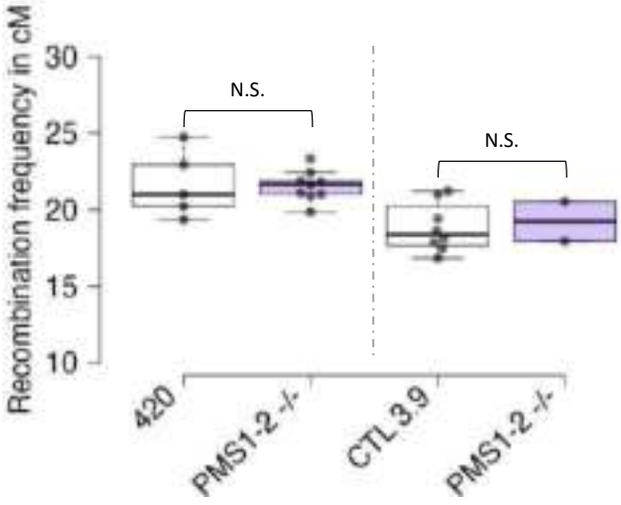


Figure 15. Recombination frequency for *pms1-2 homozygous* mutant in 420 and 3.9 intervals. One-tail T-test values, with 5% accepted error, are represented between samples connected with brackets. N.S.= Not Significant, p >0.05.

I selected homozygous mutants for *MLH1* and *MLH3* in different intervals: 420, 3.9, 1.18, and 5.1. All three *MLH1* alleles, i.e., *mlh1-2, -3,* and *-4*, showed the same decrease in recombination frequency in all tested intervals (Figure 16). *mlh1-/-* shows a 30% to 40% decrease in recombination frequency in comparison to the wildtype control (Figure 16). As for *mlh3 -/-*, I tested three different alleles, *mlh3-1, -2, and -3*. I refrained from testing the *mlh3-1 -/-* recombination frequency on chromosome 3 because of the behavior observed in the heterozygous state indicating an accompanying structural rearrangement. In 1.18 and 5.1 intervals, *mlh3-1 -/-* shows a significant decrease in crossover recombination frequency, about 40% less than the wildtype controls (Figure 17C-D). *mlh3-2* shows a significant 26% decrease in RF in 420 (Figure 17A) and a 20% decrease in 3.9 in comparison to the wildtype controls (Figure 17B). The *mlh3-3 -/-* shows no effect on recombination frequency when the whole tested populations are considered in 420 and 3.9 intervals. However, a clear segregation of crossover rate is observed in 420, where two very distinct populations can be observed (Figure 17A). The two

populations were separated and statistically tested in comparison to the wildtype control. The lower population shows a significant 25% decrease whereas the higher population shows a significant 10% increase. This phenotype can be due to the location of the T-DNA insertion in an intron, close to the end of the gene, probably within the endonuclease domain. The T-DNA could be spliced out during mRNA maturation resulting in a wildtype-like behavior, or maintained causing an inactive protein, and decreased recombination (Figure 17A-B).

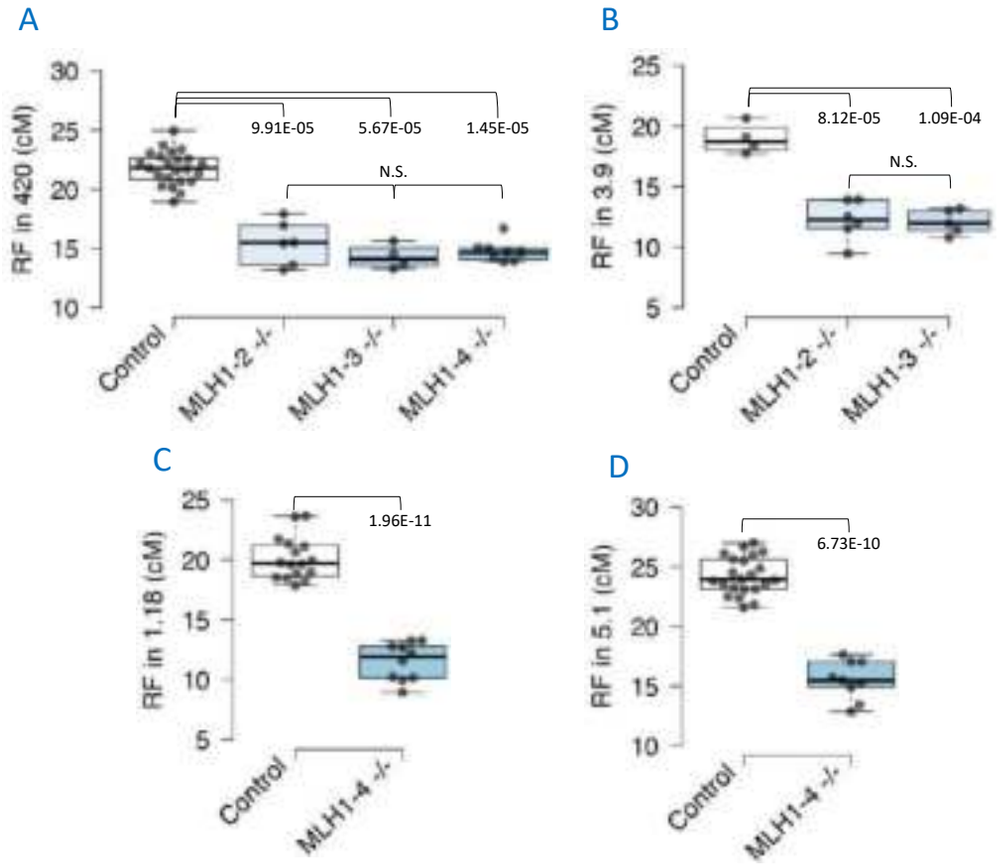


Figure 16. Recombination frequency for *MLH1* homozygous mutants. Three *MLH1* alleles are represented, the T-DNA insertion mutants *MLH1-2* and *MLH1-3* and the CRIPSR-Cas9 deletion mutant *MLH1-4*. Measurements were made in 420 and 3.9 intervals for the two T-DNA mutants and 420, 1.18, and 5.1 intervals for the CRISPR mutant. One-tail T-test values, with a 5% accepted error, are represented between samples connected with brackets. N.S.= Not Significant, p >0.05.

Overall, *mlh1 -/-* and *mlh3 -/-* show the decreased crossover recombination frequency expected from their role as the main resolvase of class I crossovers in Arabidopsis. However, only a 30% to 40% decrease is observed. This is very mild when considering that class I is responsible for 85% to 90% of the crossovers in

Arabidopsis. In contrast, the ZMM mutants show a decrease proportional to their expected role (Chelysheva et al., 2007; Macaisne et al., 2011; Chelysheva et al., 2012; Mercier et al., 2015). These results suggest that MutLγ is not the exclusive resolvase for class I crossovers in *A. thaliana*.

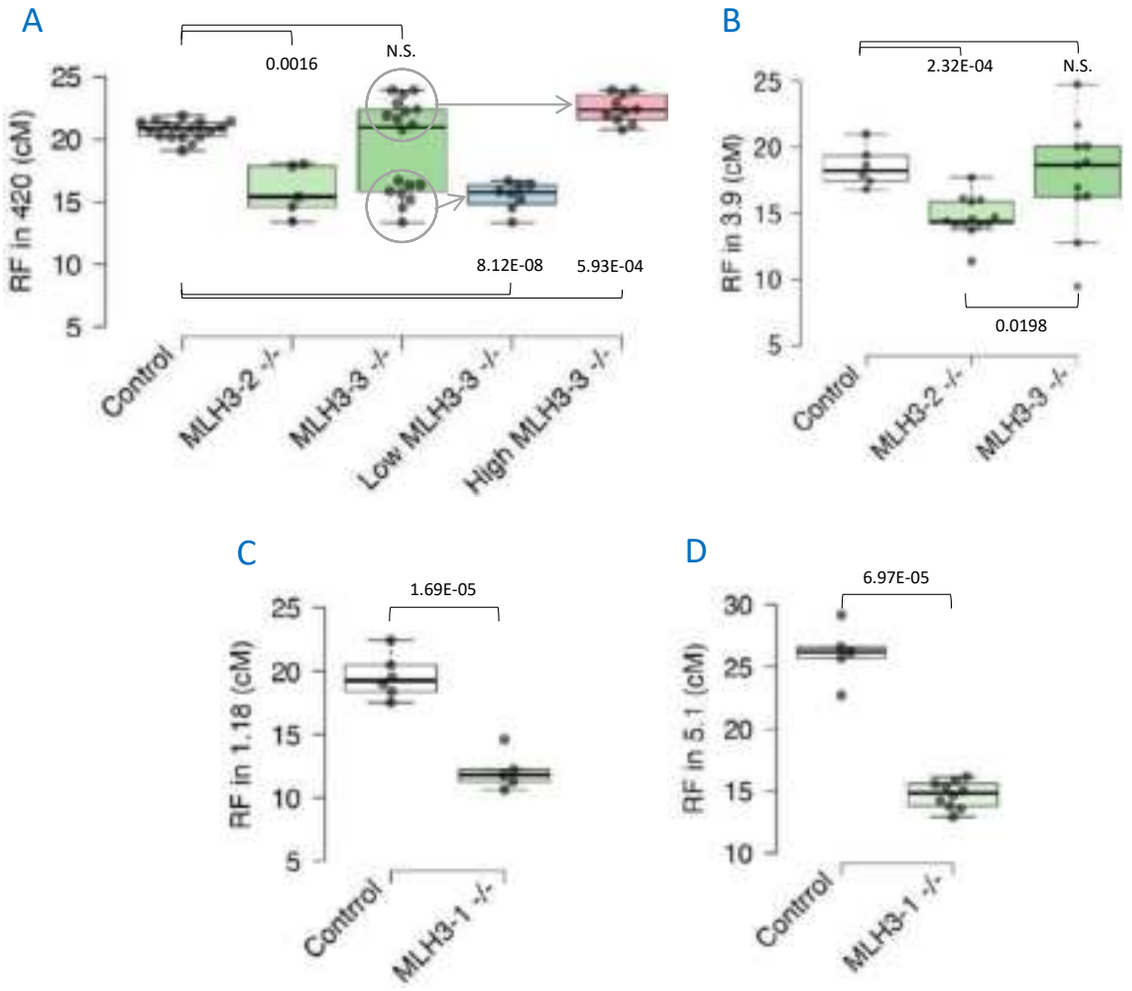


Figure 17. Recombination frequency for *MLH3* homozygous mutants. Three *MLH3* T-DNA insertion mutants are represented, *mlh3-1* and *mlh3-2* and *mlh3-3*. Measurements were made in 420, 3.9, 1.18, and 5.1 intervals. One-tail T-test values, with a 5% accepted error, are represented between samples connected with brackets. N.S.= Not Significant, $p > 0.05$.

3.3 MutLγ loss of function affects Arabidopsis fertility and chiasma formation

To further characterize the effect of *MLH1* and *MLH3* loss of function on crossover recombination, cytogenetic characterization of *mlh1-4* and fertility were assessed. Chiasmata are the physical manifestation of recombination crossover events. Fertility was quantified through silique length, seed set, pollen density, and pollen

viability. The success rate of pollen and fruit production reflects the success rate of meiotic divisions as they are products of the spores obtained at the end of meiosis.

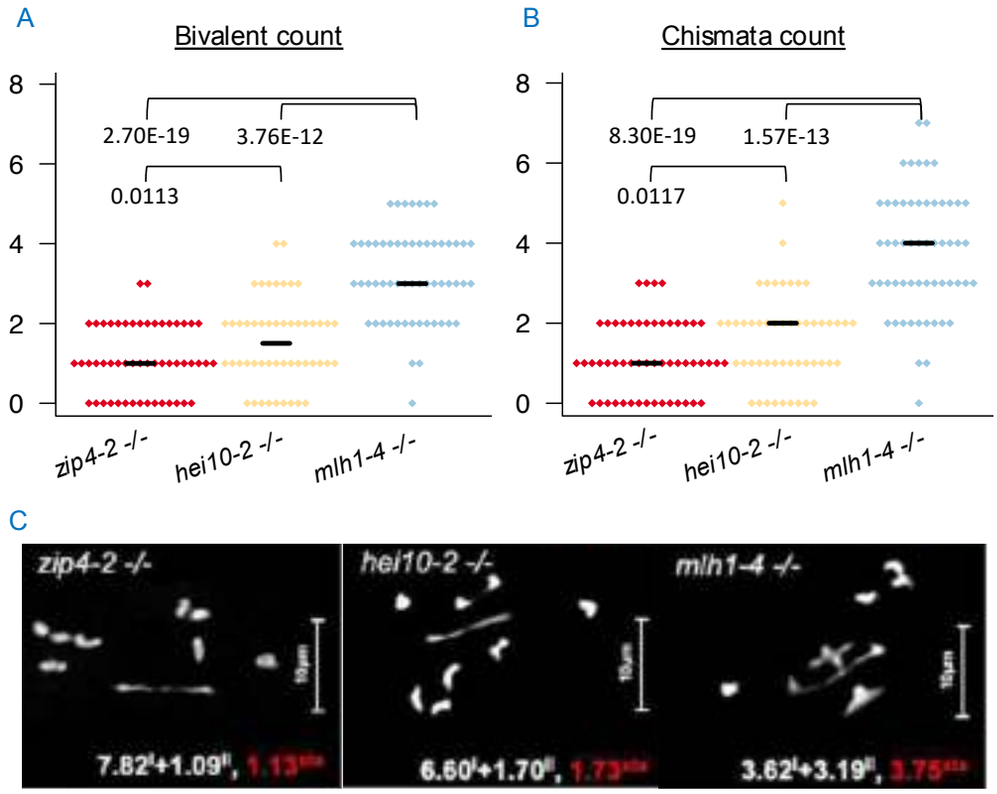


Figure 18. Cytogenetic characterization of *mlh1-4 -/-*. (A) The number of paired chromosomes observed per cell. (B) The number of chiasmata counted per cell. (C) Representative cells for *mlh1-4 -/-* and the two chosen controls *zip4-2 -/-* and *hei10-2 -/-*. T-test values, with a 5% accepted error, are represented between samples connected with brackets. scale bars are presented on the right side of the pictures, and the averaged observed behavior at the bottom. Superscript characters are read: "I" = univalent, "II" = bivalent, "xta" = chiasmata. n= 50, 53, and 57 cells in the same order on the plots.

3.3.1 *mlh1-4 -/-* shows decreased chiasmata count and loss of crossover assurance

Inflorescences from *mlh1-4 -/-* plants were collected 2h after the start of the day period. They were treated with Carnoy fixative for 24h and then conserved in 70% ethanol for 1 to 2 weeks before they were used for preparing chromosome spreads. Chromosome spreads were dyed with DAPI. Identified metaphases were then observed, interpreted, and pictured under 100X magnification. Only non-overlapping cells with a complete set of chromosomes (2n=10) were used for interpretation.

In this experiment, I used two cytogenetically characterized ZMM mutants, *zip4-2 -/-* and *hei10-2 -/-* as controls (Chelysheva et al., 2007; Chelysheva et al., 2012). Wildtype Col consistently shows 5 bivalents and about 9.2 chiasmata per cell, confirmed in the Ziółkowski lab conditions (Zhu et al., 2021). This choice was made as MutLγ is believed to be the main class I crossover resolvase. However, *mlh1* and *mlh3* null mutants do not show a severe recombination phenotype as one would expect since ZMM is responsible for about 90% of the crossovers in Arabidopsis. Consistently with their role as ZMM factors, and the published phenotypes, *zip4-2 -/-* shows 1.09 bivalents and 1.13 chiasmata per cell, and *hei10-2 -/-* shows 1.54 bivalents and 1.6 chiasmata per cell. Interestingly, *mlh1-4 -/-* shows significantly higher bivalent and chiasmata counts in comparison with *zip4-2 -/-* and *hei10-2 -/-*, 3.19 bivalents and 3.75 chiasmata per cell on average (Figure 18). This is consistent with the recombination frequency phenotype and reinforces the suggestion that MutLγ may be the main resolvase of the ZMM pathway but is probably not the only one. A similar cytogenetic phenotype was observed by Mónica Pradillo Lab at the Complutense University in Madrid for *mlh1-1 -/-*. Roughly 4 bivalents and 4 chiasmata on average per cell (Personal communications). Moreover, a similar phenotype was observed by Jackson et al., 2006 for *mlh3-1* with an average of 3.92 chiasmata per meiocyte.

3.3.2 Fertility assessment

Meiosis is a cell division specific to sexual reproduction, and so its success ultimately affects fertility. Here, I assess Arabidopsis fertility through four parameters: silique length (cm), seed per silique (count), pollen viability (percentage), and pollen density (qualitative).

3.3.2.1 *mlh1-1 -/-* and *mlh1-4 -/-* show similar fertility issues

The *mlh1-1* mutant is in a Ws background and so could not be used for recombination frequency scoring through FTLs. FTLs are only available in Col or L*er*

backgrounds. The recombination frequency of the cross between *mlh1-1* and FTLs would be a combined effect of the mutation and the heterozygosity between Ws and Col or L*er*. To use the *mlh1-1* allele as a control for the subsequent experiment, I first compared it to the *mlh1-4* allele. *mlh1-4* is the new CRISPR-Cas9 deletion allele generated in this study that shows the same crossover rate phenotype as the other two tested T-DNA insertion alleles *mlh1-2* and *mlh1-3* in the null state (Figure 16). Mature but green siliques at positions 7 to 12 of the main stem of n= 5-10 plants were collected and discolored in 96% ethanol. They were then photographed at a 6.4X magnification and measured using ImageJ. The number of seeds per silique was acquired from the same pictures. The Col and Ws controls show similar silique length and seed count, allowing for comparing *mlh1-1* and *mlh1-4* alleles without normalizing the measurements to their respective controls. *mlh1-1 -/-* and *mlh1-4 -/-* show a significant decrease in silique length, 30% and 45% respectively, and seed count, 65% and 71%, in comparison to their respective controls (Figure 19). *mlh1-4 -/-* shows significantly shorter siliques but a similar seed count.

For pollen viability, 5 open flowers were collected from n=3 plants. Their pollen was colored with Alexander staining and photographed with a 10X magnification. Round pink pollen grains were counted as viable, and green/brown deflated pollen grains as non-viable. *mlh1-1 -/-* and *mlh1-4 -/-* show the same significantly decreased viability, of about 40%, compared to the controls (Figure 19C).

Finally, for pollen density, two stage 12 flower buds were collected from three plants for each genotype. They were discolored using Carnoy fixative for 1h then colored with Alexander staining for 7h at 55 ºC. Three anthers were collected from each flower bud and photographed at 20X magnification. Representative pictures are presented in Figure 20. Loss of function mutation of *MLH1* in the Ws background shows a notable decrease in anther size in comparison to its WT control. This is not observed for the Col allele. Both *mlh1-1* and *mlh1-4* show a lower pollen density and dead pollen, with an apparently more severe effect for *mlh1-1* (Figure 19 and Figure 20).

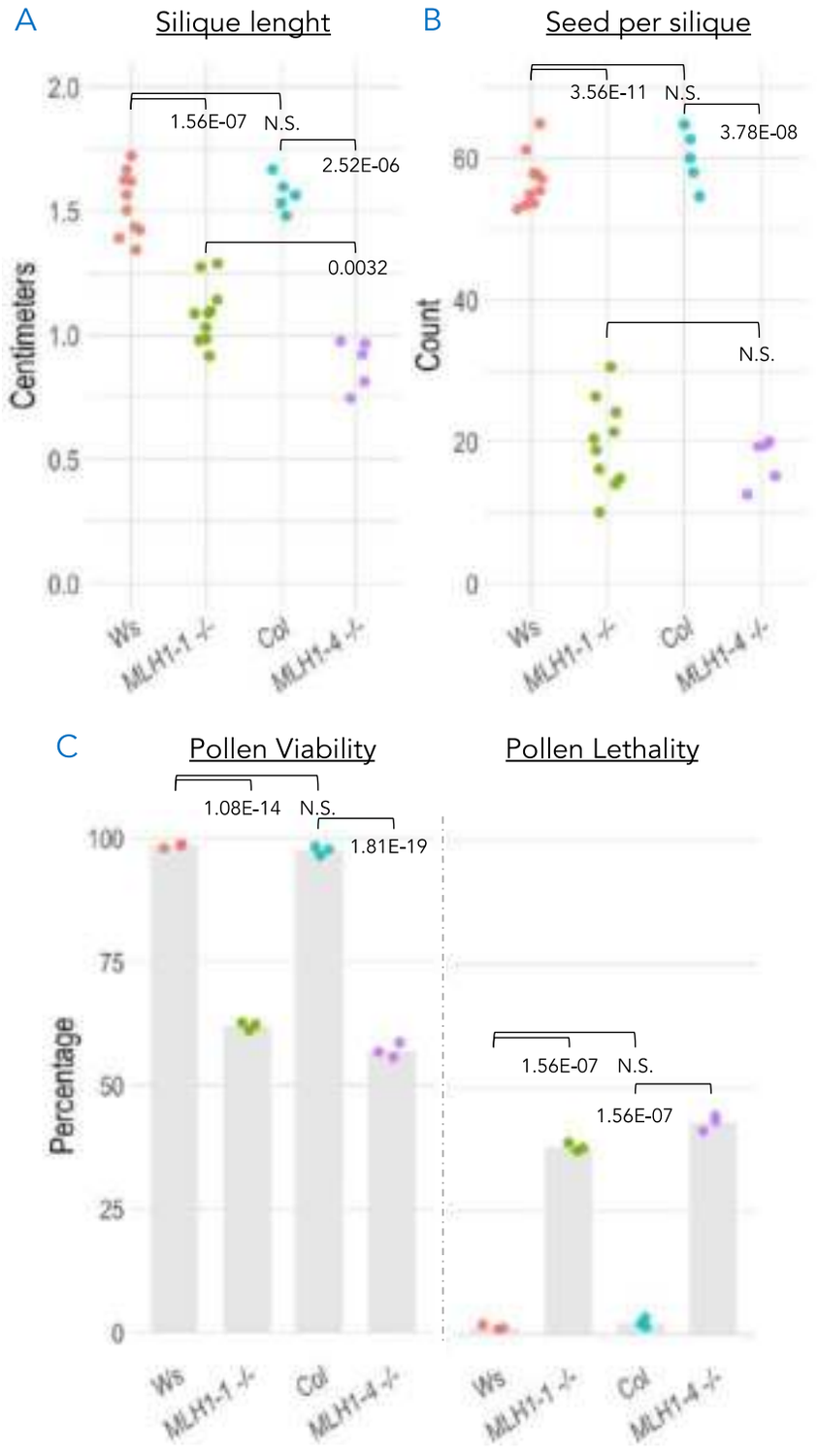


Figure 19. Seed set and pollen viability comparative assessment of *mlh1-1* vs *mlh1-4* alleles. (A) Silique length in centimeters. (B) Seed per silique count. T-test values, with a 5% accepted error, are represented between samples connected with brackets. N.S.= Not Significant, p >0.05. (C) Pollen viability and lethality in percentage. Chi-test values are represented.

Overall, *mlh1-1 -/-* and *mlh1-4 -/-* show similar effects on fertility in Arabidopsis. This means that the observed phenotypes are direct effects of *MLH1* loss of function and not potentially associated mutations.

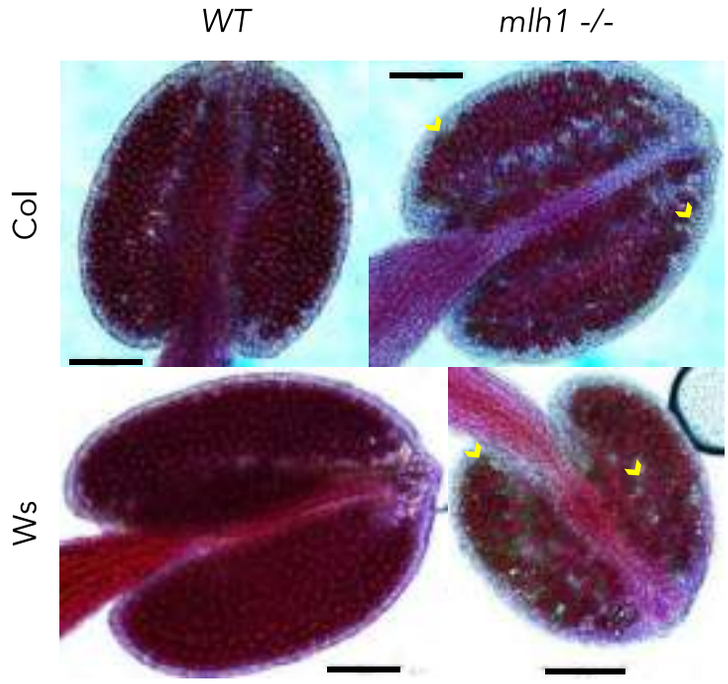


Figure 20. Representative pictures for pollen density assessment of *mlh1-4* and *mlh1-1* null mutants. *mlh1-4* is a CRISPR-cas9 mediated deletion mutant in a Col background, upper right panel. *mlh1-1* is a T-DNA insertion mutant in a Ws background, lower right panel. Their respective WTs are presented on the left panels. The scale bar represents 100 µm, and yellow arrowheads point at dead pollen.

3.3.2.2 Loss of function of both MutLγ subunits adversely affects fertility

The same fertility parameters were assessed for loss of function mutants of *MLH3, MUS81, FANCM,* and their double and triple mutant combinations. The wildtype and *mlh1-1* data presented here is the same as the data used for comparing *mlh1-1* to *mlh1-4.*

The *mlh1*, *mlh3,* and *mus81* null single mutants show a significant decrease in seed per silique count (Figure 21). Respectively, they display 65%, 66%, and 32% decrease. The *fancm* null shows about 5% non-significant decrease. *mlh1-1 mlh3-1* double null mutant shows a similar seed set to the two *mlh1-1* and *mlh3-1* single null mutants with an average of 73% decrease compared to wildtype. The triple

mlh1-1 mlh3-1 mus81 shows a significant 92% decrease in seed count. It is also lower than *mlh1-1 mlh3-1* double null. This is consistent with the loss of class I and MUS81-dependent class II crossovers. Finally, the triple *mlh1-1 mlh3-1 fancm* null mutant shows a very variable seed set averaging a 40% decrease compared to wildtype controls. This is also consistent with the uninhibited class II crossovers in the *fancm* loss of function context (Figure 21A). Silique length in centimeters was also quantified for these lines (Supplemental figure 7). The obtained data is consistent with seed count.

As mentioned above, pollen viability was assessed using Alexander staining. *mlh1-1* and *mlh3-1* show a similar and significant 38% and 29% decrease, respectively. The double *mlh1 mlh3* null mutant shows a 58% decrease, which is more severe than the single mutants. Yet, it is considerably highly viable for plants that are deprived of class I crossovers. The *mus81* single mutant shows a 12% decrease in pollen viability that is not significant in comparison to wildtype. The triple null mutant *mlh1 mlh3 mus81* shows a 70% decreased viability. It is significantly lower than the double *mlh1 mlh3.* Yet again, 30% viable pollen is considerably high. Finally, *fancm* single null mutant shows a 10% non-significant decrease. When combined with *mlh1 mlh3* double, *mlh1 mlh3 fancm* triple mutant shows a recovery in pollen viability with a 35% loss (Figure 21B and Supplemental figure 7). Chi-square test values are presented in Supplemental table 1.

Overall, fertility assays show consistent behaviors in all tested parameters, seed set, pollen viability, and pollen density (Supplemental figure 6).

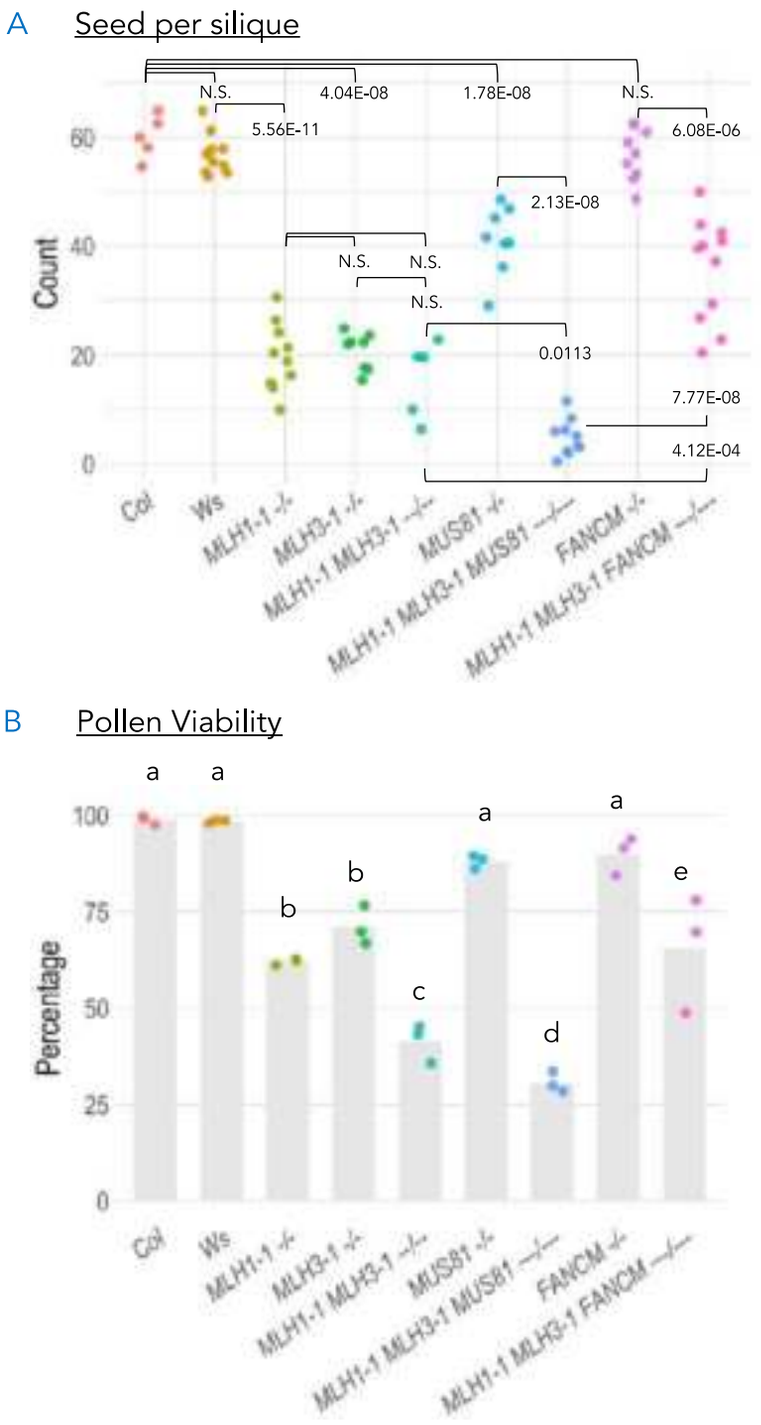


Figure 21. Seed set and pollen viability comparative assessment of *MLH1-1, MLH3-1, MUS81, FANCM,* and their combined double and multiple mutants. (A) Seed per silique count. T-test values, with a 5% accepted error, are represented between samples connected with brackets. N.S.= Not Significant, p >0.05. n~10 (B) Pollen viability in percentage. Different letters denote statistically significant differences according to a Chi test ($p < 0.05$), n=3.

3.4 *MutL* genes' overexpression effect on meiosis in an inbred context

For the purpose of investigating the effect of additional copies of the *MutL* genes, *MLH1, MLH3,* and *PMS1* were cloned under the control of their respective native promotors or the meiosis-specific *DMC1* promotor (Supplemental figure 13 and Supplemental figure 15). They were then introduced into wildtype Arabidopsis harboring the 420 interval fluorescent tags in the hemizygous state (GR/++). T1 BASTA-resistant plants, which also acquired the additional copies of the different *MutL* genes, were grown to seed and pre-selected for hemizygosity for the 420 fluorescent tags. Plants segregating for both fluorescent tags were then imaged and used for scoring the recombination frequency in 420. This is to investigate the effect of the extra copies on meiotic crossover recombination.

3.4.1 *MutL* genes overexpression under their respective native promoters does not affect recombination frequency in the T1 generation

In comparison to the Col-420 control, none of the additional copies of any of the three MutL genes affected recombination frequency in the 420 interval, subtelomeric region of the north arm of chromosome 3. The control shows an average RF of 21.2 cM, *pMLH1::MLH1* 22.55 cM, *pMLH3::MLH3* 21.74 cM, and *pPMS1::PMS1* 20.3 cM (Figure 22A).

However, some T1 *pMLH3::MLH3* individuals show interesting outlying values of 10 and 33 cM. These extreme RFs could be due to the number of copies integrated. Additionally, the overexpression of *PMS1* seems to be slightly but significantly lower than those of *MLH1* and *MLH3*.

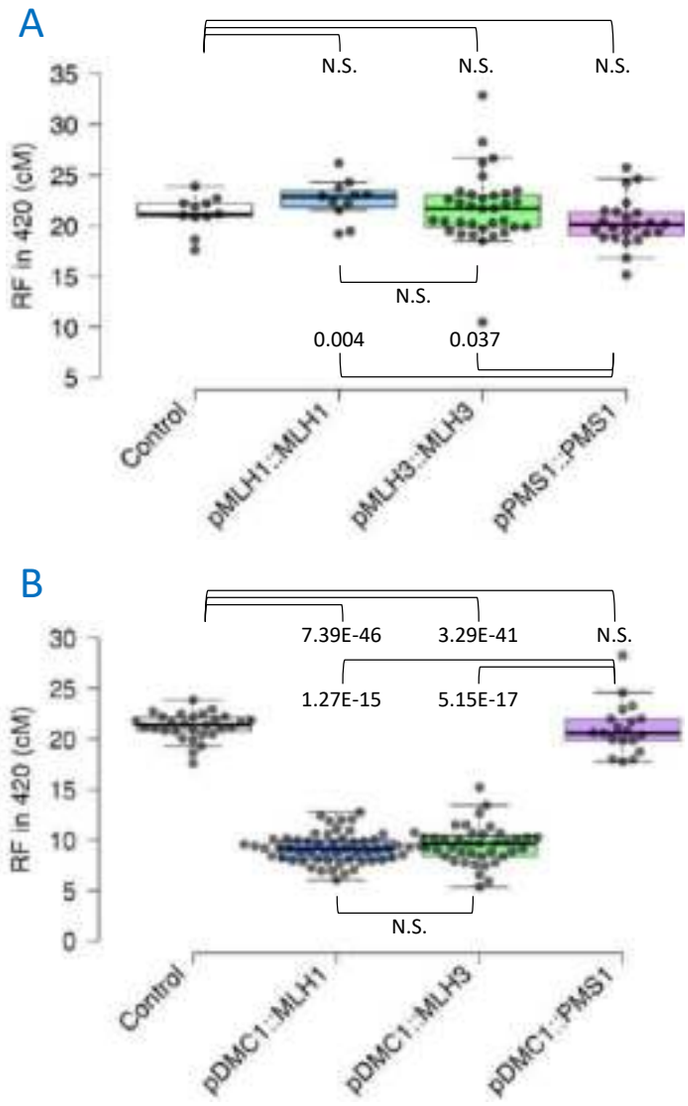


Figure 22. Recombination frequency measurement for the different MutL overexpression T1 lines in the chromosome 3 subtelomeric 420 interval. A. *MLH1, MLH3,* and *PMS1* under their respective native promoters. B. *MLH1, MLH3*, and *PMS1* under the control of the meiosis-specific *DMC1* promoter *(pDMC1).* One-tail T-test values, with a 5% accepted error, are represented between samples connected with brackets. N.S.= Not Significant, p >0.05.

3.4.2 *MLH1* and *MLH3* overexpression under the control of *DMC1* promoter affect recombination frequency in the T1 generation

Compared to the Col-420 control, lines with additional copies of *MLH1* and *MLH3* under the control of *DMC1* promoter (*pDMC1*) show a dramatic, and more interestingly similar, decrease in recombination frequency in the 420 interval. Where the control has an average RF of 21.25 cM, *pDMC1::MLH1* and *pDMC1::MLH3* respectively show an average of 9.18 cM and 9.53 cM. On the other

hand, *pDMC1::PMS1* shows a wildtype-like RF, of 21.07 cM (Figure 22B). This can be due to different possible reasons: 1) *DMC1* expression is 31, 105, and 36 times more active than *MLH1, MLH3*, and *PMS1* respectively, as shown by (Walker et al., 2017, Supplemental figure 2). MLH1 and MLH3 form the main class I crossover resolvase, the strong overexpression of their genes could affect their activity and so meiotic crossover recombination. 2) *DMC1* intervenes earlier during meiotic recombination, the untimely expression of *MLH1* and *MLH3* could hinder their activity. PMS1 does not seem to affect meiotic crossover recombination and so its overexpression does not have an effect in the tested interval.

3.4.3 *MLH1* and *MLH3* overexpression under their native promoters can increase meiotic crossover frequency in fillial generations

Five independent T2 lines were selected from *pMLH1::MLH1 (MLH1 Oe)* and *pMLH3::MLH3 (MLH3 Oe)* progenies. They were chosen for representing colder, hotter, and wildtype-like recombination frequencies. Seeds with hemizygous fluorescent tags were preselected and grown to seed. The obtained seeds were used for measuring recombination in the given interval. Crossover recombination frequency was measured for all five lines in 420 and for two lines for each of *pMLH1::MLH1* and *pMLH3::MLH3* in 3.9.

In the 420 subtelomeric interval, apart from *MLH1 Oe# 7*, which shows a low significance, none of the tested *MLH1 Oe* lines were hotter than wild type (Figure 23A). *MLH1 Oe#* 7 has an average RF = 22.29 cM *vs* 20.04 cM for the control Col-420. In the 3.9 pericentromeric interval, *MLH1 Oe# 4* shows a wildtype-like RF whereas *MLH1 Oe# 12* shows a significant increase in RF (Figure 23A). *MLH1 Oe# 12* has an average RF = 21.19 cM *vs* 19.6 cM for the control Col-3.9.

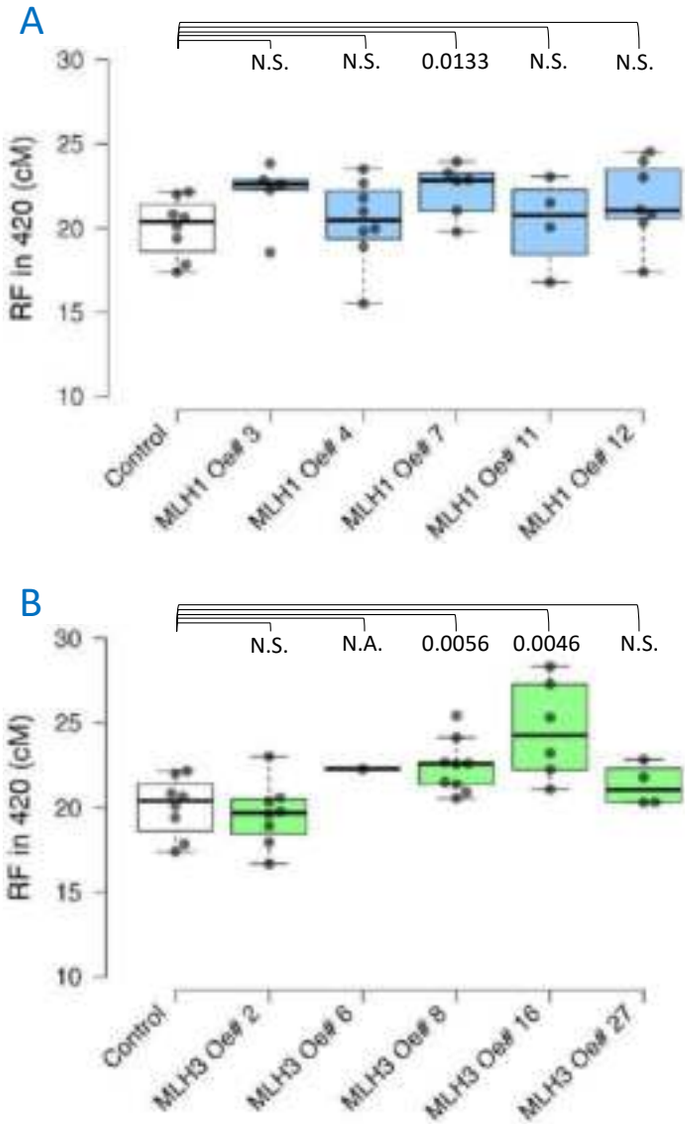


Figure 23. Recombination frequency measurement for five independent T2s of *MLH1* and *MLH3* overexpression lines under their native promoters in 420 interval. A. Five *MLH1* overexpression lines (*MLH1 Oe*) under the native promoter. B. Five *MLH3* overexpression lines (*MLH3 Oe*) under the native promoter. T-test values, with a 5% accepted error, are represented between samples connected with brackets. N.S.= Not Significant, p >0.05. N.A.= Not Applicable, no statistical test was performed because of the low number of data points.

Out of the five tested T2 lines for *MLH3 Oe* in the 420 interval, two lines *MLH3 Oe# 8* and *MLH3 Oe# 16* show significantly hotter RFs. The remaining three *MLH3 Oe* lines show wildtype-like RFs (Figure 23B). *MLH3 Oe#* 8 and *MLH3 Oe# 16* respectively have average RFs of 22.41 cM and 24.58 cM *vs* 20.04 cM for the control Col-420. In the 3.9 pericentromeric interval *MLH3 Oe# 16* shows a significantly

hotter RF whereas *MLH3 Oe# 8* shows a wildtype-like RF (Figure 24B). *MLH3 Oe# 16* has an average RF = 24.55 cM *vs* 18.85 cM for the control Col-3.9.

The expression levels of *MLH1* and *MLH3* were quantified for the ten selected T2 lines. For *MLH1 Oe* lines all five lines showed significantly higher expression levels of *MLH1*. *MLH1* overexpression did not show any correlation between expression level and recombination frequency. *MLH1 Oe# 7* and *MLH1 Oe# 12* show similar expression levels yet their effects are different in the two tested intervals (Supplemental figure 3). For the five *MLH3 Oe* T2 lines, only three of them show significantly higher expression levels. *MLH3 Oe# 8* and *MLH3 Oe# 16* are two of these three lines. Interestingly, *MLH3* expression level correlates positively with RF, r^2 = 0.78. Moreover, *MLH3 Oe# 16* the line with the hottest RF is also the line with the highest expression level of MLH3 (Supplemental figure 4).

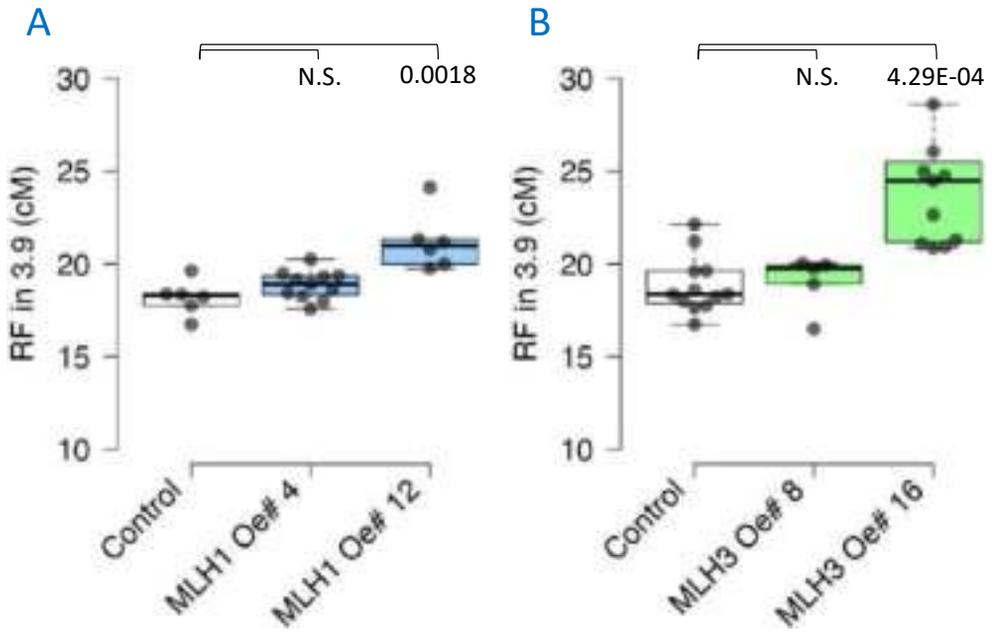


Figure 24. Recombination frequency measurement for two independent T2s of *MLH1* and *MLH3* overexpression lines under their native promoters in the 3.9 interval. A. Two *MLH1* overexpression lines (*MLH1 Oe*) under the native promoter. B. Two *MLH3* overexpression lines (*MLH3 Oe*) under the native promoter. T-test values, with a 5% accepted error, are represented between samples connected with brackets. N.S.= Not Significant, p >0.05.

3.4.4 *MLH1* and *MLH3* overexpression under the control of *DMC1* promoter in T2 generation exhibits the same phenotype as the T1

Two independent T2s for *MLH1* and *MLH3* overexpression under the control of *DMC1* were selected to quantify RF in 420. Three independent T2s were selected for scoring RF in 3.9.

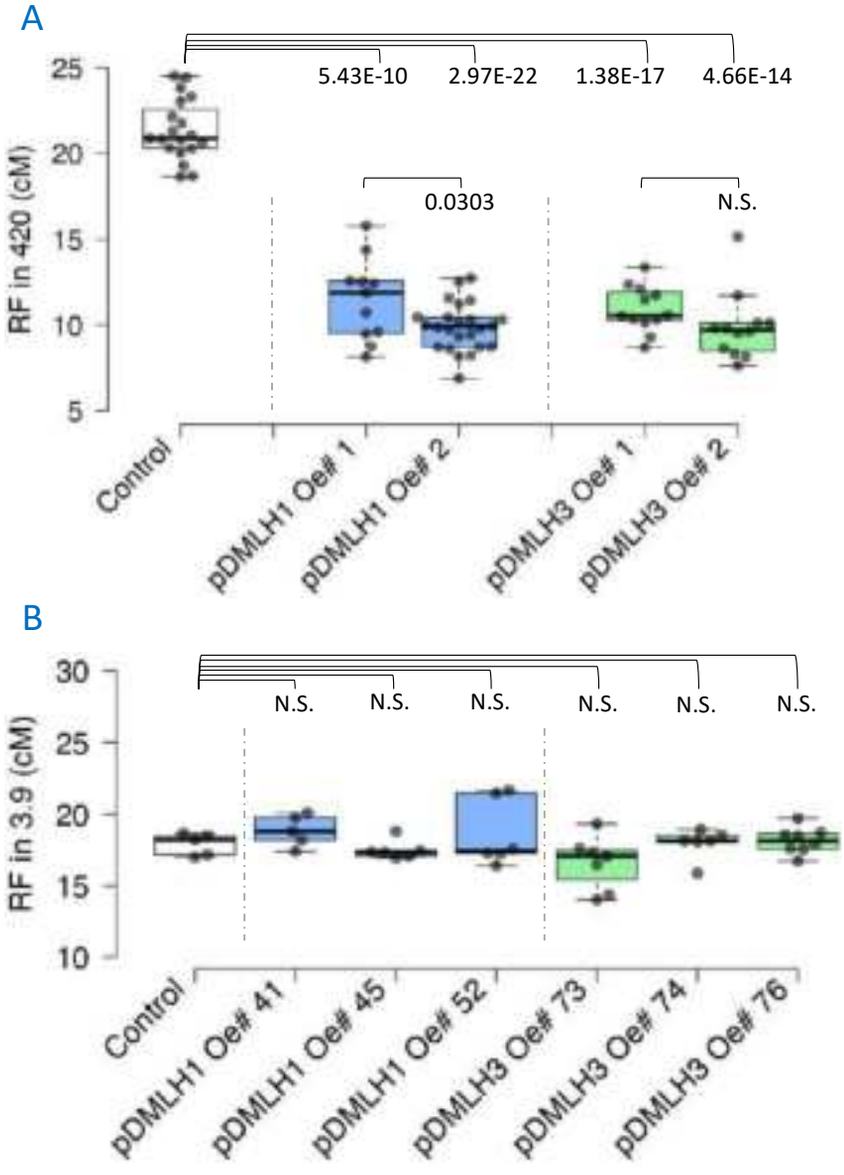


Figure 25. Recombination frequency measurement for three independent T2s of *MLH1* and *MLH3* overexpression lines under *DMC1* promotor in 420 and 3.9 intervals. A. Recombination frequency measurement in the subtelomeric region of chromosome 3, the 420 interval. B. Recombination frequency measurement in the pericentromeric region of chromosome 3, the 3.9 interval. T-test values, with a 5% accepted error, are represented between samples connected with brackets. N.S.= Not Significant, $p > 0.05$.

Both *pDMC1::MLH1 Oe (pDMLH1 Oe)* and *pDMC1::MLH3 (pDMLH3 Oe)* show a similar RF in 420 to their parental lines. *pDMLH1 Oe* average RFs of 11.48 cM and 9.89 cM, and *pDMLH3 Oe* average RFs of 10.92 cM and 9.89 cM. These crossover rates represent 46% to 53% of that of the control Col-420 which averages 21.36 cM (Figure 25A).

On the contrary, neither *pDMLH1 Oe* nor *pDMLH3 Oe* tested lines show any significant differences in RF with the control line Col-3.9 which averages an RF of 17.9 cM (Figure 25B). However, it is interesting to note that while the crossover rate was significantly decreased for the *MLH1* and *MLH3* overexpressor lines when compared to wild-type controls in subtelomeric 420 interval, they were not changed in the subtelomeric 3.9 interval.

3.4.5 *MLH1* and *MLH3* double overexpression affects crossover rate

MLH1 and *MLH3* overexpressors, under their respective native promoters and under the control of *DMC1* promoter were crossed with each other. The obtained F1s were selected with BASTA treatment and genotyped for the constructs. Very few plants with both *MLH1* and *MLH3* overexpression constructs were obtained and propagated to the F2 generation. Plants were again treated with BASTA and genotyped for the constructs. The plants that were positive for both *MLH1* and *MLH3* constructs were grown to seed, and RF was scored in the 420 interval.

MLH1/MLH3 double overexpressor (*MLH1/MLH3 dbl Oe*) under their native promoters shows a significantly higher RF than the Col-420 control with 23.94 cM *vs* 22.11 cM respectively. On the other hand, *DMLH1/DMLH3* double overexpressor (*DMLH1/DMLH3 dbl Oe*) under *DMC1* promoter shows a significantly lower 11.97 cM RF (Figure 26), though slightly higher than single overexpressor under *DMC1* promoter. *MLH1/MLH3* double overexpressors under the different promoters show similar phenotypes to the respective T2s they were generated from.

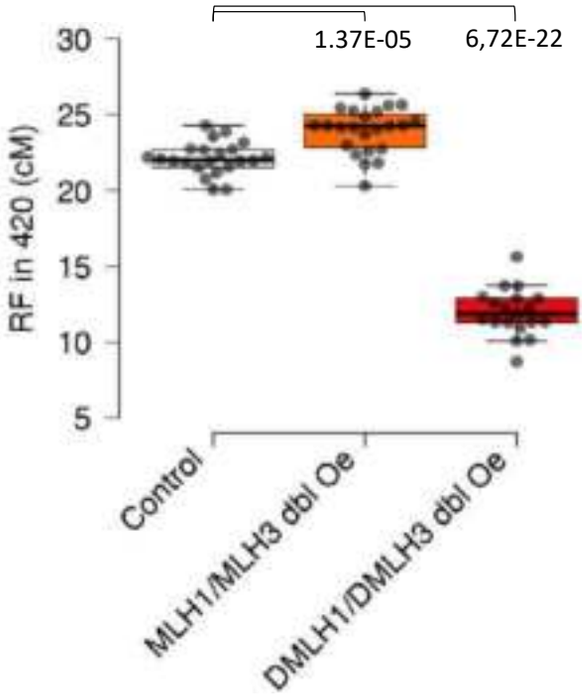


Figure 26. Recombination frequency measurement for *MLH1* and *MLH3 double* overexpression. Recombination frequency measurement in the 420 subtelomeric interval of chromosome 3 for *MLH1* and *MLH3* double overexpressor lines under their respective promoters and under the control of DMC1 promotor in 420 interval. T-test values, with a 5% accepted error, are represented between samples connected with brackets.

3.5 *MutL* genes overexpression effect on meiosis in a Col/L*er* hybrid context

MutL genes are part of the mismatch repair (MMR) family genes. These code for the proteins responsible for recognizing and correcting mismatches that accrued during DNA replication. Moreover, genetic heterozygosity is known to affect meiotic crossover recombination in Arabidopsis, with a local boost of recombination in heterozygous regions at the expense of neighboring homozygous regions, in an MMR-dependent manner (). Considering these factors, I sought to investigate how the overexpression of MutL genes would affect recombination in the Col/L*er* hybrid context.

The previously selected independent T2 lines, for *MLH1*, *MLH3,* and *PMS1* under the control of their respective native promoters or under the control of *DMC1* promoter, were crossed to Col and L*er*. Using an inbred cross to Col as a control was important to account for the reduction of the number of overexpression transgenes. The obtained F1s were preselected with BASTA, grown to seed, and scored for their crossover recombination frequency in specific intervals.

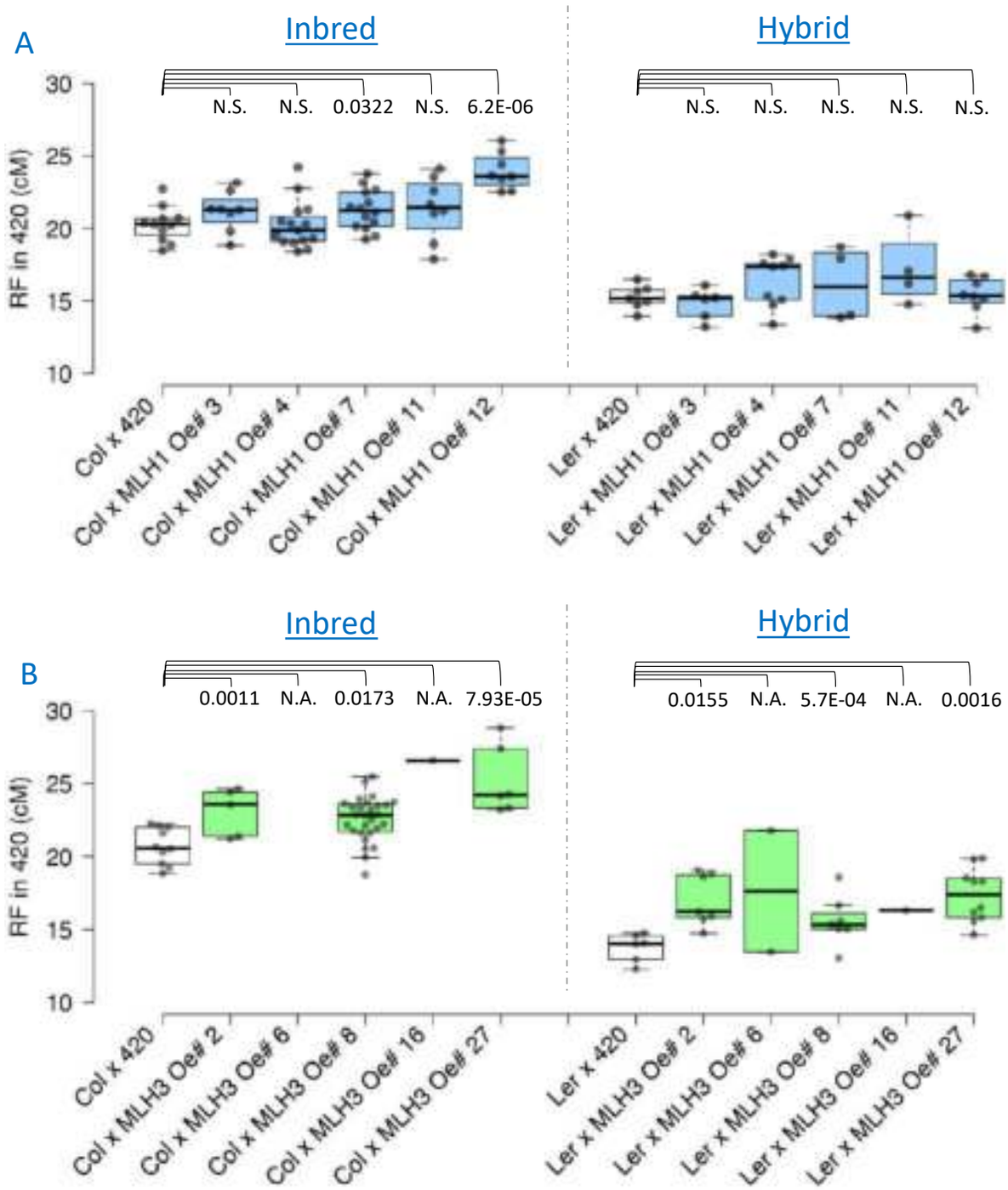


Figure 27. Hybrid context recombination frequency measurement for five independent lines of *MLH1* and *MLH3* overexpression lines under their native promoters in the 420 interval. A. Recombination frequency measurement of inbred Col/Col F1 *vs* hybrid Col/Ler *pMLH1::MLH1 (MLH1 Oe)* compared to the construct-free control. B. Recombination frequency measurement of inbred Col/Col F1 *vs* hybrid Col/Ler *pMLH3::MLH3 (MLH3 Oe)* compared to the construct-free control. T-test values, with a 5% accepted error, are represented between samples connected with brackets. N.S.= Not Significant, p >0.05. N.A.= Not Applicable, no statistical test was performed because of the low number of data points.

3.5.1 *MLH1* and *MLH3* overexpression under their respective native promoters

In inbred F1 crosses of *MLH1 Oe* lines, two lines, *MLH1 Oe# 7* and *12*, show a significantly higher RF in 420. They show an average of 21.29 cM and 23.95 cM compared to 20.28 cM for the wildtype Col x Col-420. In contrast, all the *MLH1 Oe* show an RF level similar to the hybrid L*er* x Col-420 (Figure 27A).

As for *MLH3 Oe* lines, not enough data could be collected for two lines, *MLH3 Oe# 6* and *16*. This was due to the fact that very few plants were resistant to BASTA selection, suggesting a counterselection of the transgene or silencing. The obtained data points were plotted to indicate trends.

The three remaining lines, *MLH3 Oe# 2, 8,* and *27,* show significantly higher RF in 420 in both inbred and hybrid contexts. Inbred *MLH3 Oe# 2, 8,* and *27* have average RFs of 23.05 cM, 22.58 cM, and 25.2 cM respectively compared to 20.7 cM in the control Col x Col-420 control. In hybrid, they display average RFs of 17.02 cM, 15.61 cM, and 17.35 cM compared to 13.77 cM for the control L*er* x Col-420 control (Figure 27B). By contrast, RF in 3.9 does not show a significant difference with the 21.17 cM control for *MLH1 Oe# 12* and *MLH3 Oe# 16* in the hybrid context (Figure 28).

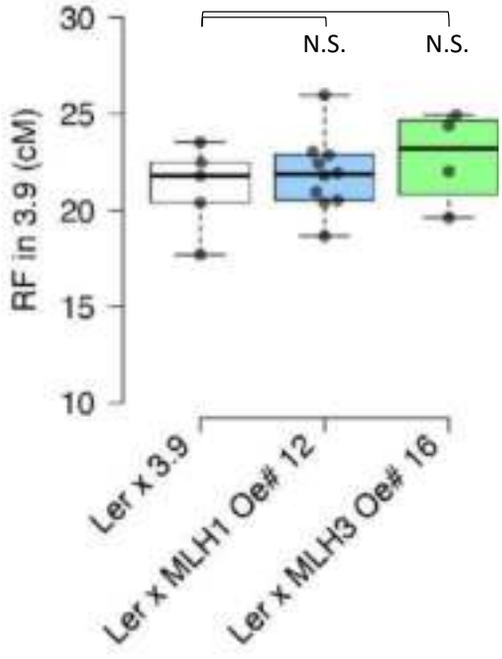


Figure 28. Hybrid context recombination frequency measurement of *MLH1* and *MLH3* overexpression lines under their native promoters in the 3.9 interval. Recombination frequency measurement in the 3.9 pericentromeric interval of chromosome 3. T-test values, with a 5% accepted error, are represented between samples connected with brackets. N.S.= Not Significant, $p > 0.05$.

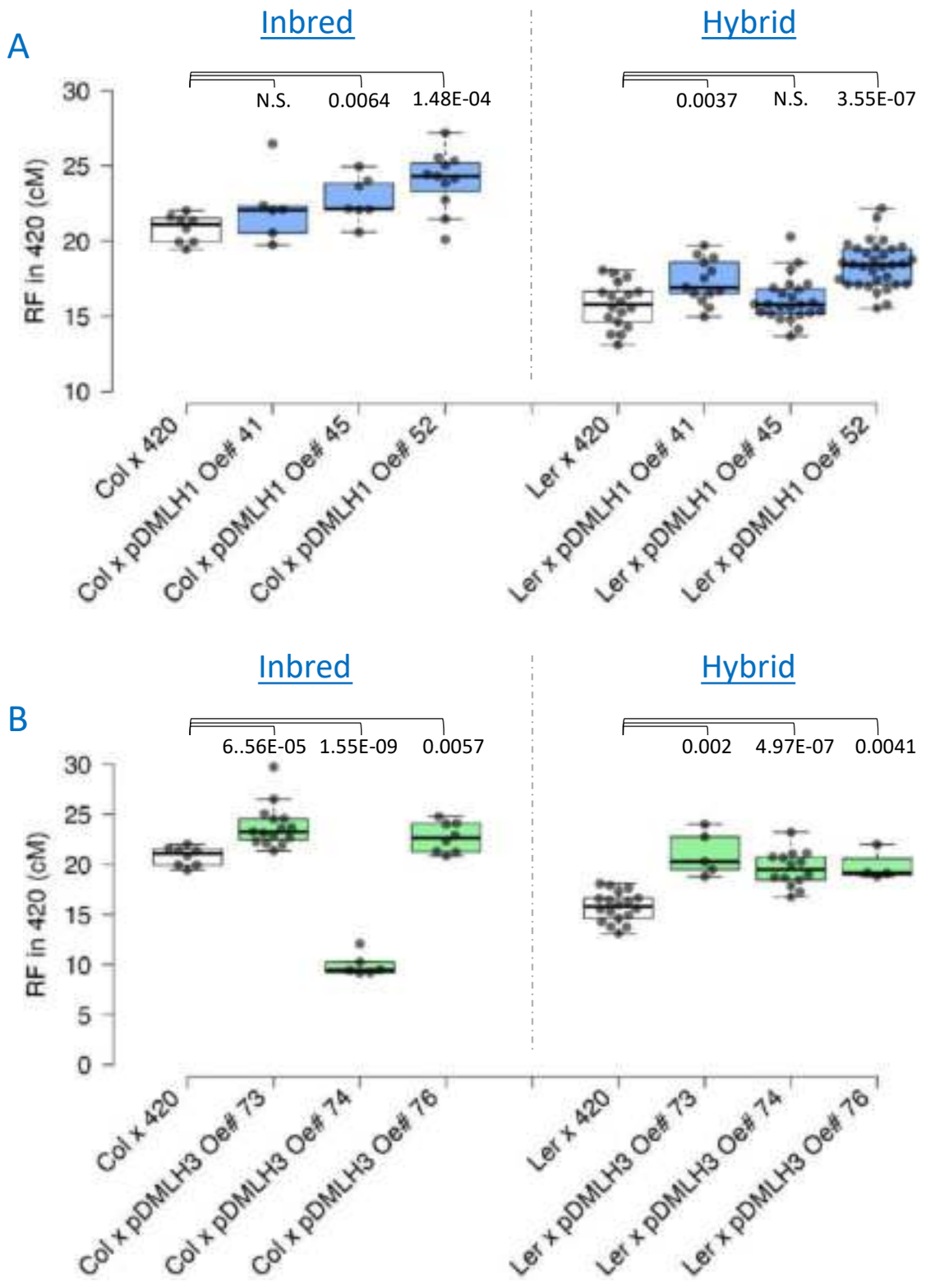


Figure 29. Hybrid context recombination frequency measurement for three independent lines of *MLH1* and *MLH3* overexpression lines under *DMC1* promotor in the 420 interval. A. Recombination frequency measurement of inbred Col/Col F1 *vs* hybrid Col/Ler *pDMC1::MLH1 (pDMLH1 Oe)* compared to the construct-free control. B. Recombination frequency measurement of inbred Col/Col F1 *vs* hybrid Col/Ler *pDMC1::MLH3 (pDMLH3 Oe)* compared to the construct-free control. T-test values, with a 5% accepted error, are represented between samples connected with brackets. N.S.= Not Significant, $p > 0.05$.

3.5.2 *MLH1* and *MLH3* overexpression under the control of *DMC1* promoter

MLH1 and *MLH3* overexpressors under the *DMC1* promoter showed a strong decrease in crossover rate in the 420 interval (Figure 22B and Figure 25A). The first interesting observation is that when the progeny of T1 plants were crossed to Col (inbred) or L*er* (hybrid), the crossover rate increased at least to the level observed in the control plants The effect of strong crossover reduction observed in T1 and T2 generations disappeared in the F1 crosses. In the inbred context, all but *MLH3 Oe# 74* show similar or higher average RF in 420 in comparison to the controls for both *pDMC1::MLH1 Oe* and *pDMC1::MLH3 Oe* (Figure 29). In the hybrid context, two out of three *pDMC1::MLH1 Oe* and all three *pDMC1::MLH3 Oe* lines show significantly higher RFs.

This observation supports the hypothesis that the expression level of *MLH1* and *MLH3* in the overexpressor lines under *DMC1* promoter in T1 and T2 generations, reached a level that became detrimental to recombination. Crossing overexpressor lines to wild-type plants lead to a reduction in *MLH1* or *MLH3* expression allowing for efficient recombination. In both inbred and hybrid, the consensus is that *MLH1* and *MLH3* overexpression, within a tolerable expression level, can boost crossover recombination frequency in the 420 subtelomeric interval and pericentromeric 3.9 interval.

3.6 *MLH1* and *MLH3* overexpression is detrimental to Arabidopsis fertility

Fertility assessment for *MLH1* and *MLH3* overexpressor lines, both under the control of their respective promotors (*MLH Oe*) and *DMC1* promoter (*DMLH Oe*), was performed. Biological material was sampled from T2 plants. Discolored siliques were used for quantifying silique length (Supplemental figure 8 and Supplemental figure 9), and the number of seeds per silique. Open flowers were used to assess pollen viability through Alexander staining.

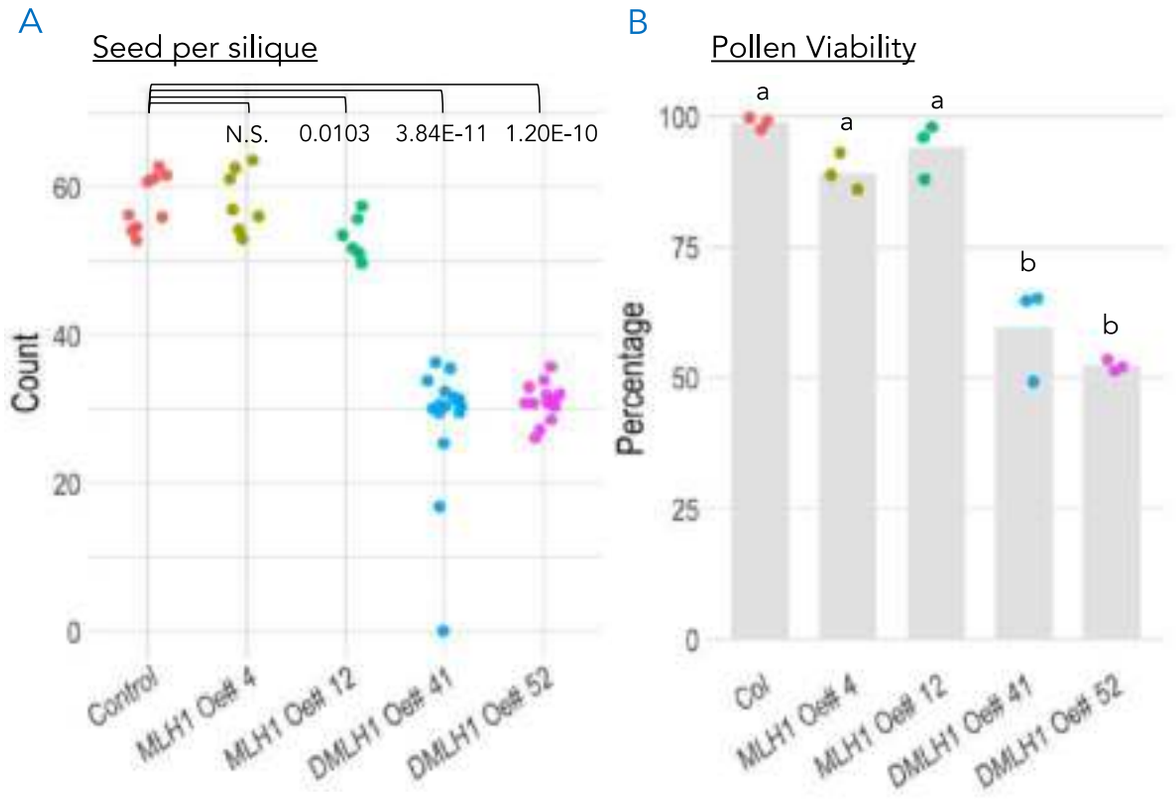


Figure 30. Seed set and pollen viability comparative assessment of *MLH1* overexpressor lines. Two native promoter overexpression lines (*MLH1 Oe*) and two *DMC1* promotor overexpression lines (*DMLH1 Oe*). (A) Seed per silique count. T-test values, with a 5% accepted error, are represented between samples connected with brackets. N.S.= Not Significant, p >0.05. n~10 (B) Pollen viability in percentage. Different letters denote statistically significant differences according to a Chi test ($p < 0.05$), n=3.

MLH1 Oe# 4 shows a seed set similar to the Col-420 control with 58.2 and 57.8 seeds per silique respectively. It however shows a small but not significant decrease in pollen viability, 89% and 99% respectively (Figure 31). *MLH1 Oe# 12* shows a slight but significant decrease in seed set, averaging 53.23 seeds per silique. Pollen viability is also lower but not significant, 94% (Figure 31). Both these overexpressor lines show however a significant increase in the proportion of dead pollen (Supplemental figure 8 and Supplemental figure 9).

DMLH1 Oe#41 and *52* both show a similar phenotype with strong decreases in both seed set and pollen viability. They respectively show an average of 28.46 and 30.98 seeds per silique, and 60% and 52% pollen viability.

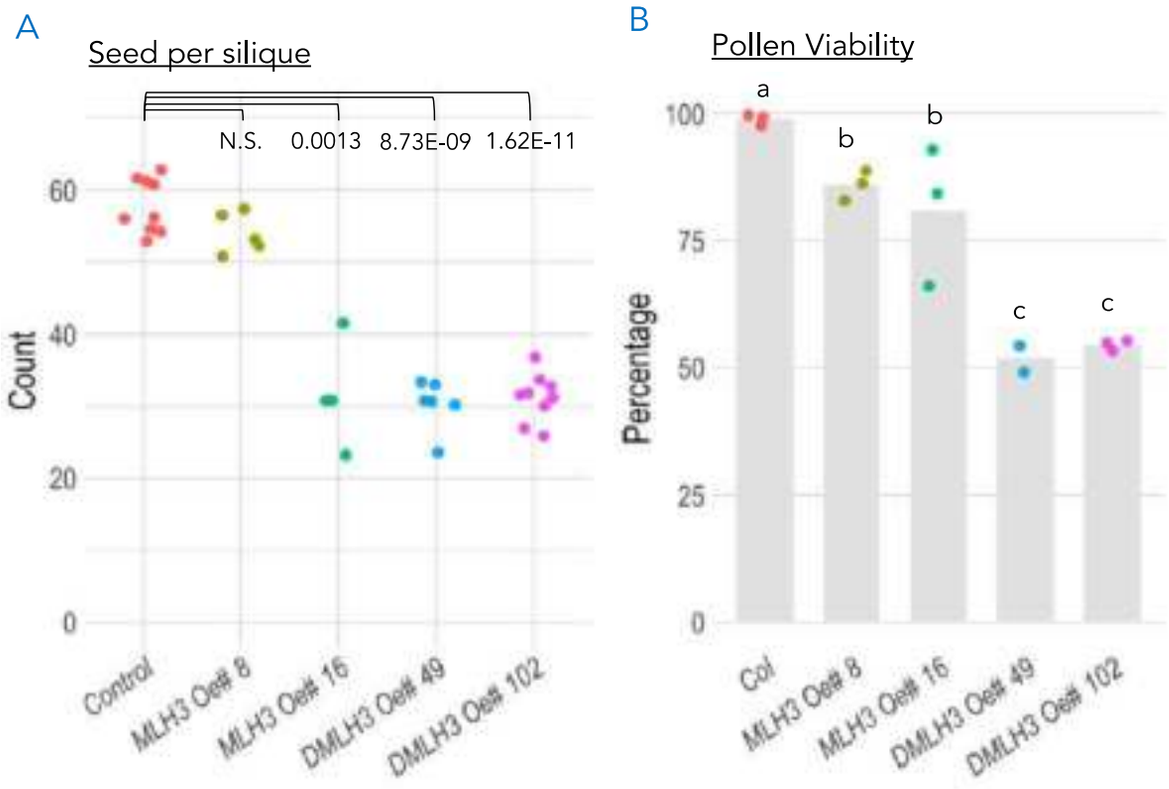


Figure 31. Seed set and pollen viability comparative assessment of *MLH3* overexpressor lines. Two native promoter overexpression lines (*MLH3 Oe*) and two *DMC1* promotor overexpression lines (*DMLH3 Oe*). (A) Seed per silique count. T-test values, with a 5% accepted error, are represented between samples connected with brackets. N.S.= Not Significant, *p* >0.05. n~10 (B) Pollen viability in percentage. Different letters denote statistically significant differences according to a Chi test ($p < 0.05$), n=3.

MLH3 Oe# 8 shows a seed set similar to the Col-420 control with 54.02 seeds per silique respectively. It also shows a slight but significant decrease in pollen viability, 86%. *MLH3 Oe# 16* shows a significant decrease in seed set with an average of 31.6 seeds per silique. Pollen viability is also significantly lower with 81% viability.

DMLH1 Oe#49 and *102* both show a similar phenotype with strong decreases in both seed set and pollen viability. They respectively show an average of 30.26 and 31.17 seeds per silique, and 52% and 55% pollen viability.

3.7 MutLγ genetic interaction with the ZMM pathway

As mentioned previously, *mlh1* and *mlh3* null mutants RF phenotype show a considerable decrease, which is however not as severe as it would be expected for the main resolvase of class I crossovers. Compared to *zmm* null mutants this phenotype is rather mild. This strongly suggests that it is not the exclusive resolvase

used for interfering crossovers. To investigate the extent of the relationship between the ZMM pathway and MutLγ, I introduced the *hei10-2* null mutation into an *MLH1* and *MLH3* overexpression background by backcrossing.

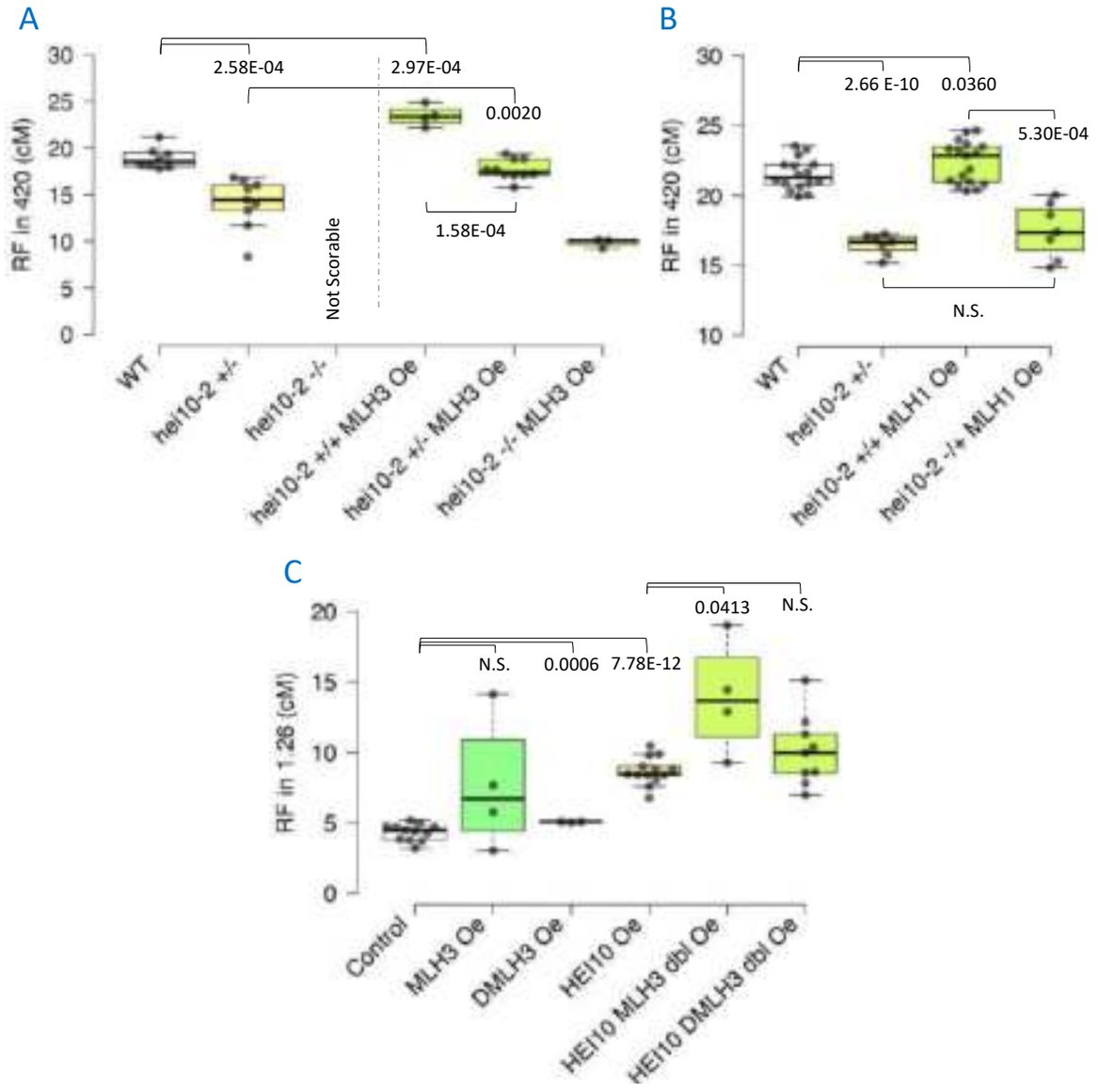


Figure 32. Recombination frequency measurement of an *MLH3* and *MLH1* overexpressors in combination with *hei10-2 -/-* or *HEI10* overexpressor. A. RF in the subtelomeric 420 interval for *hei10-2* in combination with *MLH3 Oe.* **B.** RF in the subtelomeric 420 interval for *hei10-2* in combination with *MLH1 Oe.* **C.** RF in the short subtelomeric 1.26 interval for *HE10) Oe* in combination with *MLH3 Oe.* T-test values, with a 5% accepted error, are represented between samples connected with brackets. N.S.= Not Significant, $p > 0.05$.

3.7.1 *MLH3* and *HEI10* seem to be able to operate independently

As for all *zmm* null mutants, the homozygous *hei10-/-* is not scorable using the seed based FTL technology. The *hei10-2 +/-* shows a significant decrease of RF with

an average of 14.08 cM *vs* 18.89 cM for the wildtype 420. *MLH3 Oe* shows a significant increase with an average RF of 23.44 cM. Interestingly, *hei10-2 -/- MLH3 Oe* is scorable and shows an average of 9.8 cM and the *hei10-2 +/- MLH3 Oe* is significantly hotter than *hei10-2 +/-,* with an average RF of 17.67 cM (Figure 32A). To the contrary, *MLH1 Oe* does show the same crossover recombination phenotype as *hei10-2 +/-*. *hei10-2 +/- MLH1 Oe* has an average RF of 17.5 cM *vs* 16.49 cM for the *hei10-2 +/-* grown in the same conditions (Figure 32B). This may suggest that MutLγ is able to form crossovers independently of HEI10. This preliminary data also suggests that the increase in RF in the *hei10-2* background is exclusive to the overexpression of *MLH3*, which is the only meiosis-specific MutL gene.

Additionally, I introduced both *MLH3 Oe* and *HEI10 Oe* constructs in the short subtelomeric CTL 1.26 interval (0.7 Mbp). The reason for selecting a very short interval was that the *HEI10* overexpression usually leads to very strong increases in crossover frequency and can reach up to 47 cM in the 420 interval (5.1 Mbp). Assuming an additionally increased crossover rate by *MLH3* overexpression would yield RF scores that are higher than 50 cM which is not reliable (data not shown). As expected, *HEI10 Oe* shows a higher RF than the control Col-1.26 with an average RF of 8.59 cM *vs* 4.32 cM respectively. *MLH3 Oe* and *DMLH3 Oe* show only slightly higher RFs with 7.65 cM and 5.06 cM, respectively. When combined, *HEI10 MLH3 double Oe* is significantly hotter than *HEI10 Oe* with a 13.9 cM average RF. *HEI10 DMLH3 double Oe* has an average of 10.09 cM but is not significant due to the high variability (Figure 32C). The overall trend of these measurements shows an additive effect between the overexpression of *HEI10* and *MLH3.*

3.7.2 *MLH3* overexpression seems to improve the *hei10-2 -/-* fertility phenotype

Preliminary assessment of the *hei10-2 -/- MLH3 Oe* fertility shows a small but significant improvement in Arabidopsis fertility. Qualitative assessment of anthers in comparison to wildtype Col and *hei10-2 -/-* shows a decrease in fertility in

comparison to wildtype but improved fertility in comparison to the mutant (Supplemental figure 11).

Moreover, an improvement in the seed set can be observed with more fertile plants (Supplemental figure 12). Indeed, the *hei10-2 -/- MLH3 Oe* shows an average seed set of 10.76 *vs* 5.08 seeds per silique for *hei10-2 -/-* (Figure 33). The improvement in fertility is not homogenous between plants of the same population. This is most probably due to the variability of *MLH3* overexpression in independent crosses and propagation of the plants.

To further investigate this phenotype, I plan to introduce other *zmm* mutations in the *MLH3 Oe* background (*zip4* and *msh4).* I would also like to further quantify the fertility phenotype in the *hei10-2 -/- MLH3 Oe*, conduct a cytogenetic confirmation, and confront the obtained results with the phenotype observed in *zip4 -/- MLH3 Oe* and *msh4 -/- MLH3 Oe*.

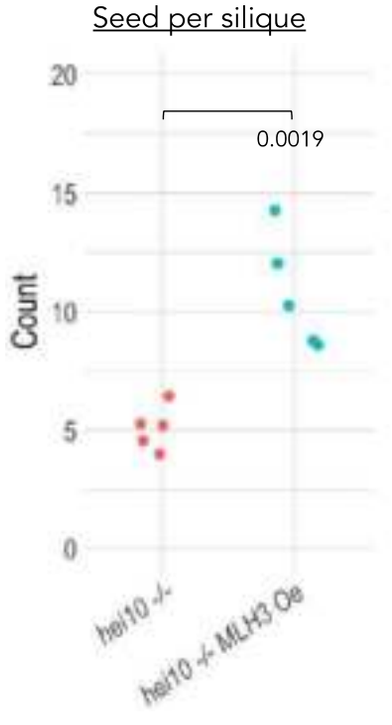


Figure 33. Seed set comparative assessment of *hei10 -/-* to *hei10 -/- MLH3 Oe*. T-test values, with a 5% accepted error, are represented between samples connected with brackets. N.S.= Not Significant, p >0.05.

3.8 Additional copies of a nuclease-dead *EXO1b* can boost the crossover rate

In addition to MLH1 and MLH3, the MutLγ complex has a regulatory add-on subunit, EXO1. EXO1 is a 5′-3′ exonuclease that is responsible for resecting double-

strand breaks for further processing. It has an additional, exonuclease-independent activity which function is to stimulate the nicking activity of MutLγ (Cannavo et al., 2020; Kulkarni et al., 2020). Arabidopsis has two homologs, EXO1a and EXO1b. Both *EXO1a* and *EXO1b* are expressed in both leaf and flower bud tissue. Recombination frequency measurement of the *exo1b* null mutant showed a similar rate to the wildtype control (Supplemental figure 17). The loss of function of *EXO1b* is most probably compensated by EXO1a and/or other exonucleases.

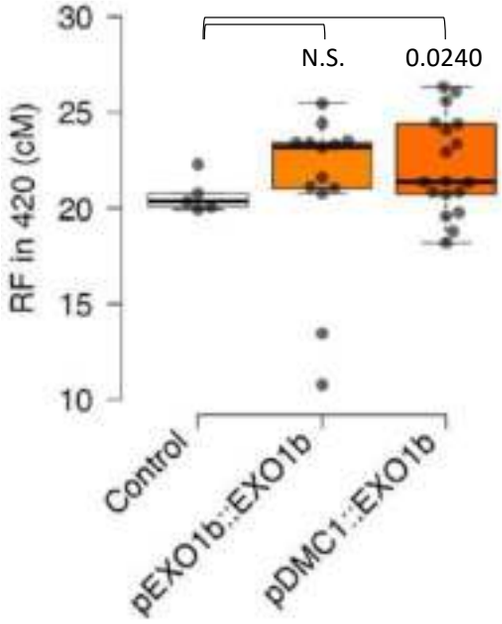


Figure 34. Recombination frequency measurement for EXO1b overexpression at the T1 generation. Recombination frequency measurement in the 420 subtelomeric interval of chromosome 3 for EXO1b under its endogenous promoter and under the control of DMC1 promotor. T-test values, with a 5% accepted error, are represented between samples connected with brackets. N.S.= Not Significant, p >0.05.

I further sought to introduce additional copies of a nuclease-dead *EXO1b*, *EXO1b DA*. I only used the nuclease-dead variant, because the overexpression of the functional gene is likely to affect early stages of DSB processing during resection (Tomimatsu et al., 2014; Mercier et al., 2015; Sanchez et al., 2020). *EXO1b* was cloned under the control of its native promoter and under the control of *DMC1* promoter. The exonuclease-dead variant was achieved through mutagenesis PCR by substituting the asparagine 180 (D) for alanine (A) in the active domain of EXO1 (…ITEDSDL… to …ITEDSAL…) (Wang et al., 2022). For the sake of simplicity, I will refer to the cloned copy as *EXO1b.* Similarly to the *MLH* overexpression lines, Col-

420 plants (GR/++) were transformed with the different constructs (Supplemental figure 16). The T0 plants were grown to seed, the seeds were sown and selected with BASTA. The surviving transformants were grown to seed and the plants with segregating fluorescent tags were used for scoring RF in the 420 interval.

At the T1 generation, both *pEXO1b::EXO1b* and *pDMC1::EXO1b,* hereafter *EXO1b Oe* and *DEXO1b Oe,* show variability in comparison to the Col-420 control with a trend for higher RF (Figure 34). *DEXO1b Oe* shows a significantly higher RF, 22.22 cM, than the control, 20.68 cM, but not *EXO1b Oe,* 21.19 cM (Figure 34).

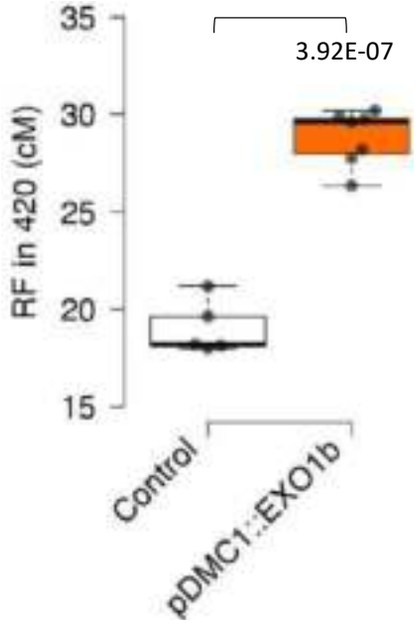


Figure 35. Recombination frequency measurement for *DEXO1b* overexpression at the T2 generation. Recombination frequency measurement in the 420 subtelomeric interval of chromosome 3 for *EXO1b* under the control of *DMC1* promotor. T-test values, with a 5% accepted error, are represented between samples connected with brackets. N.S.= Not Significant, $p > 0.05$.

At the T2 generation, *DEXO1b Oe* shows a significantly higher RF, 28.92 cM, when compared to the control, 19.02 cM (Figure 35). This result suggests that additional copies of *EXO1b* are sufficient for boosting RF in Arabidopsis in the tested subtelomeric interval. More interestingly, its effect is stronger than the effects of *MLH1* and *MLH3* overexpression, which may suggest that EXO1 is a limiting factor in Arabidopsis MutL complex activity.

4 Discussion

In this chapter, I investigated the effect of expression levels of *MutL* genes, *MLH1, MLH3,* and *PMS1,* on meiotic crossover recombination in Arabidopsis. I used null mutants in homozygous and heterozygous states (Figure 12 - Figure 17), and overexpressor lines with different levels of overexpression triggered by different promoters (Figure 22 - Figure 29). Additional copies were introduced either under their respective native promoters or under the control of the meiosis-specific *DMC1* promoter, which is between 30 and 100 folds more active than the MutL native promoters (Supplemental figure 2). For this characterization, I assessed crossover recombination frequency in different intervals and quantified fertility. For male fertility, pollen viability and pollen density were assessed. Female meiosis was inferred from seed set as pollen grains outnumber the number of eggs by far (~500 pollen grains/anther x 6 *vs* ~60 eggs/stigma in *A. thaliana*) (Figure 19 - Figure 21). Finally, most recombination measurements were made within the subtelomeric 420 and pericentromeric 3.9 intervals. Subtelomeric and pericentromeric regions are the most active regions in Arabidopsis, which often show opposite trends with respect to changes in recombination frequency. These intervals were extensively used, published, and vetted making them reliable as indicators for genome-wide trends.

4.1 PMS1 does not affect meiotic crossover recombination in Arabidopsis

PMS1 is a core subunit of the MutLα (MLH1/PMS1) heterodimer. After DNA replication, the role of PMS1 within the MMR system is to nick the mismatches that were recognized by the upstream MutS dimers (Iyer et al., 2006; Li, 2008; Fukui, 2010; Larrea et al., 2010; Han et al., 2022). In the mouse and fission yeast models, meiotic defects were observed in the *pms1* null mutants. In mice, its loss of function translates into sterility and improper chromosome synapsis (Baker et al., 1995). In fission yeast, a 12% decrease in spore viability, and a 50% decrease in meiotic

division success are observed (Schär et al., 1997). In *A. thaliana*, PMS1 has been characterized as a limiting factor of somatic homeologous recombination. Severe fertility issues were also observed with decreased seed set and pollen viability in both the heterozygous and homozygous mutants (Li et al., 2009). The fertility phenotype was very compelling and contributed to my interest in the role of PMS1 in meiotic crossover recombination. The alleles used in Li et al., 2009 are not commercially available and could not be obtained.

In my work, recombination frequency assessment for *pms1-1* null mutant and *PMS1* overexpression under its native promoter and *DMC1* promoter did not show any significant differences with the wildtype controls in both tested intervals, 420 and 3.9 (Figure 12 - Figure 15, and Figure 22). Moreover, contrary to Li et al., 2009, no major effect was observed when assessing the fertility of the *pms1* null mutant in the Ziolkowski lab plant growth conditions (Supplemental figure 10). The minor observed issues seem more likely to be due to a less efficient repair of accumulated mismatches combined with environmental factors. Indeed, both measurements of crossover rate and fertility assessments did not indicate any differences in comparison to wild-type controls (Figure 12 - Figure 15, and Supplemental figure 10).

Additionally, *mlh1* and *mlh3* loss of function phenotypes are similar to each other, suggesting that the observed meiotic effects are caused by the loss of function of MutLγ (MLH1-MLH3), not MutLα (MLH1-PMS1). My results assert that PMS1 is not directly involved in Arabidopsis meiotic crossover recombination.

4.2 MutLγ is required for class I crossover formation and crossover assurance

MLH1 and MLH3 form the MutLγ endonuclease which is believed to be the main resolvase for the class I crossovers (Hunter, 2007; Hunter, 2015; Mercier et al., 2015; Dluzewska et al., 2018; Ziolkowski, 2022). Crossover recombination frequency measurements show consistent values between both *mlh1* and *mlh3* null mutants (Figure 16 and Figure 17). The obtained values are milder than what would be

expected for the loss of 90% of Arabidopsis crossovers. This phenotype is further confirmed by the cytological phenotype of *mlh1-4,* where an average of 3.78 chiasmata/meiocyte is observed. The *mlh1-4 -/-* contrasts with *ZMM* null mutants that show the expected ~1 chiasma/meiocyte. Moreover, even when the number of chiasmata/cell is higher than 5, univalents are still observed, with an average of only 3 bivalents/meiocyte (Figure 18). The cytological phenotype of *mlh1-4 -/-* also shows that the loss of MutLγ causes a loss of crossover assurance. Crossover assurance refers to the assurance that every pair of chromosomes receives at least one crossover (Hunter, 2007; Shinohara et al., 2008; Li et al., 2021). The first conclusion from all these observations is that MutLγ is indispensable for class I crossovers and crossover assurance in *A. thaliana*.

4.3 MutLγ is the main resolvase but not the only resolvase of class I crossovers

The cytogenetics data discussed above are congruent with the obtained recombination and fertility data. Again, the *mlh1* and *mlh3* null mutants are significantly affected but recombine more than a *zmm* null mutant and are still relatively fertile (Figure 19 - Figure 21). These results also suggest that in the case of MutLγ loss of function, (an)other resolvase(s) can process dHJ determined by ZMM to yield crossovers. One possibility is that MUS81 may take over the function of MutLγ at least in a subset of ZMM-stabilized intermediates. This is supported by the observation that the triple *mlh1-1 mlh3-1 mus81 ---/---* does show a decreased fertility compared to *mlh1-1*, *mlh3-1*, *mus81* single and *mlh1-1 mlh3-1* double mutants (Figure 21). However, I was not able to assess recombination frequency in this line because the *mlh1-1* mutant allele used to construct it was in a Ws background whereas the *mlh3-1* and *mus81* were in a Col background. A fully homogenous background is required to assess recombination without accounting for the effect of the heterozygosity of the background. Therefore, in the future, I plan to cross the new *mlh1-4 +/-* allele I generated in this work to *mus81* +/-, which both are in the Col background. The *mlh1-4 -/-* allele has an average of 3.78

chiasmata per meiocyte. If MUS81 is responsible for the additional ~3 chiasmata, I expect to observe 10 univalents and 0 chiasmata. If MUS81 is not responsible for them, I would expect ~3 chiasmata per cell.

Additionally, in my work, when *MLH3* is overexpressed in a *hei10* mutant background, an increase in crossover rate is observed and plant fertility is improved. Furthermore, *MLH3 HEI10* double overexpressors display an additive effect on crossover recombination frequency in the 1.26 interval when compared to the two single overexpressors. This further suggests that MutLγ can resolve or promote the resolution of crossover intermediates that were not designated by the ZMM machinery.

Interestingly, additional evidence supporting the view that class I crossovers can be resolved by other endonucleases than MutLγ, comes from the *MLH3* nuclease dead (*mlh3 DN/DN*) mutant analysis in mice. Crossover recombination in *mlh3 DN/DN* mice is significantly affected but is less severe than in the null mutant. The MLH3 DN/DN molecules can still bind DNA, suggesting that they may play a signaling role where they recruit other endonucleases to resolve the joint molecules (Lipkin et al., 2002; Toledo et al., 2019). Moreover, in other Eukaryotes, MutSγ, the heterodimer responsible for recruiting MutLγ to dHJs, is indirectly involved in regulating noncrossover and class II crossover factors. Evidence shows that the loss of function of MutSγ hindrance SC formation which results in negative effects on both class I and II crossover formation (Milano et al., 2019). Considering the available literature and my results together, similarly to other eukaryotes (Edelmann et al., 1999; Agarwal and Roeder, 2000; Kneitz et al., 2000; Shodhan et al., 2014; Pattabiraman et al., 2017), Arabidopsis class I crossovers may also be resolved by other resolvases than MutLγ, and MutLγ may not be class I exclusive, suggesting that this *modus operandum* could be a general feature of meiotic crossover resolution and ZMM pathway.

4.4 MutLγ's ability to boost recombination frequency is limited

I generated two levels of overexpression lines for *MLH1* and *MLH3* to test the tolerated levels of their expression. The overexpression lines under the control of the endogenous promoters show no or only small increases in RF in the tested intervals (Figure 23). Moreover, the overexpression lines under the control of *DMC1* promoter show a very drastic decrease in RF (Figure 25). This was until they were backcrossed to Col, and the MutL transgene expression level was halved resulting in RFs warmer than the Col-420 control (Figure 27). This denotes that *MLH1* and *MLH3* expression levels are very tightly regulated to maintain them at physiologically tolerated levels (Figure 36, Veitia, 2002; Veitia et al., 2008; Veitia et al., 2013; Johnson et al., 2019; Morrill and Amon, 2019). This is further comforted by the empirical observations made while caring for the plants. Indeed, I observed an active counterselection against high expression levels of these genes: The BASTA-resistant plants' proportion decreases after every filial generation. As BASTA resistance is an indicator of the transcriptional activity of the transgene, this suggests extensive silencing/counterselection of the MutL transgene additional copies.

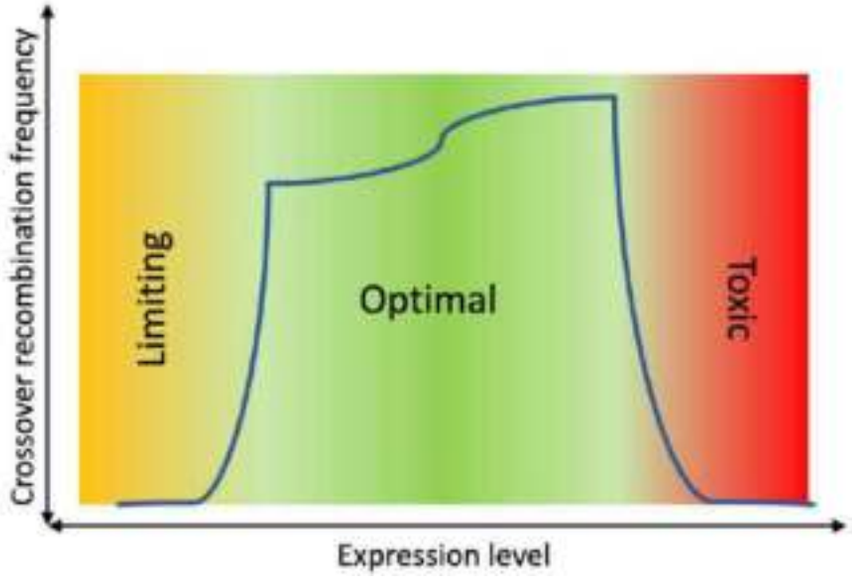


Figure 36. Dosage stabilization hypothesis. Under this hypothesis the expression levels of *MLH1* and *MLH3* would be sufficient to maintain a wildtype level, or slightly higher, crossover rate within a definite interval. A lower or higher dosage of transcripts is detrimental. Adapted from Morrill and Amon, 2019.

The limited effect of *MLH1* and *MLH3* overexpression could also be due to the interfering nature of class I crossovers (Jones and Franklin, 2006; Berchowitz and Copenhaver, 2010; Wang et al., 2015; von Diezmann and Rog, 2021; Li et al., 2021). Indeed, MutLγ is not part of the ZMM proteins and operates downstream of this machinery. ZYP1 is indispensable for maintaining interference in Arabidopsis (Capilla-Pérez et al., 2021; France et al., 2021). HEI10 is also believed to be involved in regulating interference through its coarsening to crossover sites (Morgan et al., 2021). As such, the activity of the additional molecules of MutLγ, which happens after crossover site designation, could be limited by the scarcity of substrates to be resolved as crossovers.

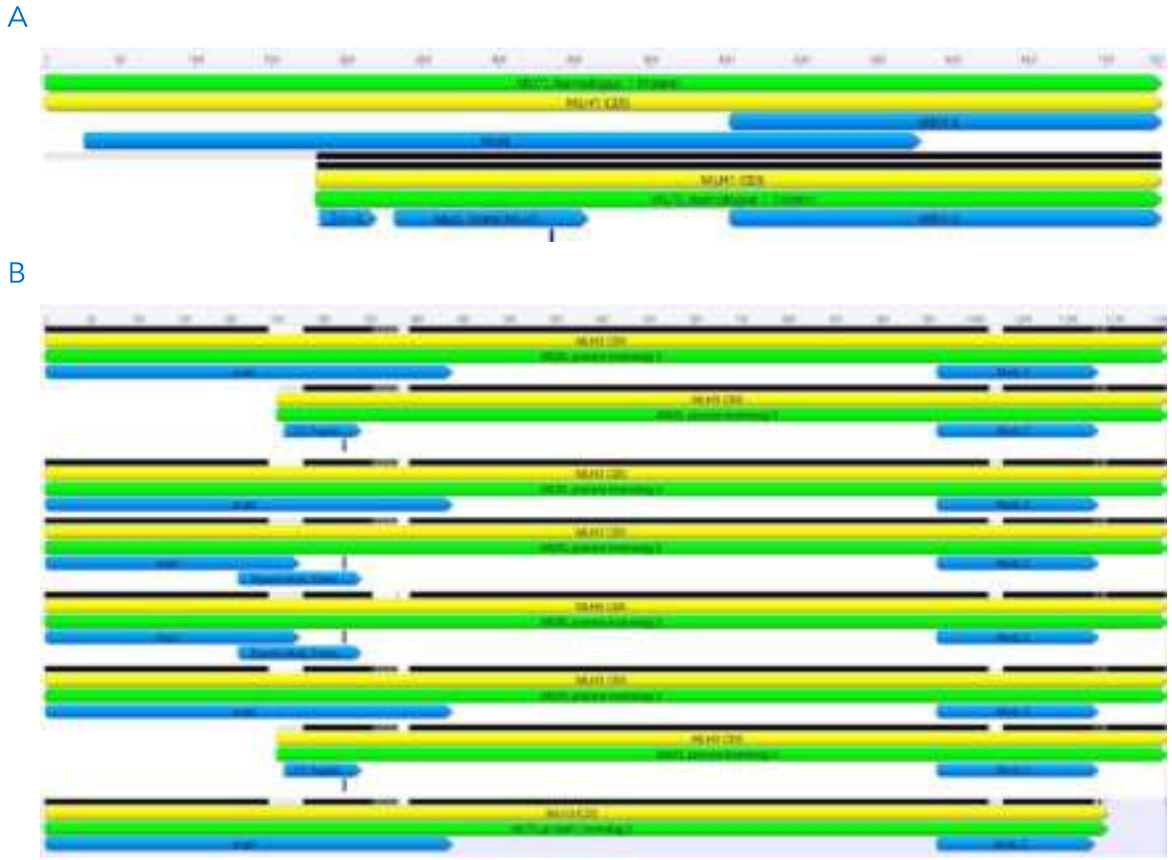


Figure 37. Representation of the different variants of *MLH1* and *MLH3* transcripts. (A) *MLH1*. (B) *MLH3*. Green arrowheads represent the produced protein. Yellow arrowheads represent the coding DNA sequence (CDS). Blue arrowheads represent the positions of active domains. The black lines indicate the homology between the variants, with black representing 100% homology, shades of grey representing lower homology, and white representing gaps. Generated using Geneious databases and alignment pipeline.

Finally, it is also possible that *MLH1* and *MLH3* are subjected to a very strong regulation that maintains the final pool of functional molecules within physiologically tolerated levels (Smith et al., 2003; Reddy et al., 2013; Wang and Zhou, 2014). RT-qPCR results showed that the additional copies did yield higher expression levels, but the measured crossover recombination frequencies were not proportional. Both *MLH1* and *MLH3* contain alternative start sites and have several predicted splicing variants (Figure 37), some of which were confirmed. The predicted variants would produce significantly different versions of the proteins with a non-functional ATPase domain for MLH1 and TopoII MutL trans domain for MLH3. One can speculate that the production of these protein variants can be used as a way to limit the number of available fully functional molecules.

4.5 MutLγ expression level alteration hinders Arabidopsis fertility

The assessment of fertility in *MLH1* and *MLH3* mutants and overexpression lines showed decreased values for all the tested parameters, seed set, silique length, pollen viability, and pollen density (Figure 19 - Figure 21, and Supplemental figure 6 - Supplemental figure 10). This is coherent with the dosage stabilization hypothesis. Loss of function and excessive overexpression both yield a similar phenotype of loss of fitness (Figure 36). In the loss of function situation, the lack of crossover events causes a failure in proper chromosome segregation (Hunter, 2007; Shinohara et al., 2008; Hunter, 2015; Mercier et al., 2015; Li et al., 2021). In the toxic overexpression context, it could be an overactivated negative feedback loop also leading to missegregation or damage to the chromatin due to the endonuclease activity of the MutLγ heterodimer. The latter can be verified through cytology.

4.6 Can *EXO1* overexpression be used to increase the global crossover rate in Arabidopsis?

EXO1, the exonuclease that is responsible for resecting DSBs in the early stages of their processing, is also involved in class I crossover resolution. It can physically

interact with MLH1 and so the MutLγ endonuclease (Cannavo et al., 2020; Kulkarni et al., 2020; Sanchez et al., 2020). It is believed to be a regulatory add-on as this interaction is independent of its enzymatic activity.

Contrary to *MLH1* and *MLH3* overexpression lines, *EXO1b* nuclease-dead overexpression lines are more stable. The construct is not counter-selected and shows a multiple insertion behavior. The increase in RF as measured in the 420 subtelomeric interval is also more stable and higher than what was observed in *MLH1* and *MLH3* overexpressors (Figure 23 - Figure 25, Figure 34, and Figure 35).

To further investigate this phenotype, I plan to characterize the effect of *EXO1b* nuclease-dead on Arabidopsis fertility and generate double and triple overexpressors with *MLH1* and *MLH3.* Expression levels of *EXO1b* will be quantified along with other meiotic factors. A whole genome crossover mapping can also be considered after measuring recombination in other intervals. This would provide information about the distribution of the additional crossovers across the genome. *EXO1a* characterization will also be attempted.

5 Conclusions

Plants, including Arabidopsis, possess three homologs of the MutL protein, MLH1, PMS1, and MLH3. In this work, I showed that unlike other organisms, like yeast and mice, only Arabidopsis MLH1 and MLH3 are directly involved in meiotic crossover recombination. Arabidopsis PMS1 does not have any significant effect for all tested meiotic parameters. *mlh1* and *mlh3* loss of function mutants display a significant decrease in meiotic crossover recombination in the tested intervals, and a significant decrease in fertility. Interestingly, the overexpression of *MLH1* and *MLH3,* at two different promoter-determined levels, through additional copies of the genes, also induces a decrease in fertility, but a small but significant increase in crossover recombination rate in the tested intervals. This suggests that *MLH1* and *MLH3* display a dosage stabilization behavior in Arabidopsis. Both loss of function and excessive activity are detrimental to plant fitness.

Cytogenetic characterization of an *mlh1* null mutant showed that the MutLγ loss of function induces a strong decrease in chiasma and bivalent numbers coupled with a loss of crossover assurance. Indeed, univalents were observed in all cells even in the ones with more than 5 chiasmata. Considering ZMM as the pathway responsible for the majority of crossover events in Arabidopsis, this observation shows that MutLγ is required for the resolution of ZMM intermediates into crossovers. The *mlh1* null cytogenetic phenotype is however less severe than *zmm* null mutants, suggesting that some of the ZMM intermediates could be processed by another endonuclease.

Genetic interactions between *mlh1/3* and other meiotic factors showed that *mutlγ* loss of function has an additive effect with *mus81* and a partial recovery when combined with *fancm* loss of function when assessed using plant fertility. When assessing the crossover recombination rate, *MLH3,* but not *MLH1,* shows an additive effect when overexpressed together with *HEI10*. Moreover, *MLH3* overexpression increases slightly the crossover rate and seed set in a *hei10* null

background. These results suggest that MutLγ is not strictly active downstream of the ZMM pathway and may also resolve intermediates generated by other types of machinery.

Finally, the overexpression of *EXO1b* nuclease-dead can increase the meiotic crossover rate to a higher level than the MutLγ subunits in the tested interval. EXO1 is an add-on subunit of the MutLγ complex. one can assume that the observed phenotype is due to EXO1's ability to foster MutLγ endonuclease activity (Cannavo et al., 2020; Kulkarni et al., 2020; Sanchez et al., 2020). Further testing is required to properly understand the observed phenotype.

6 Acknowledgments

mlh1-1, mlh1-1 mlh3-1, mlh1-1 mlh3-1 mus81, and *mlh1-1 mlh3-1 fancm* seed lines were generously shared by Dr Mónica Pradillo from the Complutense University in Madrid. *MLH1* endogenous promoter overexpression construct was cloned by Dr Weronika Sura. *MLH1* and *MLH3* overexpression constructs under the control of *DMC1* promoter were cloned by MSc. Julia Dluzewska. *PMS1* overexpression constructs, both endogenous and *DMC1* promoters, were cloned and transformed into plants by MSc. Olga Maria Wienskowska. Meiotic chromosome spreads and interpretation were conducted with the help of Dr Alexandre Pele.

7 Bibliography

Agarwal S, Roeder GS (2000) Zip3 Provides a Link between Recombination Enzymes and Synaptonemal Complex Proteins. Cell **102**: 245–255

Aguilera A, Rothstein R (2007) Molecular Genetics of Recombination. doi: 10.1007/978-3-540-71021-9

Alexander MP (1969) Differential staining of aborted and nonaborted pollen. Biotechnic and Histochemistry **44**: 117–122

Alou AH, Azaiez A, Jean M, Belzile FJ (2004) Involvement of the Arabidopsis thaliana AtPMS1 gene in somatic repeat instability. Plant Mol Biol **56**: 339–349

Baker SM, Bronner CE, Zhang L, Plug AW, Robatzek M, Warren G, Elliott EA, Yu J, Ashley T, Arnheim N, et al (1995) Male mice defective in the DNA mismatch repair gene PMS2 exhibit abnormal chromosome synapsis in meiosis. Cell **82**: 309–319

Bębenek A, Ziuzia-Graczyk I (2018) Fidelity of DNA replication—a matter of proofreading. Current Genetics 2018 64:5 **64**: 985–996

Berchowitz L, Copenhaver G (2010) Genetic Interference: Dont Stand So Close to Me. Curr Genomics **11**: 91–102

Bieluszewski T, Szymanska-Lejman M, Dziegielewski W, Zhu L, Ziolkowski PA (2022) Efficient Generation of CRISPR/Cas9-Based Mutants Supported by Fluorescent Seed Selection in Different Arabidopsis Accessions. Methods in Molecular Biology. Humana Press Inc., pp 161–182

Blackwell AR, Dluzewska J, Szymanska-Lejman M, Desjardins S, Tock AJ, Kbiri N, Lambing C, Lawrence EJ, Bieluszewski T, Rowan B, et al (2020) MSH 2 shapes the meiotic crossover landscape in relation to interhomolog polymorphism in Arabidopsis. EMBO J. doi: 10.15252/embj.2020104858

Cannavo E, Sanchez A, Anand R, Ranjha L, Hugener J, Adam C, Acharya A, Weyland N, Aran-Guiu X, Charbonnier J-B, et al (2020) Regulation of the MLH1–MLH3 endonuclease in meiosis. Nature **586**: 618–622

Capilla-Pérez L, Durand S, Hurel A, Lian Q, Chambon A, Taochy C, Solier V, Grelon M, Mercier R (2021) The synaptonemal complex imposes crossover interference and heterochiasmy in Arabidopsis. Proceedings of the National Academy of Sciences **118**: e2023613118

Chelysheva L, Gendrot G, Vezon D, Doutriaux MP, Mercier R, Grelon M (2007) Zip4/Spo22 Is Required for Class I CO Formation but Not for Synapsis Completion in Arabidopsis thaliana. PLoS Genet **3**: 802–813

Chelysheva L, Vezon D, Chambon A, Gendrot G, Pereira L, Lemhemdi A, Vrielynck N, Le Guin S, Novatchkova M, Grelon M (2012) The Arabidopsis HEI10 is a new ZMM protein related to Zip3. PLoS Genet. doi: 10.1371/journal.pgen.1002799

Chirinos-Arias MC, Spampinato CP (2020) Growth and development of AtMSH7 mutants in Arabidopsis thaliana. Plant Physiology and Biochemistry **146**: 329–336

Culligan KM, Meyer-Gauen G, Lyons-Weiler J, Hays JB (2000) Evolutionary origin, diversification and specialization of eukaryotic MutS homolog mismatch repair proteins.

von Diezmann L, Rog O (2021) Let's get physical – mechanisms of crossover interference. J Cell Sci. doi: 10.1242/jcs.255745

Dion É, Li L, Jean M, Belzile F (2007) An Arabidopsis MLH1 mutant exhibits reproductive defects and reveals a dual role for this gene in mitotic recombination. Plant Journal **51**: 431–440

Dluzewska J, Szymanska M, Ziolkowski PA (2018) Where to Cross Over ? Defining Crossover Sites in Plants. **9**: 1–20

Edelmann W, Cohen PE, Kneitz B, Winand N, Lia M, Heyer J, Kolodner R, Pollard JW, Kucherlapati R (1999) Mammalian MutS homologue 5 is required for chromosome pairing in meiosis. Nat Genet **21**: 123–127

Fishel R (2015) Mismatch repair. Journal of Biological Chemistry **290**: 26395–26403

France MG, Enderle J, Röhrig S, Puchta H, Franklin FCH, Higgins JD (2021) ZYP1 is required for obligate cross-over formation and cross-over interference in Arabidopsis. Proc Natl Acad Sci U S A **118**: e2021671118

Fukui K (2010) DNA mismatch repair in eukaryotes and bacteria. J Nucleic Acids. doi: 10.4061/2010/260512

Groothuizen FS, Winkler I, Cristóvão M, Fish A, Winterwerp HHK, Reumer A, Marx AD, Hermans N, Nicholls RA, Murshudov GN, et al (2015) MutS/MutL crystal structure reveals that the MutS sliding clamp loads MutL onto DNA. Elife. doi: 10.7554/eLife.06744.001

Han X-P, Yang X-W, Liu J (2022) MutS and MutL sliding clamps in DNA mismatch repair. Genome Instab Dis **1**: 3

Harada H, Nie Y, Araki I, Soeno T, Chuman M, Washio M, Sakuraya M, Ushiku H, Niihara M, Hosoda K, et al (2021) Haploinsufficiency by minute MutL homolog 1 promoter DNA methylation may represent unique phenotypes of microsatellite instability-gastric carcinogenesis. PLoS One **16**: e0260303

Harfe BD, Jinks-Robertson S (2000) DNA MISMATCH REPAIR AND GENETIC INSTABILITY. Annu Rev Genet **34**: 359–399

He W, Rao HBDP, Tang S, Bhagwat N, Kulkarni DS, Ma Y, Chang MAW, Hall C, Bragg JW, Manasca HS, et al (2020) Regulated Proteolysis of MutSγ Controls Meiotic Crossing Over. Mol Cell **78**: 168-183.e5

He W, Verhees GF, Bhagwat N, Yang Y, Kulkarni DS, Lombardo Z, Lahiri S, Roy P, Zhuo J, Dang B, et al (2021) SUMO fosters assembly and functionality of the MutSγ complex to facilitate meiotic crossing over. Dev Cell **56**: 2073-2088.e3

Hegan DC, Narayanan L, Jirik FR, Edelmann W, Liskay RM, Glazer PM (2006) Differing patterns of genetic instability in mice deficient in the mismatch repair genes Pms2, Mlh1, Msh2, Msh3 and Msh6. Carcinogenesis **27**: 2402–2408

Hord CLH, Sun YJ, Pillitteri LJ, Torii KU, Wang H, Zhang S, Ma H (2008) Regulation of Arabidopsis early anther development by the mitogen-activated protein

kinases, MPK3 and MPK6, and the ERECTA and related receptor-like kinases. Mol Plant **1**: 645–658

Hunter N (2015) Meiotic recombination: The essence of heredity. Cold Spring Harb Perspect Biol **7**: 1–35

Hunter N (2007) Meiotic recombination. Molecular Genetics of Recombination. Springer, Berlin, Heidelberg, pp 381–442

Iyer RR, Pluciennik A, Burdett V, Modrich PL (2006) DNA Mismatch Repair: Functions and Mechanisms. Chem Rev **106**: 302–323

Jackson N, Sanchez-Moran E, Buckling E, Armstrong SJ, Jones GH, Franklin FCH (2006) Reduced meiotic crossovers and delayed prophase I progression in AtMLH3-deficient Arabidopsis. EMBO J **25**: 1315–1323

Jiang Y, Marszalek PE (2011) Atomic force microscopy captures MutS tetramers initiating DNA mismatch repair. EMBO Journal. doi: 10.1038/emboj.2011.180

Jiricny J (2013) Postreplicative Mismatch Repair. Cold Spring Harb Perspect Biol **5**: a012633–a012633

Johnson AF, Nguyen HT, Veitia RA (2019) Causes and effects of haploinsufficiency. Biological Reviews **94**: 1774–1785

Jones GH, Franklin FCH (2006) Meiotic Crossing-over: Obligation and Interference. Cell **126**: 246–248

Kneitz B, Cohen PE, Avdievich E, Zhu L, Kane MF, Hou H, Kolodner RD, Kucherlapati R, Pollard JW, Edelmann W (2000) MutS homolog 4 localization to meiotic chromosomes is required for chromosome pairing during meiosis in male and female mice. Genes Dev **14**: 1085–1097

Kulkarni DS, Owens SN, Honda M, Ito M, Yang Y, Corrigan MW, Chen L, Quan AL, Hunter N (2020) PCNA activates the MutLγ endonuclease to promote meiotic crossing over. Nature **586**: 623–627

Kunkel TA, Bebenek K (2000) DNA Replication Fidelity. Annu Rev Biochem **69**: 497–529

Lambing C, Franklin FCH, Wang C-JR (2017) Understanding and Manipulating Meiotic Recombination in Plants. Plant Physiol **173**: 1530–1542

Lario LD, Botta P, Casati P, Spampinato CP (2015) Role of AtMSH7 in UV-B-induced DNA damage recognition and recombination. J Exp Bot **66**: 3019–3026

Larrea AA, Lujan SA, Kunkel TA (2010) SnapShot: DNA Mismatch Repair. Cell **141**: 730-730.e1

Li G-M (2008) Mechanisms and functions of DNA mismatch repair. Cell Res **18**: 85–98

Li L, Dion E, Richard G, Domingue O, Jean M, Belzile FJ (2009) The Arabidopsis DNA mismatch repair gene PMS1 restricts somatic recombination between homeologous sequences. Plant Mol Biol **69**: 675–684

Li X, Zhang J, Huang J, Xu J, Chen Z, Copenhaver GP, Wang Y (2021) Regulation of interference-sensitive crossover distribution ensures crossover assurance in Arabidopsis. Proc Natl Acad Sci U S A. doi: 10.1073/PNAS.2107543118/-/DCSUPPLEMENTAL

Lin Z, Nei M, Ma H (2007a) The origins and early evolution of DNA mismatch repair genes - Multiple horizontal gene transfers and co-evolution. Nucleic Acids Res **35**: 7591–7603

Lin Z, Nei M, Ma H (2007b) The origins and early evolution of DNA mismatch repair genes—multiple horizontal gene transfers and co-evolution. Nucleic Acids Res **35**: 7591–7603

Lipkin SM, Moens PB, Wang V, Lenzi M, Shanmugarajah D, Gilgeous A, Thomas J, Cheng J, Touchman JW, Green ED, et al (2002) Meiotic arrest and aneuploidy in MLH3-deficient mice. Nat Genet **31**: 385–390

Lloyd AH, Milligan AS, Langridge P, Able JA (2007) TaMSH7: A cereal mismatch repair gene that affects fertility in transgenic barley (Hordeum vulgare L.). BMC Plant Biol **7**: 67

Lu X, Liu X, An L, Zhang W, Sun J, Pei H, Meng H, Fan Y, Zhang C (2008) The Arabidopsis MutS homolog AtMSH5 is required for normal meiosis. Cell Research 2008 18:5 **18**: 589–599

Lynn A, Soucek R, Börner GV (2007) ZMM proteins during meiosis: Crossover artists at work. Chromosome Research **15**: 591–605

Macaisne N, Vignard J, Mercier R (2011) SHOC1 and PTD form an XPF–ERCC1-like complex that is required for formation of class I crossovers. J Cell Sci **124**: 2687–2691

Martín AC, Shaw P, Phillips D, Reader S, Moore G (2014) Licensing MLH1 sites for crossover during meiosis. Nat Commun **5**: 1–5

Melamed-Bessudo C, Yehuda E, Stuitje AR, Levy AA (2005) A new seed-based assay for meiotic recombination in Arabidopsis thaliana. Plant Journal **43**: 458–466

Mercier R, Mézard C, Jenczewski E, Macaisne N, Grelon M (2015) The Molecular Biology of Meiosis in Plants. Annu Rev Plant Biol **66**: 297–327

Milano CR, Kim Holloway J, Zhang Y, Jin B, Smith C, Bergman A, Edelmann W, Cohen PE (2019) Mutation of the ATPase domain of mutS homolog-5 (MSH5) reveals a requirement for a functional MutSG complex for all crossovers in mammalian meiosis. G3: Genes, Genomes, Genetics **9**: 1839–1850

Modrich P, Lahue R (1996) Mismatch repair in replication fidelity, genetic recombination, and cancer biology. Annu Rev Biochem **65**: 101–133

Morgan C, Fozard JA, Hartley M, Henderson IR, Bomblies K, Howard M (2021) Diffusion-mediated HEI10 coarsening can explain meiotic crossover positioning in Arabidopsis. Nature Communications 2021 12:1 **12**: 1–11

Morrill SA, Amon A (2019) Why haploinsufficiency persists. Proceedings of the National Academy of Sciences **116**: 11866–11871

Pattabiraman D, Roelens B, Woglar A, Villeneuve AM (2017) Meiotic recombination modulates the structure and dynamics of the synaptonemal complex during C. elegans meiosis. PLoS Genet **13**: e1006670

Peterson R, Slovin JP, Chen C (2010) A simplified method for differential staining of aborted and non-aborted pollen grains. International Journal of Plant Biology **1**: 66–69

Peterson SE, Keeney S, Jasin M (2020) Mechanistic Insight into Crossing over during Mouse Meiosis. Mol Cell **78**: 1252-1263.e3

Putnam CD (2020) MutS sliding clamps on an uncertain track to DNA mismatch repair. Proceedings of the National Academy of Sciences **117**: 20351–20353

Pyatnitskaya A, Borde V, De Muyt A (2019) Crossing and zipping: molecular duties of the ZMM proteins in meiosis. Chromosoma **128**: 181–198

Qiu R, Sakato M, Sacho EJ, Wilkins H, Zhang X, Modrich P, Hingorani MM, Erie DA, Weninger KR (2015) MutL traps MutS at a DNA mismatch. Proc Natl Acad Sci U S A **112**: 10914–10919

Reddy ASN, Marquez Y, Kalyna M, Barta A (2013) Complexity of the Alternative Splicing Landscape in Plants. Plant Cell **25**: 3657–3683

Reyes GX, Schmidt TT, Kolodner RD, Hombauer H (2015) New insights into the mechanism of DNA mismatch repair. Chromosoma **124**: 443–462

Sachadyn P (2010) Conservation and diversity of MutS proteins. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis **694**: 20–30

Sanchez A, Adam C, Rauh F, Duroc Y, Ranjha L, Lombard B, Mu X, Wintrebert M, Loew D, Guarné A, et al (2020) Exo1 recruits Cdc5 polo kinase to MutLγ to ensure efficient meiotic crossover formation. Proceedings of the National Academy of Sciences **117**: 30577–30588

Schär P, Baur M, Schneider C, Kohli J (1997) Mismatch Repair in *Schizosacchromyces pombe* Requires the *mutL* Homologous Gene *pms1:* Molecular Cloning and Functional Analysis. Genetics **146**: 1275–1286

Serra H, Svačina R, Baumann U, Whitford R, Sutton T, Bartoš J, Sourdille P (2021) Ph2 encodes the mismatch repair protein MSH7-3D that inhibits wheat

homoeologous recombination. Nat Commun. doi: 10.1038/s41467-021-21127-1

Shinohara M, Oh SD, Hunter N, Shinohara A (2008) Crossover assurance and crossover interference are distinctly regulated by the ZMM proteins during yeast meiosis. Nat Genet **40**: 299–309

Shodhan A, Lukaszewicz A, Novatchkova M, Loidl J (2014) Msh4 and Msh5 Function in SC-Independent Chiasma Formation During the Streamlined Meiosis of *Tetrahymena*. Genetics **198**: 983–993

Shrestha KS, Aska EM, Tuominen MM, Kauppi L (2021) Tissue-specific reduction in MLH1 expression induces microsatellite instability in intestine of Mlh1+/− mice. DNA Repair (Amst) **106**: 103178

Shrestha KS, Aska E-M, Tuominen MM, Kauppi L (2020) Mlh1 haploinsufficiency induces microsatellite instability specifically in intestine. bioRxiv 652198

Smith CWJ, Patton JG, Nadal-Ginard B (2003) ALTERNATIVE SPLICING IN THE CONTROL OF GENE EXPRESSION. https://doi.org/101146/annurev.ge23120189002523 **23**: 527–577

Sugawara N, Goldfarb T, Studamire B, Alani E, Haber JE (2004) Heteroduplex rejection during single-strand annealing requires Sgs1 helicase and mismatch repair proteins Msh2 and Msh6 but not Pms1. Proc Natl Acad Sci U S A **101**: 9315–9320

Tam SM, Samipak S, Britt A, Chetelat RT (2009) Characterization and comparative sequence analysis of the DNA mismatch repair MSH2 and MSH7 genes from tomato. Genetica **137**: 341–354

Toledo M, Sun X, Brieño-Enríquez MA, Raghavan V, Gray S, Pea J, Milano CR, Venkatesh A, Patel L, Borst PL, et al (2019) A mutation in the endonuclease domain of mouse MLH3 reveals novel roles for mutlγ during crossover formation in meiotic prophase I. PLoS Genet. doi: 10.1371/journal.pgen.1008177

Tomimatsu N, Mukherjee B, Catherine Hardebeck M, Ilcheva M, Vanessa Camacho C, Louise Harris J, Porteus M, Llorente B, Khanna KK um, Burma S (2014) Phosphorylation of EXO1 by CDKs 1 and 2 regulates DNA end resection and repair pathway choice. Nat Commun **5**: 3561

Umar A, Kunkel TA (1996) DNA-replication fidelity, mismatch repair and genome instability in cancer cells. Eur J Biochem **238**: 297–307

Veitia RA (2002) Exploring the etiology of haploinsufficiency. BioEssays **24**: 175–184

Veitia RA, Bottani S, Birchler JA (2008) Cellular reactions to gene dosage imbalance: genomic, transcriptomic and proteomic effects. Trends in Genetics **24**: 390–397

Veitia RA, Bottani S, Birchler JA (2013) Gene dosage effects: nonlinearities, genetic interactions, and dosage compensation. Trends in Genetics **29**: 385–393

Walker J, Gao H, Zhang J, Aldridge B, Vickers M, Higgins JD, Feng X (2017) Sexual-lineage-specific DNA methylation regulates meiosis in Arabidopsis. Nature Genetics 2017 50:1 **50**: 130–137

Wang L, Tsutsumi S, Kawaguchi T, Nagasaki K, Tatsuno K, Yamamoto S, Sang F, Sonoda K, Sugawara M, Saiura A, et al (2012a) Whole-exome sequencing of human pancreatic cancers and characterization of genomic instability caused by MLH1 haploinsufficiency and complete deficiency. Genome Res **22**: 208–219

Wang Q, Zhou T (2014) Alternative-splicing-mediated gene expression. Phys Rev E Stat Nonlin Soft Matter Phys **89**: 012713

Wang S, Lee K, Gray S, Zhang Y, Tang C, Morrish RB, Tosti E, van Oers J, Amin MR, Cohen PE, et al (2022) Role of EXO1 nuclease activity in genome maintenance, the immune response and tumor suppression in *Exo1D173A* mice. Nucleic Acids Res **50**: 8093–8106

Wang S, Zickler D, Kleckner N, Zhang L (2015) Meiotic crossover patterns: Obligatory crossover, interference and homeostasis in a single process. Cell Cycle **14**: 305–314

Wang Y, Cheng Z, Huang J, Shi Q, Hong Y, Copenhaver GP, Gong Z, Ma H (2012b) The DNA Replication Factor RFC1 Is Required for Interference-Sensitive Meiotic Crossovers in Arabidopsis thaliana. PLoS Genet **8**: e1003039

Wang Y, Copenhaver GP (2018) Meiotic Recombination: Mixing It Up in Plants. Annu Rev Plant Biol **69**: 577–609

Ward JO, Reinholdt LG, Motley WW, Niswander LM, Deacon DC, Griffin LB, Langlais KK, Backus VL, Schimenti KJ, O'Brien MJ, et al (2007) Mutation in mouse Hei10, an E3 ubiquitin ligase, disrupts meiotic crossing over. PLoS Genet. doi: 10.1371/journal.pgen.0030139

Weigel Detlef, Glazebrook Jane (2002) Arabidopsis : a laboratory manual. Cold Spring Harbor Laboratory Press

Wu S-Y (2003) Dissimilar mispair-recognition spectra of Arabidopsis DNA-mismatch-repair proteins MSH2{middle dot}MSH6 (MutS) and MSH2{middle dot}MSH7 (MutS). Nucleic Acids Res **31**: 6027–6034

Yang M, Hsieh P (2016) DNA Mismatch Repair in Mammals. Genome Stability. Elsevier, pp 303–319

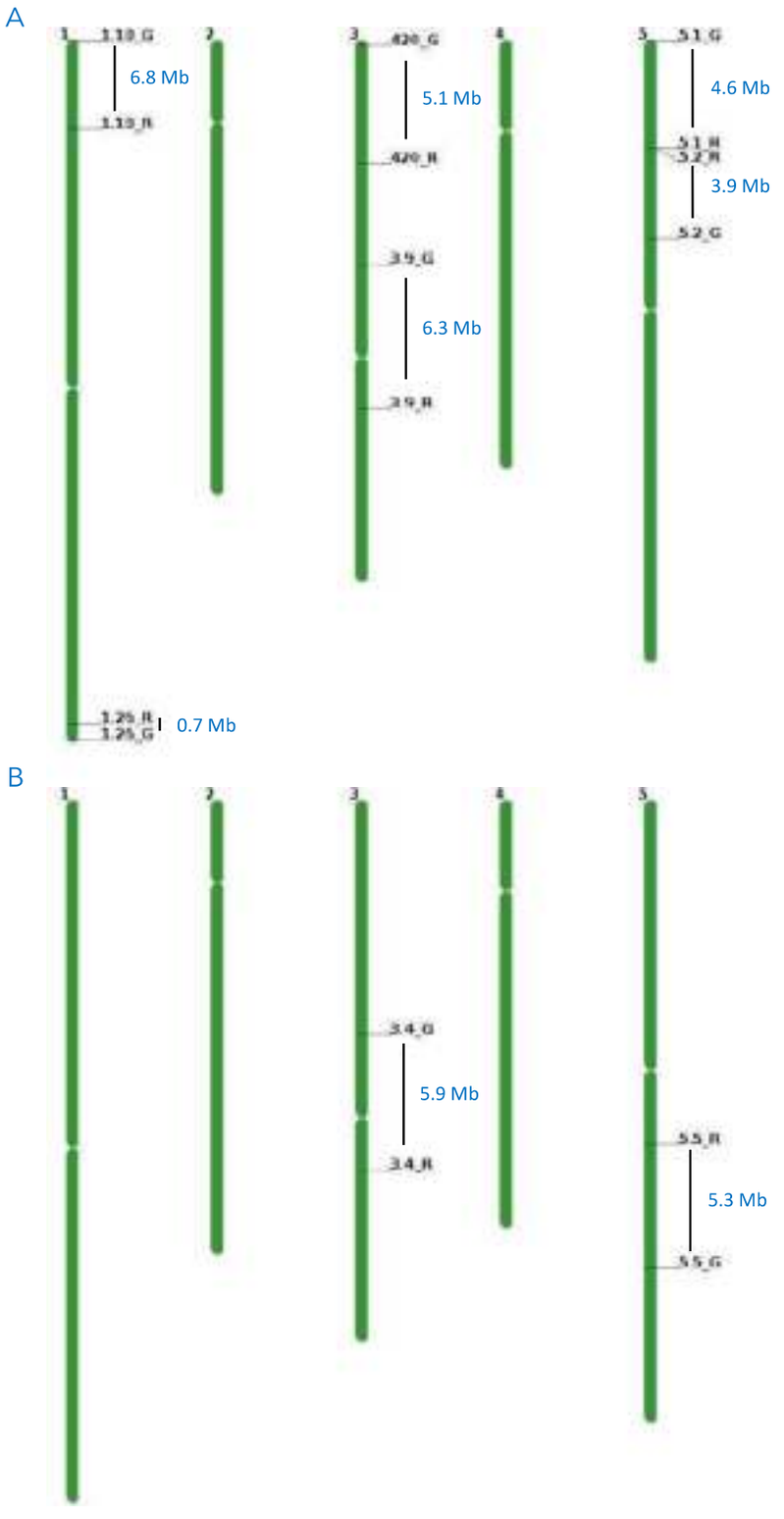
Zhu L, Fernández-Jiménez N, Szymanska-Lejman M, Pelé A, Underwood CJ, Serra H, Lambing C, Dluzewska J, Bieluszewski T, Pradillo M, et al (2021) Natural variation identifies SNI1, the SMC5/6 component, as a modifier of meiotic crossover in Arabidopsis. Proc Natl Acad Sci U S A. doi: 10.1073/PNAS.2021970118/-/DCSUPPLEMENTAL

Ziolkowski PA (2022) Why do plants need the ZMM crossover pathway? A snapshot of meiotic recombination from the perspective of interhomolog polymorphism. Plant Reprod **1**: 3

Ziolkowski PA, Berchowitz LE, Lambing C, Yelina NE, Zhao X, Kelly KA, Choi K, Ziolkowska L, June V, Sanchez-Moran E, et al (2015) Juxtaposition of heterozygous and homozygous regions causes reciprocal crossover remodelling via interference during Arabidopsis meiosis. Elife **4**: 1–29

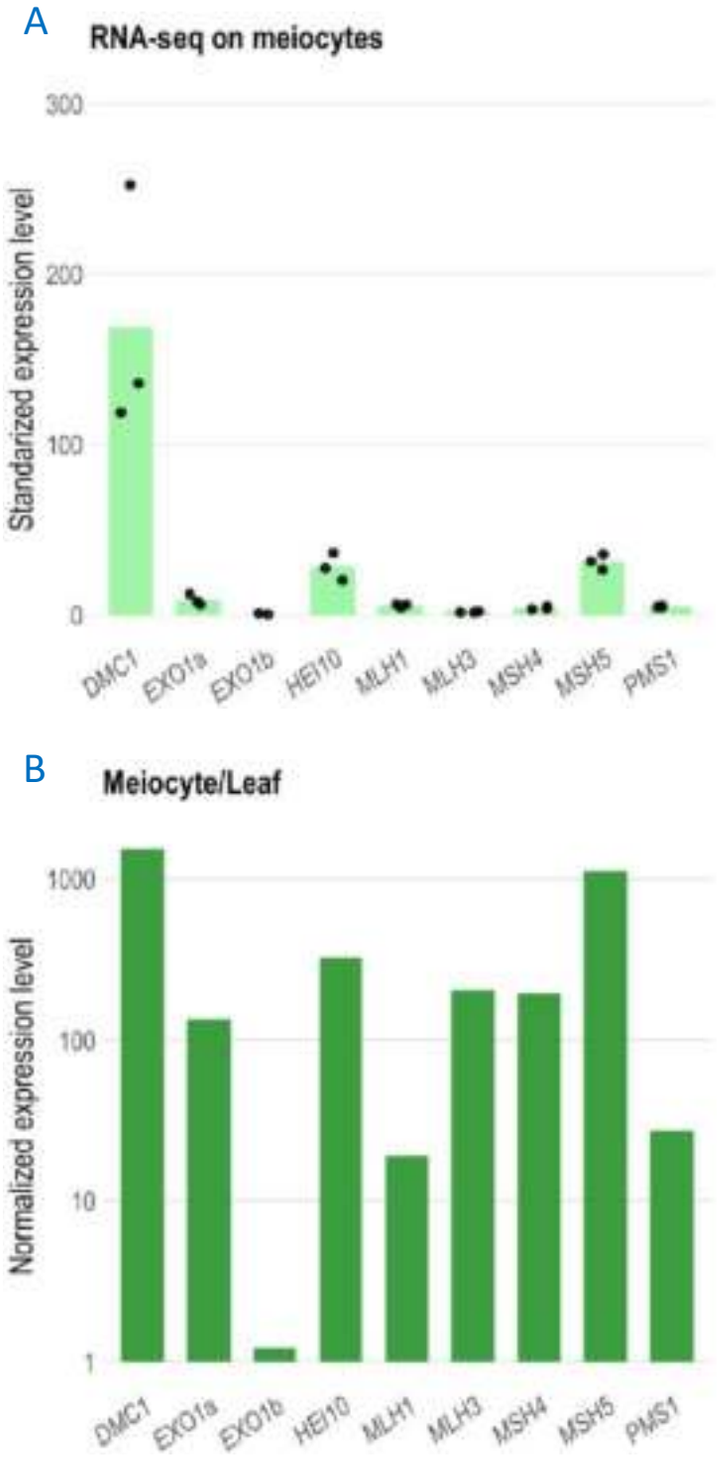
8 Supplemental data

8.1 Chromosome map presentation of the used fluorescent traffic lines



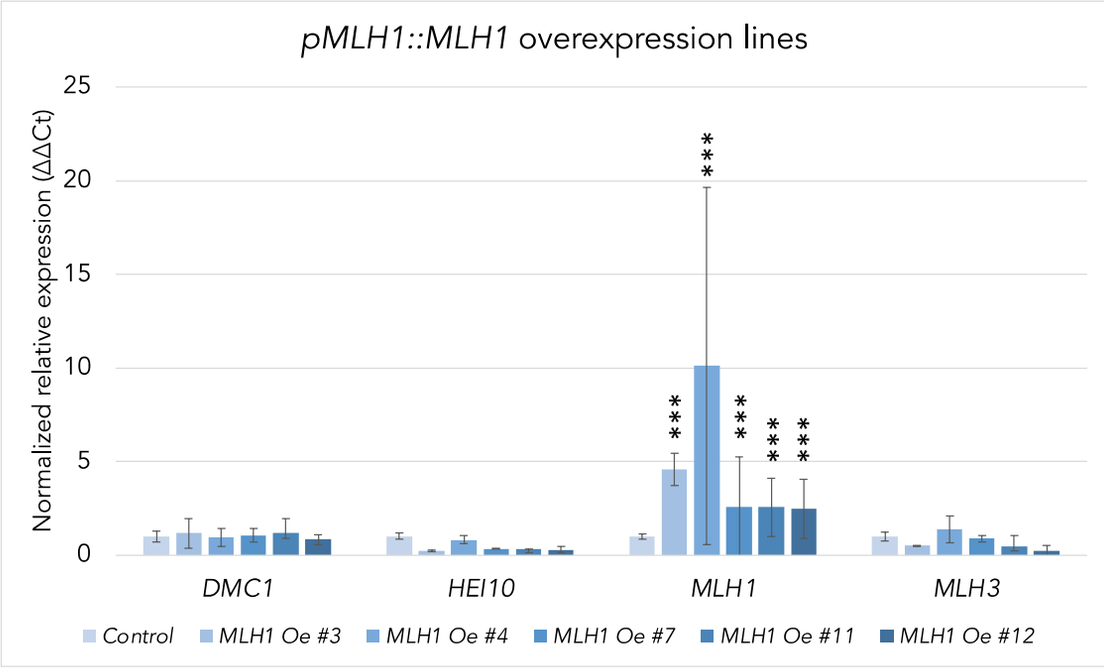
Supplemental figure 1. Representation of the position of the fluorescent tags on Arabidopsis five chromosomes. A. Col traffic lines (CTLs). B. L*er* traffic lines (LTLs). R= dsRed, G= eGFP. The distance between the fluorescent tags of each interval is indicated in Mb.

8.2 Expression level supplemental data

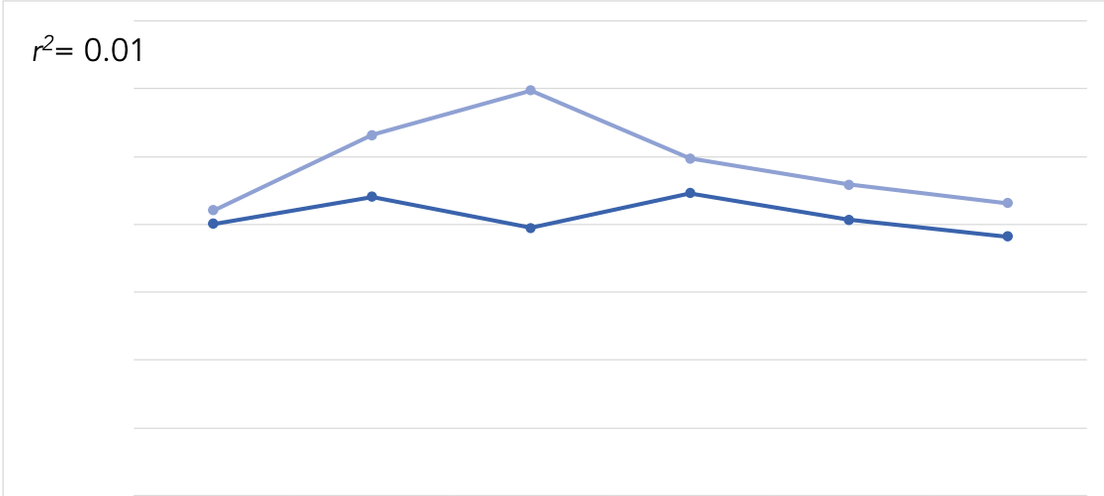


Supplemental figure 2. Graphic representation of several meiotic genes' expression levels in wildtype Arabidopsis. A. Standardized expression levels of *DMC1, EXO1a, EXO1b, HEI10, MLH1, MLH3, MSH4, MSH5,* and *PMS1* in meiocytes. B. Average relative expression level of the same genes in meiocytes normalized to leaf tissue. The data represents n=3 biological replicates. Open access data from Walker et al., 2017.

A



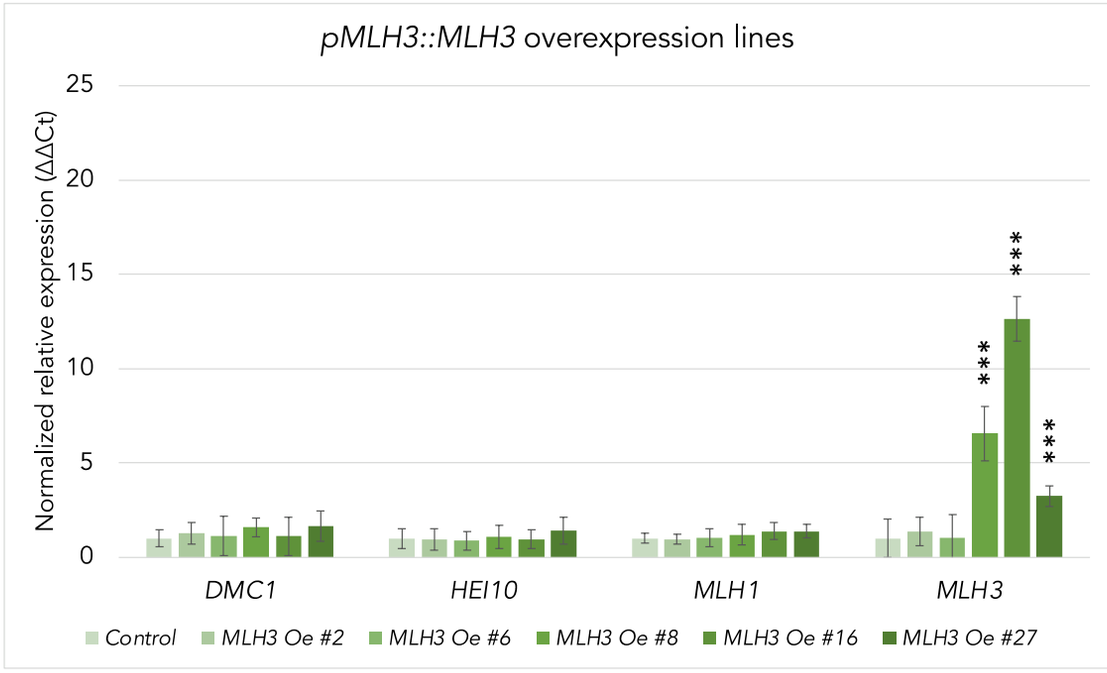
B



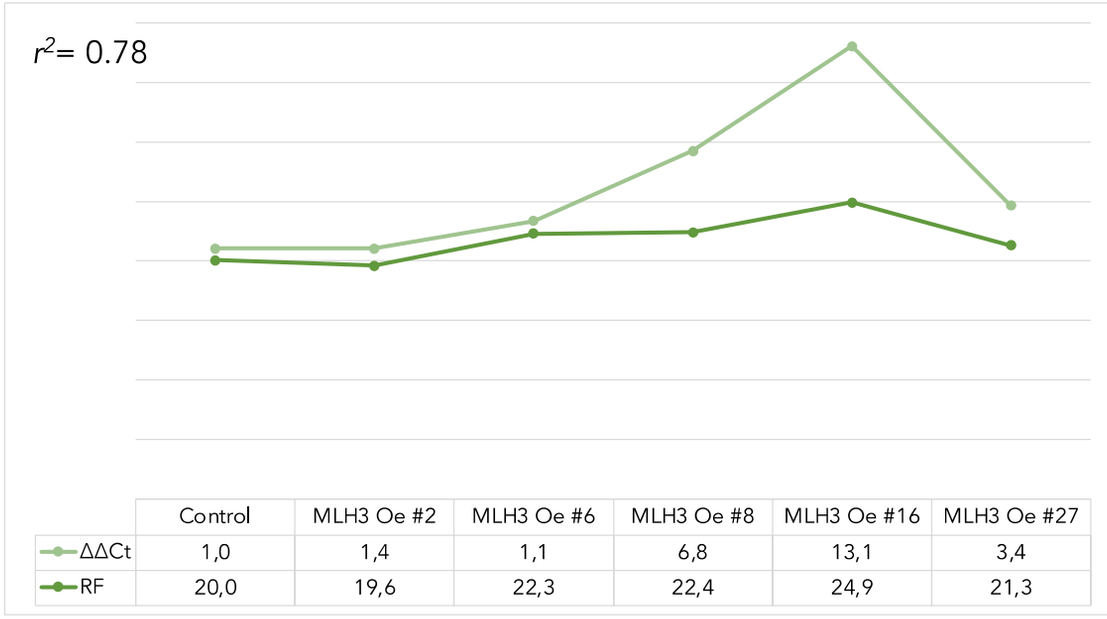
	Control	MLH1 Oe #3	MLH1 Oe #4	MLH1 Oe #7	MLH1 Oe #11	MLH1 Oe #12
ΔΔCt	1	4,57	10,10	2,55	2,58	2,46
RF	20,0	22,0	19,8	22,3	20,4	19,1

Supplemental figure 3. Quantitative assessment of *MLH1* expression level in five T2 generation *pMLH1::MLH1* overexpression lines. A. quantitative expression level of *MLH1*, its partner *MLH3* and two meiotic genes *DMC1* and *HEI10*. *MLH1* is significantly overexpressed in all tested lines. The other genes do not show any significant changes. B. *MLH1* expression level does not correlate with crossover recombination frequency in the 420 interval, r^2= 0.01. The RT-qPCR data represents n=3 biological replicates. Each biological replicate received two technical replicates. The plotted data was standardized to the housekeeping gene *KUP9* and normalized to the wildtype controls. *** $P < 0.001$.

A



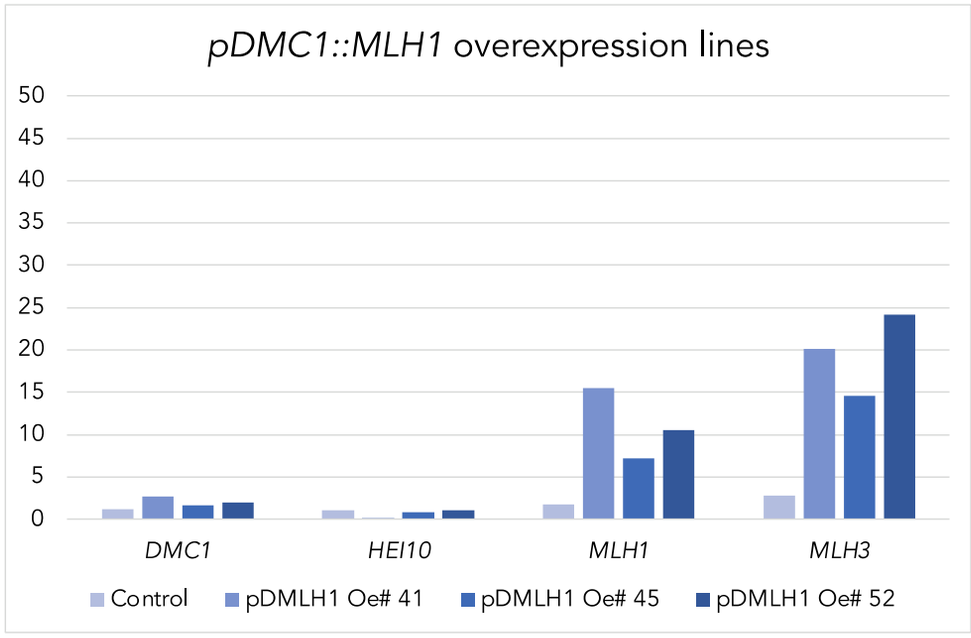
B



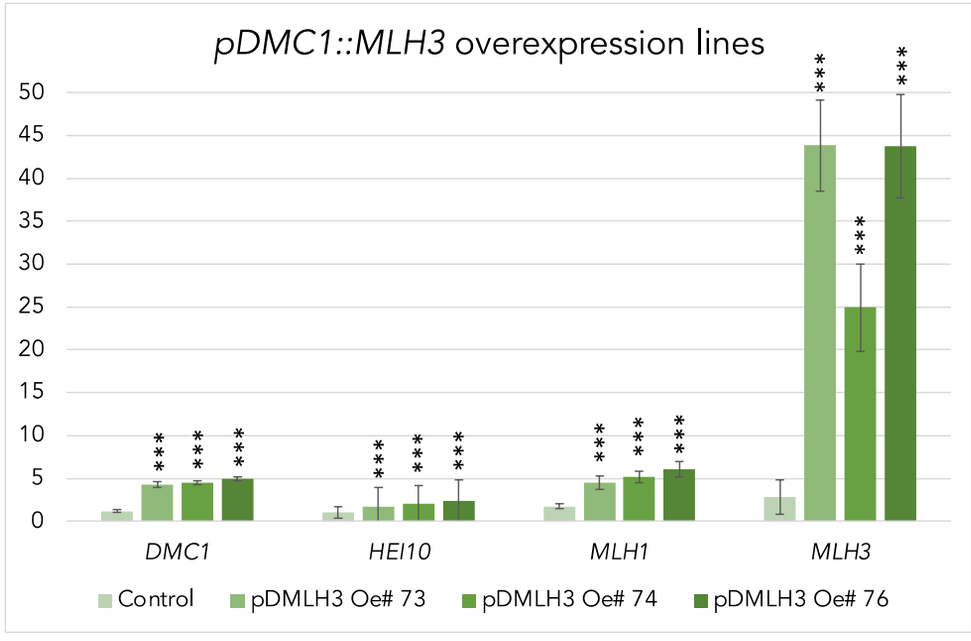
	Control	MLH3 Oe #2	MLH3 Oe #6	MLH3 Oe #8	MLH3 Oe #16	MLH3 Oe #27
ΔΔCt	1,0	1,4	1,1	6,8	13,1	3,4
RF	20,0	19,6	22,3	22,4	24,9	21,3

Supplemental figure 4. Quantitative assessment of *MLH3* expression level in five T2 generation *pMLH3::MLH3* overexpression lines. A. Quantitative expression level of *MLH3*, its partner *MLH1* and two meiotic genes *DMC1* and *HEI10*. *MLH3* is significantly overexpressed in 3/5 tested lines. The other genes do not show any significant changes. B. *MLH3* expression level correlates positively with crossover recombination frequency in the 420 interval, r^2= 0.78. The RT-qPCR data represents n=3 biological replicates. Each biological replicate received two technical replicates. The plotted data was standardized to the housekeeping gene *KUP9* and normalized to the wildtype controls. *** $P < 0.001$.

A

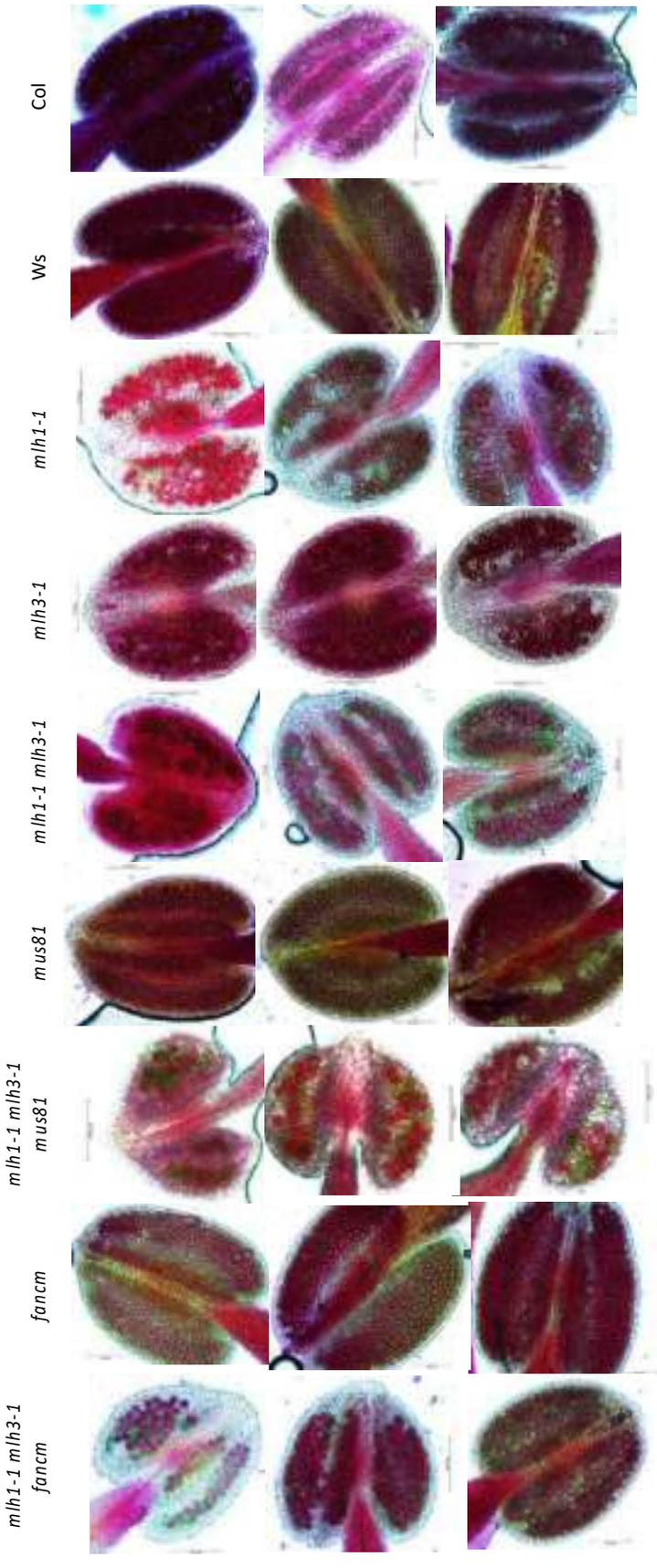


B

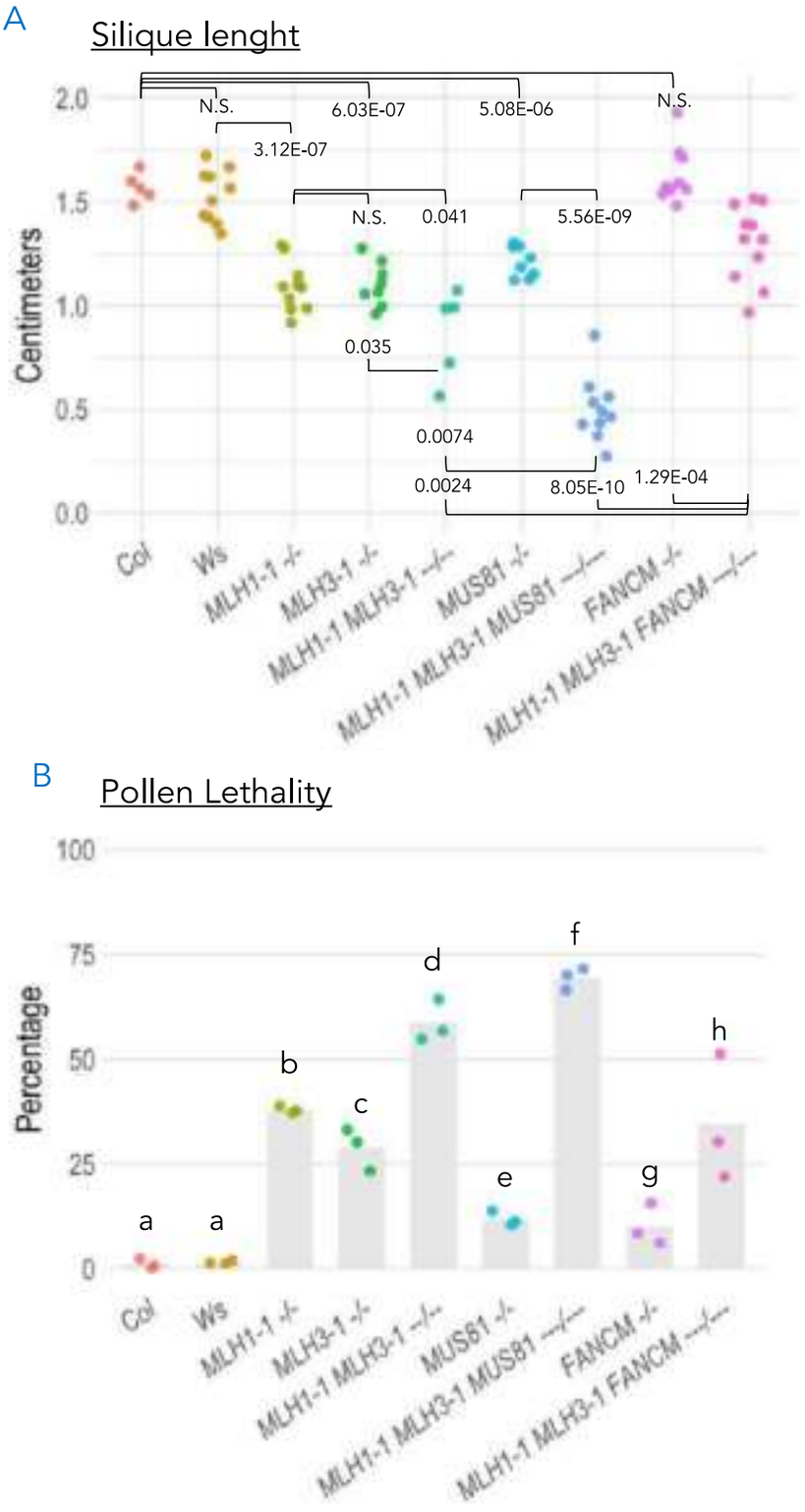


Supplemental figure 5. Quantitative assessment of *MLH1* and *MLH3* expression levels in three T2 generation *pDMC1::MLH1* and *pDMC1::MLH3* overexpression lines. Expression levels of *MLH1*, its partner *MLH3* and two meiotic genes *DMC1* and *HEI10* in: A. *pDMC1::MLH1*. *MLH1* and *MLH3* are overexpressed in all tested lines. The other genes do not show any major changes. B. *pDMC1::MLH3*. All tested genes are overexpressed in all tested lines. The RT-qPCR data represents n=3 biological replicates. Each biological replicate received three technical replicates. The plotted data was standardized to the housekeeping gene *KUP9* and normalized to the wildtype controls. *pDMC1::MLH1* data represents only one biological replicate. *** $P < 0.001$.

8.3 Fertility assessment supplemental data



Supplemental figure 6. Pollen density assessment for *MLH1-1, MLH3-1, MUS81, FANCM*, and their combined multiple mutants.



Supplemental figure 7. Silique length and pollen lethality comparative assessment of *MLH1-1, MLH3-1, MUS81, FANCM,* and their combined multiple mutants. (A) Silique length in centimeters. T-test values, with a 5% accepted error, are represented between samples connected with brackets. N.S.= Not Significant, p >0.05. n~10 (B) Pollen lethality in percentage. Different letters denote statistically significant differences according to a Chi test (p < 0.05), n=3. Detailed values are presented in Supplemental table 1.

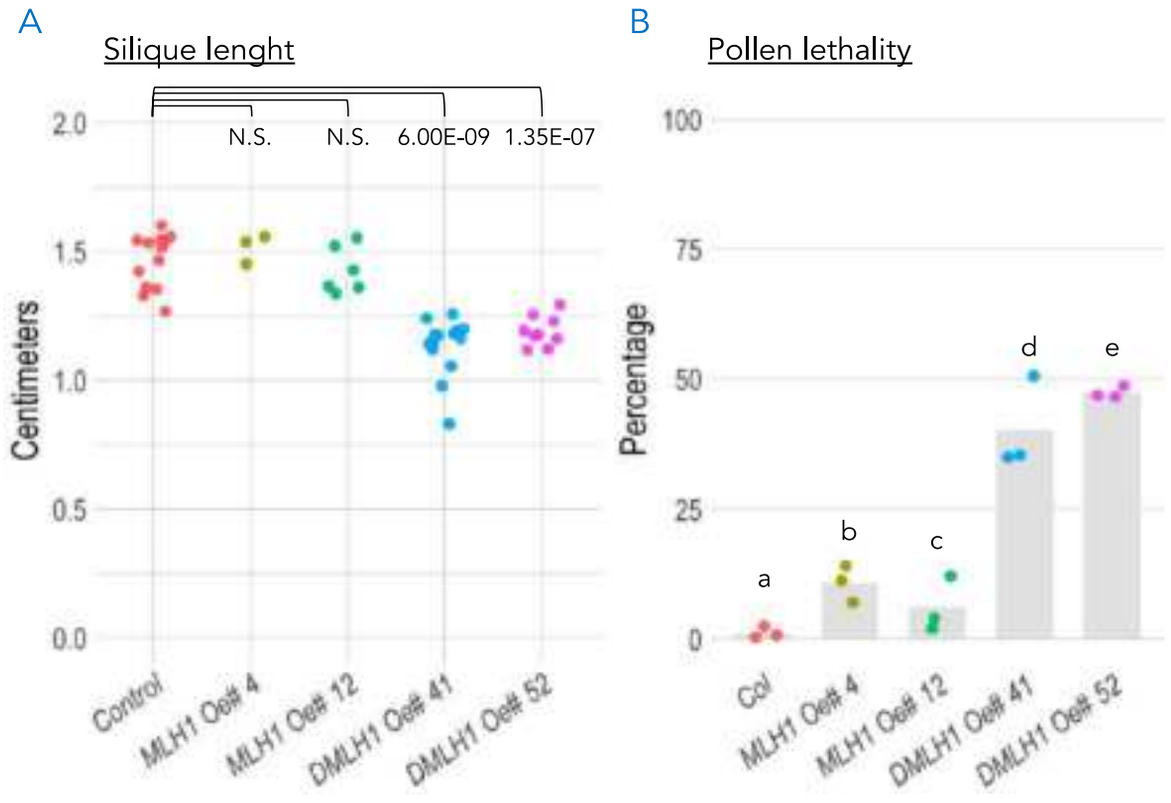
Supplemental table 1. Detailed cross Chi-test values for pollen viability and lethality in *MLH1-1, MLH3-1, MUS81, FANCM,* and their combined multiple mutants. Red $p > 0.05$, yellow $p < 0.05$, green $p < 0.001$.

Viable pollen

	Col	Ws	MLH1-1 -/-	MLH3-1 -/-	MLH1-1 MLH3-1 --/--	MUS81 -/-	MLH1-1 MLH3-1 MUS81 ---/---	FANCM -/-	MLH1-1 MLH3-1 FANCM ---/---
Col									
Ws	0,98								
MLH1-1 -/-	5,44E-15	1,08E-14							
MLH3-1 -/-	5,01E-08	5,01E-08	0,142						
MLH1-1 MLH3-1 --/--	7,07E-55	8,24E-54	3,04E-08	8,39E-17					
MUS81 -/-	0,128	0,152	9,43E-06	0,005	4,67E-17				
MLH1-1 MLH3-1 MUS81 ---/---	4,34E-101	3,81E-100	3,84E-22	7,02E-36	3,99E-04	6,10E-72			
FANCM -/-	0,213	0,205	2,04E-06	0,003	4,35E-18	0,638	1,59E-76		
MLH1-1 MLH3-1 FANCM ---/---	5,17E-16	2,27E-15	0,014	2,07E-04	9,41E-07	3,05E-09	1,26E-32	5,46E-11	

Non-Viable pollen

	Col	Ws	MLH1-1 -/-	MLH3-1 -/-	MLH1 MLH3-1 --/--	MUS81 -/-	MLH1 MLH3-1 MUS81 ---/---	FANCM -/-	MLH1 MLH3-1 FANCM ---/---
Col									
Ws	0,277								
MLH1-1 -/-	4,75E-24	1,50E-23							
MLH3-1 -/-	3,60E-18	1,10E-17	3,38E-03						
MLH1-1 MLH3-1 --/--	1,49E-37	4,73E-37	1,32E-05	5,63E-11					
MUS81 -/-	3,73E-07	1,17E-06	7,28E-39	2,27E-18	1,03E-124				
MLH1-1 MLH3-1 MUS81 ---/---	1,50E-44	4,75E-44	4,88E-10	2,29E-39	3,97E-02	1,23E-185			
FANCM -/-	6,05E-06	1,52E-05	9,33E-66	1,01E-27	5,43E-203	0,032	5,98E-34		
MLH1-1 MLH3-1 FANCM ---/---	7,97E-22	2,56E-21	8,91E-05	9,40E-05	8,36E-18	5,88E-11	1,78E-37	1,03E-12	



Supplemental figure 8. Silique length and pollen lethality comparative assessment of *MLH1* overexpression lines under its endogenous promoter and *DMC1* promoter. (A) Silique length in centimeters. T-test values, with a 5% accepted error, are represented between samples connected with brackets. N.S.= Not Significant, p >0.05. n~10 (B) Pollen lethality in percentage. Different letters denote statistically significant differences according to a Chi test (p < 0.05), n=3. Detailed values are presented in **Supplemental table 2**.

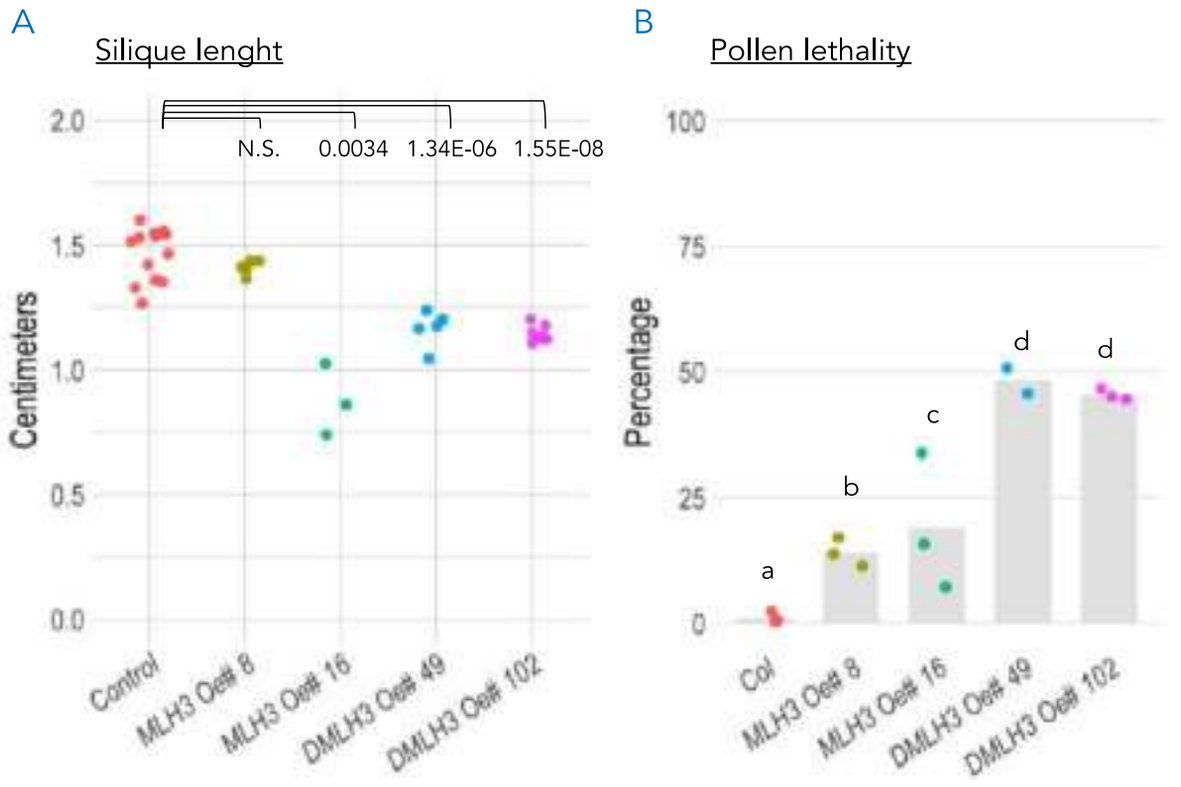
Supplemental table 2. Detailed cross Chi-test values for pollen viability and lethality of *MLH1* overexpression lines under its endogenous promoter and *DMC1* promoter. Red $p > 0.05$, yellow $p < 0.05$, green $p < 0.001$.

Pollen viability

	Col	*MLH1 Oe# 4*	*MLH1 Oe# 12*	*DMLH1 Oe# 41*	*DMLH1 Oe# 52*
Col		0,2237	0,5742	3,87E-11	5,25E-15
MLH1 Oe# 4			0,6379	2,37E-07	1,04E-10
MLH1 Oe# 12				5,09E-09	8,16E-13
DMLH1 Oe# 41					0,0599
DMLH1 Oe# 52					

Pollen lethality

	Col	*MLH1 Oe# 4*	*MLH1 Oe# 12*	*DMLH1 Oe# 41*	*DMLH1 Oe# 52*
Col		0,00	0,00	0,00	0,00
MLH1 Oe# 4			0,0038	5,88E-71	3,02E-76
MLH1 Oe# 12				4,39E-148	4,87E-258
DMLH1 Oe# 41					0,0691
DMLH1 Oe# 52					



Supplemental figure 9. Silique length and pollen lethality comparative assessment of *MLH3* overexpression lines under its endogenous promoter and *DMC1* promoter. (A) Silique length in centimeters. T-test values, with a 5% accepted error, are represented between samples connected with brackets. N.S.= Not Significant, p >0.05. n~10 (B) Pollen lethality in percentage. Different letters denote statistically significant differences according to a Chi test (p < 0.05), n=3. Detailed values are presented in **Supplemental table 3**.

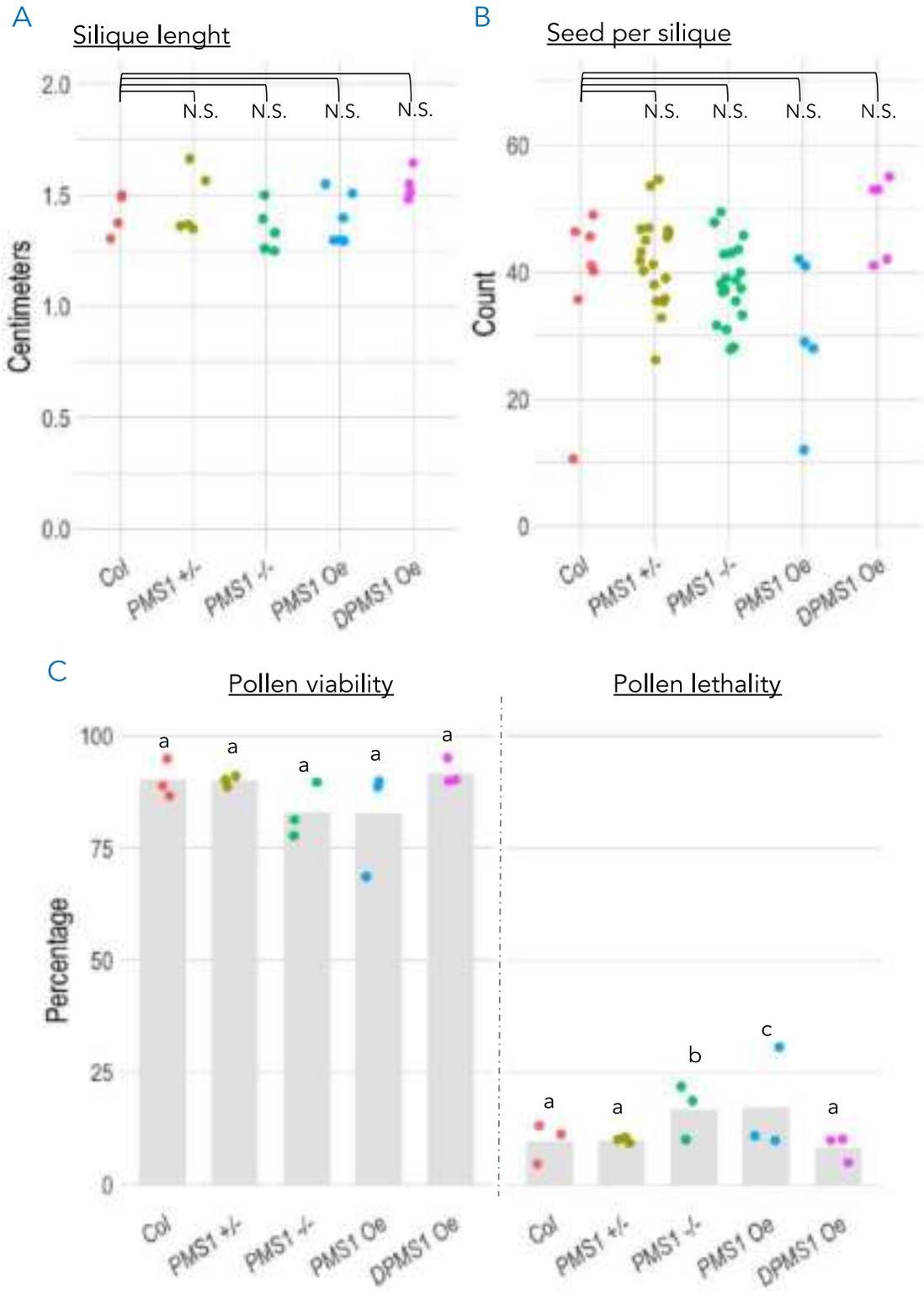
Supplemental table 3. Detailed cross Chi-test values for pollen viability and lethality of *MLH1* overexpression lines under its endogenous promoter and *DMC1* promoter. Red $p > 0.05$, yellow $p < 0.05$, green $p < 0.001$.

Pollen viability

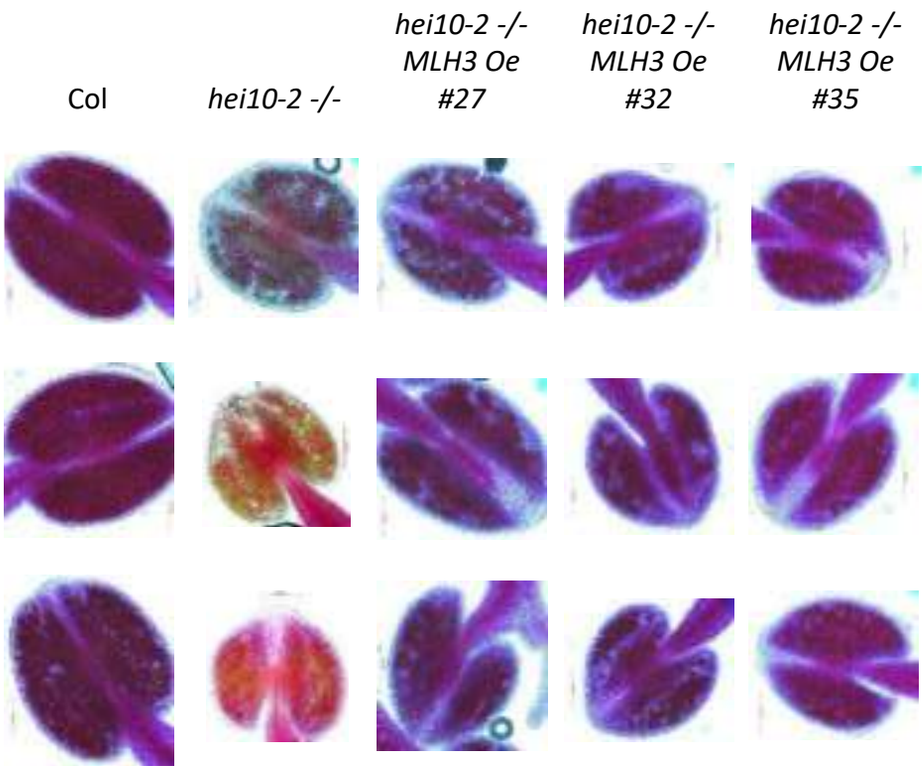
	Col	*MLH3 Oe# 8*	*MLH3 Oe# 16*	*DMLH3 Oe# 49*	*DMLH3 Oe# 102*
Col		1,15E-22	5,31E-23	4,76E-27	1,62E-26
MLH3 Oe# 8			4,68E-20	5,70E-23	1,60E-22
MLH3 Oe# 16				5,20E-03	1,31E-02
DMLH3 Oe# 49					0,9321
DMLH3 Oe# 102					

Pollen lethality

	Col	*MLH3 Oe# 8*	*MLH3 Oe# 16*	*DMLH3 Oe# 49*	*DMLH3 Oe# 102*
Col		0,00E+00	7,23E-246	0,00	0,00
MLH3 Oe# 8			2,23E-02	6,83E-41	4,47E-33
MLH3 Oe# 16				4,93E-62	5,05E-59
DMLH3 Oe# 49					0,6858
DMLH3 Oe# 102					



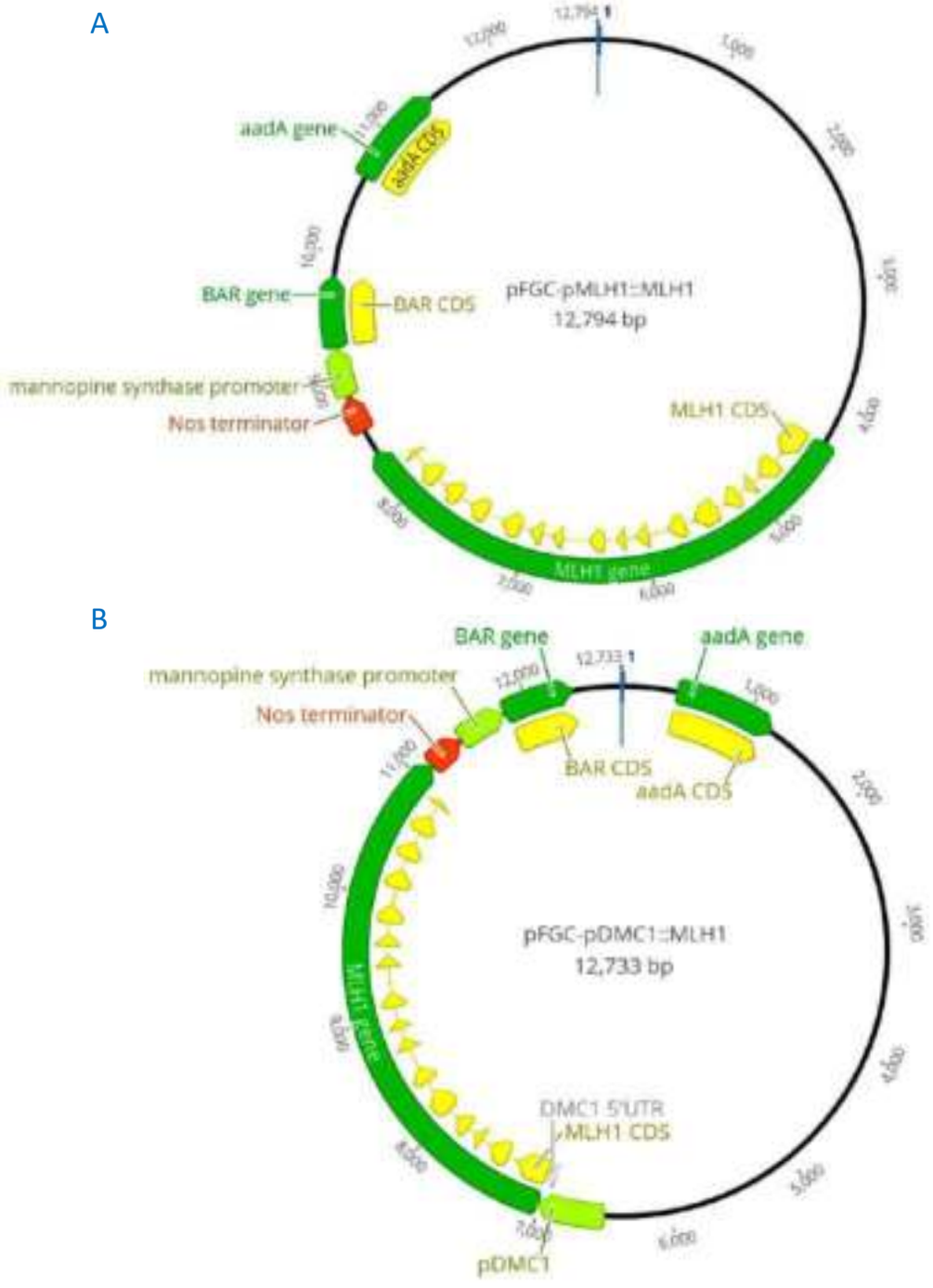
Supplemental figure 10. Silique length and pollen viability comparative assessment of PMS1 at different expression levels. (A) Silique length in centimeters. T-test values, with a 5% accepted error, are represented between samples connected with brackets. N.S.= Not Significant, p >0.05. n~10. (B) Seed per silique count. T-test values, with a 5% accepted error, are represented between samples connected with brackets. N.S.= Not Significant, p >0.05. (C) Pollen viability and lethality in percentage. Different letters denote statistically significant differences according to a Chi test ($p < 0.05$), n=3. This data was generated by MSc. Olga Maria Wienskowska for her Master thesis.



Supplemental figure 11. Pollen density assessment for *hei10-2* in combination with *MLH3 overexpression*. Anthers were dissected from stage 12 flower buds, discolored using Carnoy fixative, and colored with Alexander staining. They were then mounted and pictured at a 20X magnification.

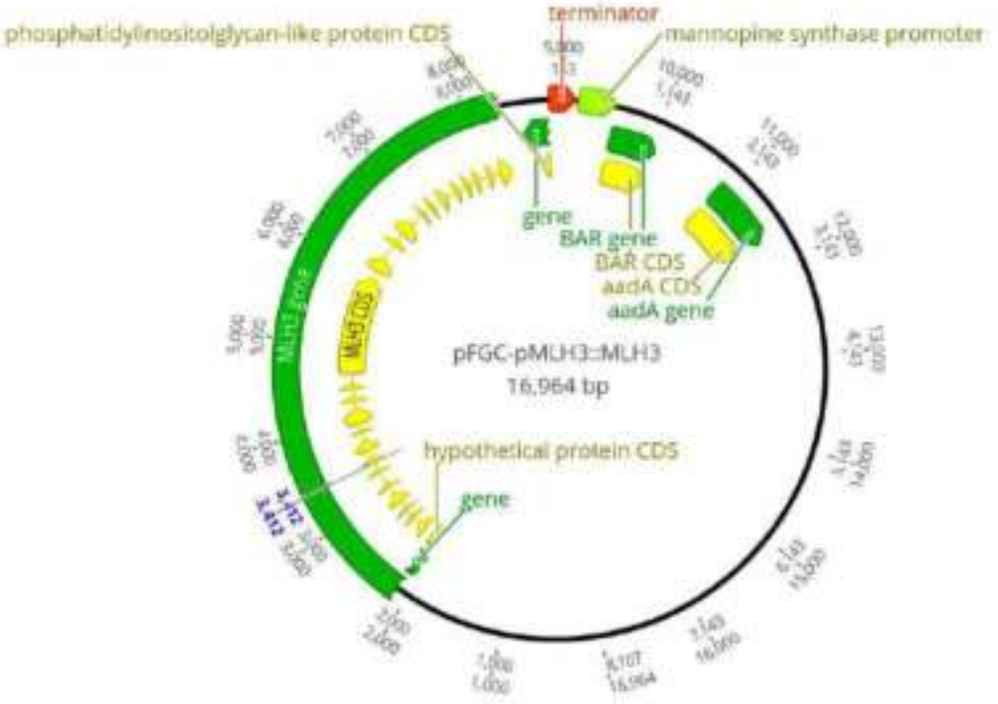
Supplemental figure 12. Representative picture of *hei10-2 -/-* and *hei10-2 -/- MLH3 Oe* where the fertility phenotype is improved.

8.4 Overexpression constructs maps

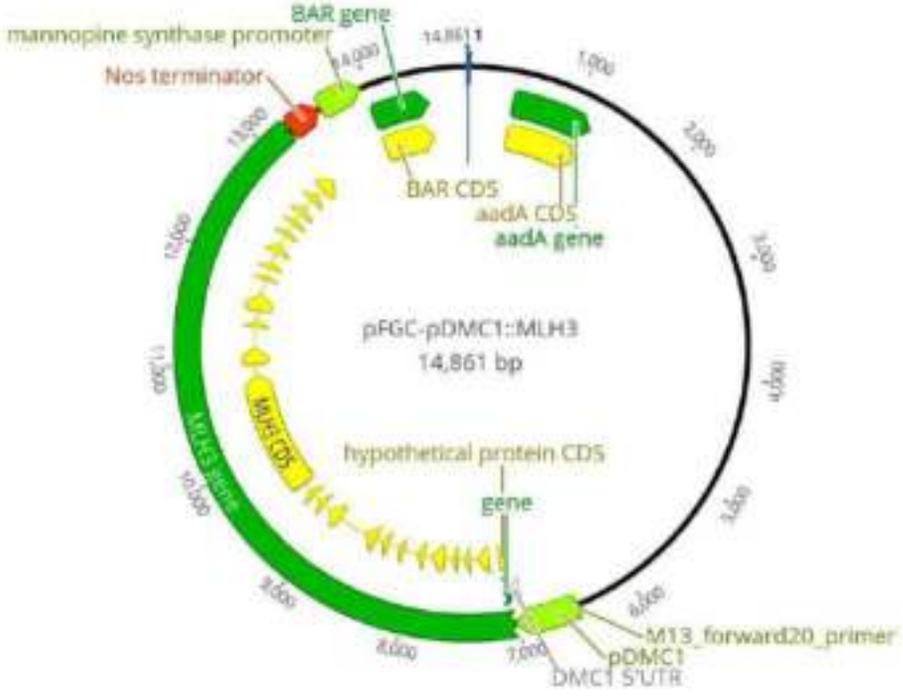


Supplemental figure 13. *MLH1* overexpression constructs. A. *MLH1* under the control of its endogenous promoter. B. *MLH1* under the control of the meiosis-specific *DMC1* promoter. pFGC confers a Kanamycin resistance to the transformant bacteria and a BASTA resistance to the transformant plant subsequently.

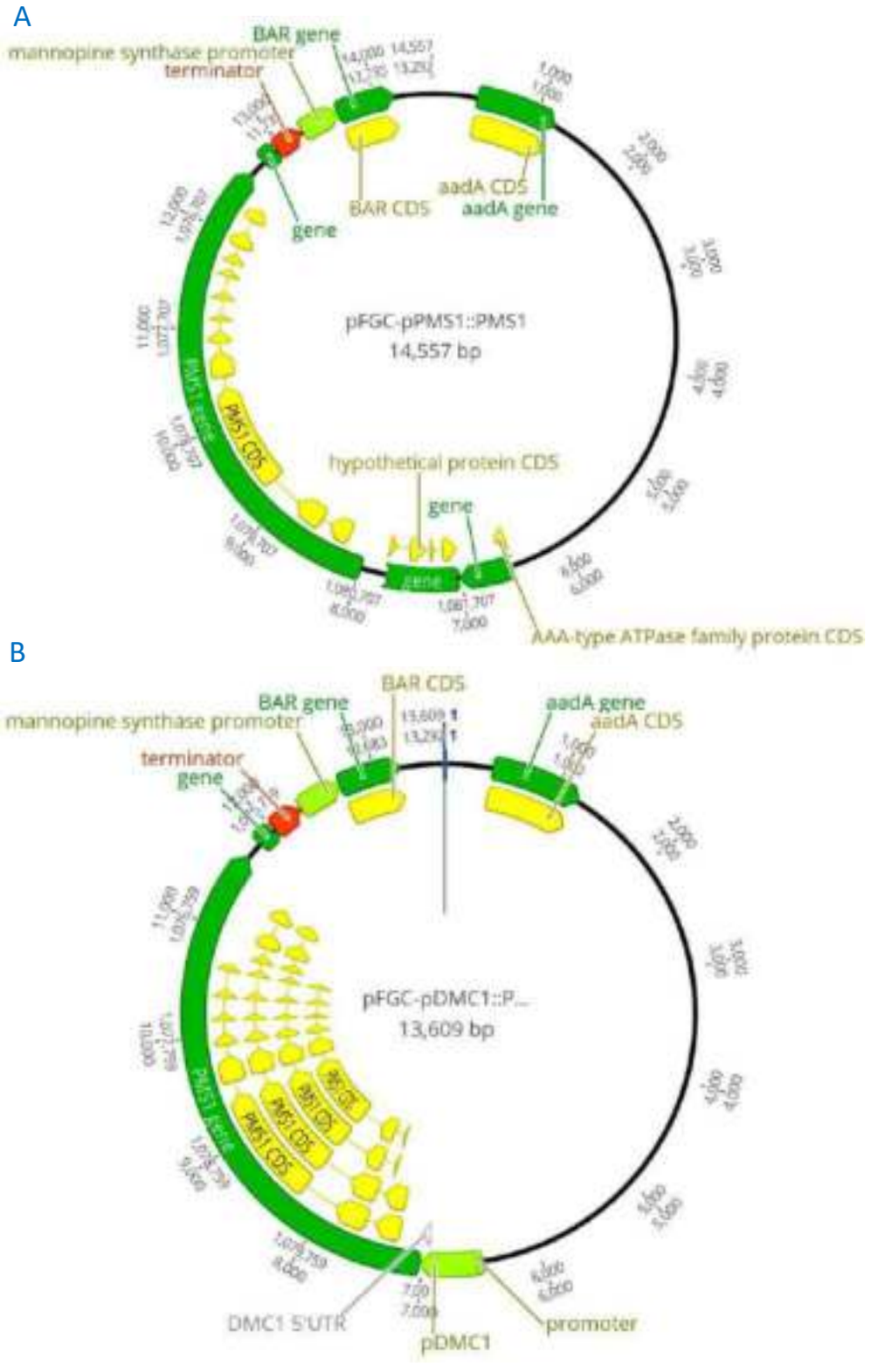
A



B

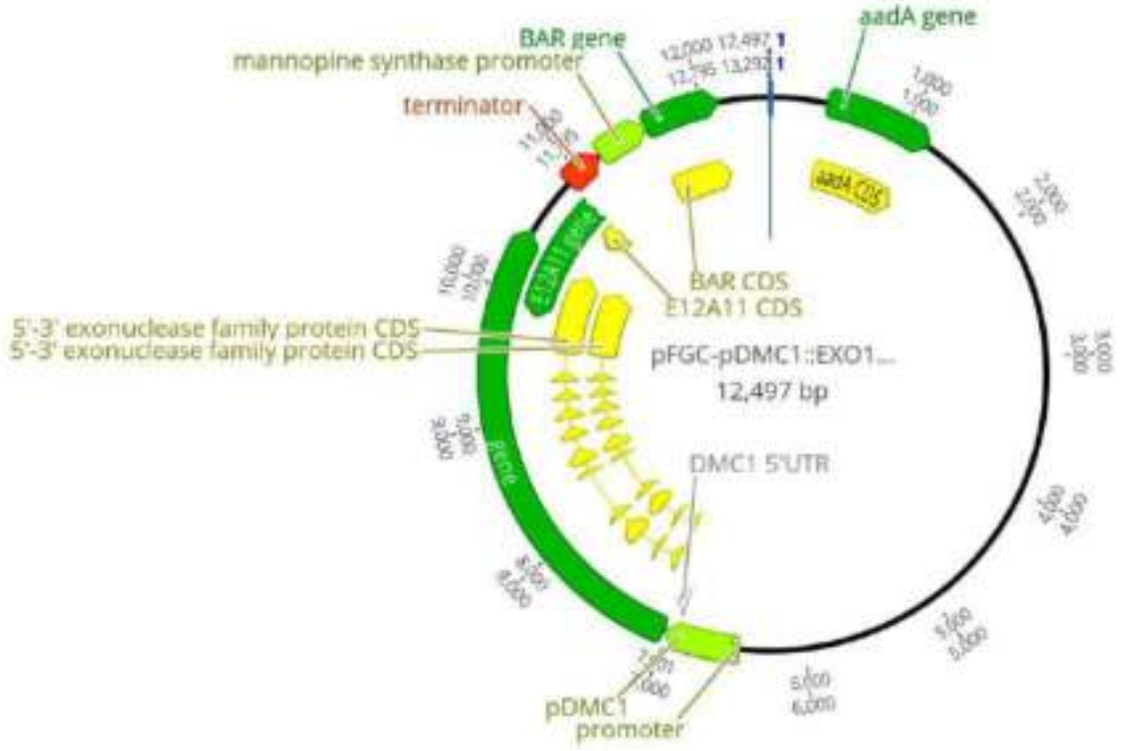


Supplemental figure 14. *MLH3* overexpression constructs. A. *MLH3* under the control of its endogenous promoter. B. *MLH3* under the control of the meiosis-specific *DMC1* promoter. pFGC confers a Kanamycin resistance to the transformant bacteria and a BASTA resistance to the transformant plant subsequently.

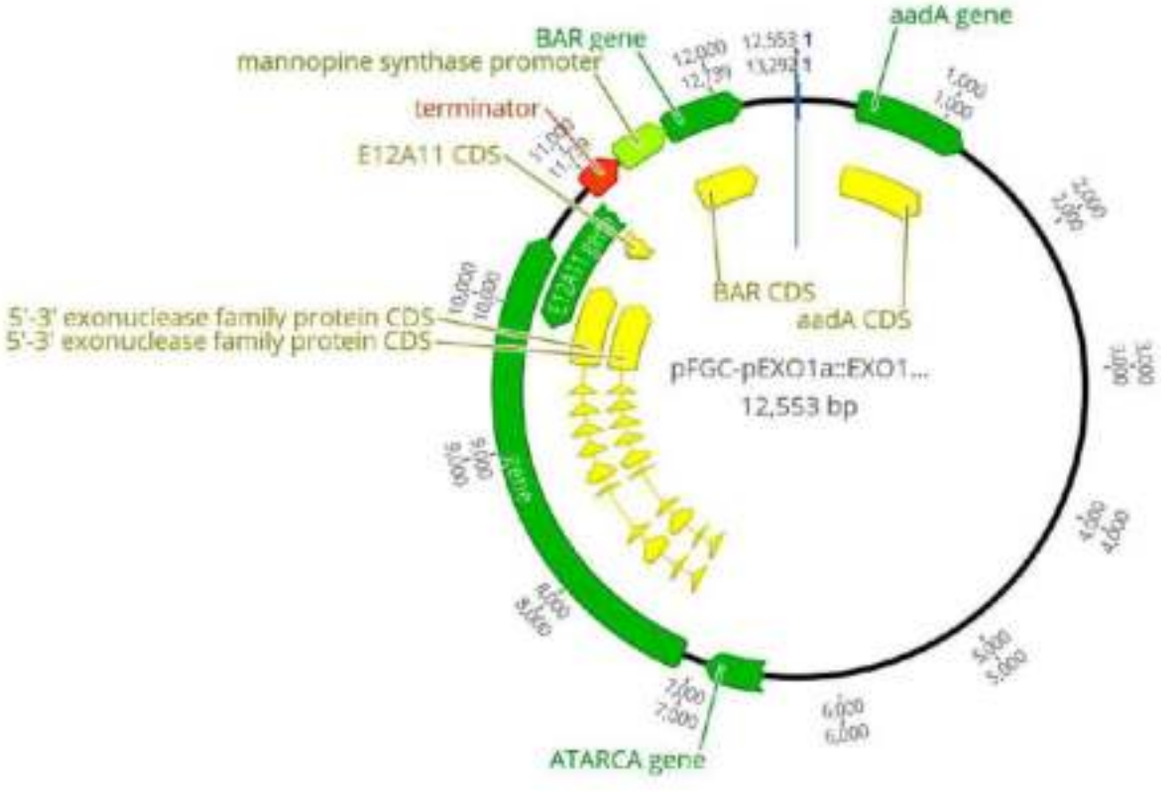


Supplemental figure 15. *PMS1* overexpression constructs. A. *PMS1* under the control of its endogenous promoter. B. *PMS1* under the control of the meiosis-specific *DMC1* promoter. pFGC confers a Kanamycin resistance to the transformant bacteria and a BASTA resistance to the transformant plant subsequently.

A

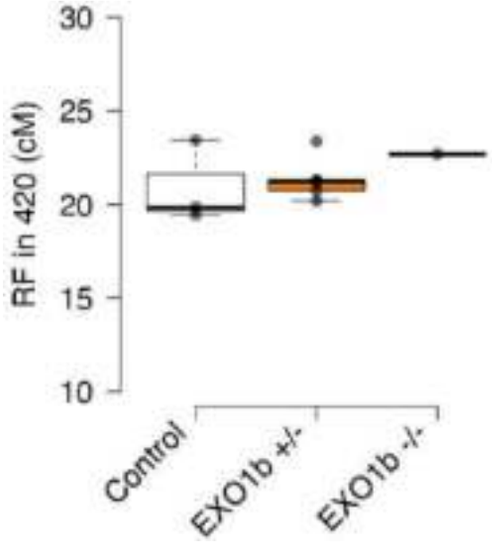


B



Supplemental figure 16. *EXO1* overexpression constructs. A. *EXO1* under the control of its endogenous promoter. B. *EXO1* under the control of the meiosis-specific *DMC1* promoter. pFGC confers a Kanamycin resistance to the transformant bacteria and a BASTA resistance to the transformant plant subsequently.

8.5 Recombination frequency assessment supplemental data



Supplemental figure 17. Recombination frequency measurement for *EXO1a* mutant. Recombination frequency measurement in the 420 subtelomeric interval of chromosome 3. No statistical assessment was conducted because of the low number of data points.

8.6 Primer tables

Supplemental table 4. Genotyping primers

Name	Target	Sequence
BAR-prom-R1	*aaAd*	CCATGTCCTACACGCCGAAA
EXO1b-Geno-F1	*EXO1b*	CGACAAAGAGAGTGCGTGGA
EXO1b-Geno-R1	*EXO1b*	AAGCATCGATTCCCACCTGG
ExoA-1	*EXO1a*	CATTCCCGTCCTTCAGATTCGTA
ExoA-2	*EXO1a*	GGACCTCCATCAAAGACCATGAT
ExoB-1	*EXO1b*	GCTCATGCATTCATCTCCAAGTA
ExoB-2	*EXO1b*	CCTTCAGCAATTGCAACAGCAA
Fancm dCAPS-F	*FANCM*	ACAATATATGTTTCGTGCAGGTAAGACATTGGAAG
Fancm dCAPS-R	*FANCM*	CACCAATAGATGTTGCGACAAT
fancm-MP-LP	*FANCM*	GGATCTAGGGTTCCAATAG
fancm-MP-RP	*FANCM*	CCTCAATCTGCTGCATCAC
GABI-08474	GK T-DNA	ATAATAACGCTGCGGACATCTACATTTT
GABI-1	GK T-DNA	GATGTTAGGCCAGGACTTTGAA
Geno M13OX F1	pFGC / MLH	CGTTTCACTTTGGTGGTCTGTACC
Geno M13OX F2	pFGC / MLH	GCCGTCGTTTAGCTAAACCCTAAC
Geno M13OX F3	pFGC / MLH	CAAGGAGTTTCTGCAGCTATTGGG
Geno M13OX R1	pFGC / MLH	TCTTGCTGTAAAGCGTTGTTTGGT
Geno M13OX R2	pFGC / MLH	TGGATTACTTCACCAGCTGCGATA
GK-mlh1_L	*MLH1*	CTCCTGTGACTCCTCTGGTTG
GK-mlh1_R	*MLH1*	GTTCTTTTGCGAGCATACCTG
hei10-2 For	*HEI10*	AAGGAGTTCCCAGAGATGCTC
hei10-2 Rev	*HEI11*	GCCAGCAAGACAGAACAGTTC
LBb1.3	SALK T-DNA	ATTTTGCCGATTTCGGAAC
LBc-1	SALK T-DNA	TGGACCGCTTGCTGCAACTCT
LBd-1	SALK T-DNA	GAACCACCATCAAACAGGATTT
M13-F (-20)	Plasmid	GTAAAACGACGGCCAGT
MLH1_0F	*MLH1*	CACCGAAGATTCAACGCTTAGAAG
MLH1_1F	*MLH1*	TCGAAGCTGACTAAGTTTGAGGA
MLH1_1R	*MLH1*	GCAGAGCATTCCACCAATCTATC
MLH1_2F	*MLH1*	CTGTGACTCCTCTGGTTGTACTT
MLH1_2R	*MLH1*	GGAACTTTCTGTGTCTTTTGTCCT
MLH1_3F	*MLH1*	TTGCATGCTACAGAAAGTGGAAT
MLH1_3R	*MLH1*	GCTGCAAAACCCAATAAGATGGT
MLH1_4F	*MLH1*	AGGTATGCTGGAGACTGTAAGGA
MLH1_4R	*MLH1*	ATAGAACTGAATACCGTCACCCG
MLH1_5R	*MLH1*	AAAAGTCCCATTTGAAGCCATGG

MLH1_del1 F	*MLH1*	CTTCAGGTTCTTTTGCGAGCA
MLH1_del2 R	*MLH1*	CGGGGGAAACGATTTTCTTCG
MLH1-2	*MLH1*	TCCGCTCGAGTTAGCATCGTTCGAATATCTTGTACAG
MLH1-538 F	*MLH1*	CGCGGATCCATGATTGCTAGAAGGAAGACACTTCA
mlh1-bel-F	*MLH1*	TTGTGCCCATGCGTTTTCAG
mlh1-bel-R	*MLH1*	AGGAGTATTTCAGCGTGCACA
MLH3_del1 F	*MLH1*	TGCTCCACTTGTGGGATTCAA
MLH3_del2 R	*MLH1*	ATAGCTTCTGGCCAATCTGCA
mlh3-1_L	*MLH3*	CGAAGCTGTAAATTCGCTTTG
mlh3-1_R	*MLH3*	ATACCTTGAACTCAACGTGCG
mlh3-LP2	*MLH3*	CAAAACTTTCTTGGGGCTACC
mlh3-LP3	*MLH3*	ATGCATGGAACCTACAAGTGG
mlh3-MP-LP	*MLH3*	GTAGCCCCAAGAAAGTTTTGG
mlh3-MP-RP	*MLH3*	GCCTAGGAATGTCAAAGGGAC
mlh3-RP2	*MLH3*	GATCAGGCGTTTCAAGAGATG
mlh3-RP3	*MLH3*	TTACGATCCGATGAATCCTTG
mus81-MP-LP	*MUS81*	GACAGTTGAAGGTCGGGAAG
mus81-MP-RP	*MUS82*	AATTTTCCACAAACCCTTTGG
NST-R	NOS terminator	ACCGGCAACAGGATTCAATCTTA
pJET for	pJet	CGACTCACTATAGGGAGAGCGGC
pJET Rev	pJet	CTGCCATGGAAAATCGATGTTCTT
R-MLH3ck	*MLH3*	TCTCCACGTTGGTGAAGTCG
R-MLH3cs	*MLH3*	CAGGAACTGCGTCCTCCATT
R-MLH3qf	*MLH3*	ATTTGAAAAATCTCAGAATTCCAGGAACTGCGTCCTCCATT
T3 promoter (20pb)	Plasmid	ATTAACCCTCACTAAAGGGA

Supplemental table 5. Cloning primers

Name	Target	Sequence
EXO1a-DA-F	*EXO1a*	ATCACTGAGGATTCTGCTCTCATACC
EXO1a-DA-R	*EXO1a*	GGTATGAGAGCAGAATCCTCAGTGAT
EXO1b-DA-F	*EXO1b*	ATAACCGAAGACAGCGCTTTACTT
EXO1b-DA-R	*EXO1b*	ATATGCAAGTAAAGCGCTGTCTTC
fgPMS1-1F	*PMS1*	ACGACGGCCAGTGCCAAGCTTCTAGAGTATGCGCAAGTGTGTCTTC
fgPMS1-R	*PMS1*	TCTATCGATCAATCAGGATCCTCATGTTTACTGGAAAACTGTTG
gbPMS1 For2	*PMS1*	GAGAGTCTTAGAGCGAGAAATCTAGAATGCAAGGAGATTCTTCTCCGT

gbPMS1 Rev2	*PMS1*	TGATTGATCGATAGAGCTCGGCGGCCGCTCATGCCAATGAGATGGTTGC
gbPMS1 Rev3	*PMS1*	ACGGAGAAGAATCTCCTTGCAT
gbPMS1-2F	*PMS1*	TCTTAGAGCGAGAAATCTAGAGCATTGTGCAGCTGCGTC
gEXO1a-F	*EXO1a*	CAACTGTTGACGACGACGAC
gEXO1a-R	*EXO1a*	GGTCAAATGGCGTTTTCCGT
gEXO1b-1F	*EXO1b*	TAAAACGACGGCCAGTGCCATAACTGGAGCGCTGATGGAAG
gEXO1b-1R	*EXO1b*	TTCCATACAGAAGTTGGGGAAACA
gEXO1b-2F	*EXO1b*	TGTATGATATGTTTCCCCAACTTC
gEXO1b-2R	*EXO1b*	TCTATCGATCAATCAGGATCATACATGGAACCAAGGCCACC
gMLH3 For 1	*MLH3*	TAAAACGACGGCCAGTGCCAAGATTCACTTCTTCTTGGAAGGATTTTG
gMLH3 rev 1	*MLH3*	CGAGCTCTATCGATCAATCACAGGAACTGCGTCCTCCATTG
gPMS1 For 1	*PMS1*	TAAAACGACGGCCAGTGCCAGTATGCGCAAGTGTGTCTTCAG
gPMS1 Rev 1	*PMS1*	GAGCTCTATCGATCAATCACTCATGTTTACTGGAAAACTGTTGAAG
gMLH1-1F	*MLH1*	GCGTAGCAGCTTGAGATACTCCAATTTGTATCTGGCGCGAG
gMLH1-F6	*MLH1*	AATCGGACGTCCCAATTTGTATCTGGCGCGAG
gMLH1-F7	*MLH1*	CCTTAATTAAGGTTAATTAAGGCCAATTTGTATCTGGCGCGAG
gMLH1-R6	*MLH1*	TCCCCCCGGGGGGACCCGGGGGGACGAGCCATATCTACGTCGCT
gMLH1-R7	*MLH1*	CCTTAATTAAGGTTAATTAAGGCGAGCCATATCTACGTCGCT
pDMC1 F	*DMC1*	ACGACGGCCAGTGCCAAGCTTAAAATTAATTTGATTAGTGGATCCGC
pDMC1 R	*DMC2*	TTTCTCGCTCTAAGACTCTCTAAG
gPMS1 1For	*PMS1*	TCCGCTCGAGTCATGCCAATGAGATGGTTGCATCAT
gPMS1 1Rev	*PMS1*	CGCGGATCCATGCAAGGAGATTCTTCTCCGTCTCCG
gPMS1 2For	*PMS1*	TCCGCTCGAGTCACCTGATAATATCCATGTGCATTAACACAG
gPMS1 2Rev	*PMS1*	CGCGGATCCATGGTTGTGGCATTTCCCCAACCAA
R1VF	pJET-U3/6	AGAATTCCCATGGAAGGATCCTCGAGGCTGCAGGAATTCGATATCAAGC
R2	pJET-U3/6	CCATGATTACGCCAAGCTCG
R2F1	pJET-U3/6	CTTGGCGTAATCATGGGGATGGCTCGAGTTTTCAGC
VRF1	pJET-U3/6	ATGTTACTAGATCGGGGATCCGGATGGCTCGAGTTTTCAGC

Supplemental table 6. Sequencing primers

Name	Target	Sequence
MLH1_0F	*MLH1*	CACCGAAGATTCAACGCTTAGAAG
MLH1_1F	*MLH1*	TCGAAGCTGACTAAGTTTGAGGA
MLH1_1R	*MLH1*	GCAGAGCATTCCACCAATCTATC

MLH1_2F	*MLH1*	CTGTGACTCCTCTGGTTGTACTT
MLH1_2R	*MLH1*	GGAACTTTCTGTGTCTTTTGTCCT
MLH1_3F	*MLH1*	TTGCATGCTACAGAAAGTGGAAT
MLH1_3R	*MLH1*	GCTGCAAAACCCAATAAGATGGT
MLH1_4F	*MLH1*	AGGTATGCTGGAGACTGTAAGGA
MLH1_4R	*MLH1*	ATAGAACTGAATACCGTCACCCG
MLH1_5R	*MLH1*	AAAAGTCCCATTTGAAGCCATGG
PMS1 seq 1	*PMS1*	GTGGTAACCGCATCAATCGC
PMS1 seq 3	*PMS2*	CGAAAGGAGTGCGGTTTGTC
PMS1 seq 4	*PMS3*	TCCACGAGACACCTTGAAGC
PMS1 seq 5	*PMS4*	CTCTATCCTGGCTCGGTCGA
PMS1 seq 6	*PMS5*	GATGCTAGCATTGCACGGAC
seq MLH1-1	*MLH1*	AGGTCTTCTGTAAGGCAAAGAAG
seq MLH1-1	*MLH1*	GCTGAGATTCACCACATTTGCTAG
seq MLH1-724 R	*MLH1*	CATATACAGACCTAATTGAATCAAGCCTTG
seqMLH3-1	*MLH3*	TGAAGACTTCCCACAAGTTACTGAC
seqMLH3-1	*MLH3*	TTCTCCAATTCCAACACATCCCC
seqMLH3-2	*MLH3*	AACCTGATGATCTGGAGTGTTTGA
seqMLH3-2	*MLH3*	GCACATCCTCATCTTGGGTCTC
seqMLH3-590 F	*MLH3*	GTGATGAAGAGCTTTTCCAAACCA
seqMLH3-991 R	*MLH3*	ACTTCTTGAACTCAACGTGCGT

Supplemental table 7. qPCR primers

Name	Traget	Sequence
qDMC1-1F	*DMC1*	ATCTGGAAACTGGCTTCAGCTT
qDMC1-1R	*DMC1*	CCACCATCAGGCTCTTGTTCA
qDMC1-2F	*DMC1*	GACATTTTTGCAGCAAGGCCA
qDMC1-2R	*DMC1*	AATGGATCCAGGAGCTGTGC
qDMC1-897 F	*DMC1*	GCTTTTTCACATCTCCTGCGTT
qDMC1-994 R	*DMC1*	GCTCGTTGAGCGTGAAGAAAAT
qEXO1a-1F	*EXO1a*	GAGTGCTACTCAAAGGCCGT
qEXO1a-1R	*EXO1a*	ATCTGCGCATCAGCTTCGTA
qEXO1a-2F	*EXO1a*	TCCTGATGACAACACACCAGA
qEXO1a-2R	*EXO1a*	GCAACTTCGCTAACACATGGA
qEXO1a-3F	*EXO1a*	GCAACAGGGAAGAGGAAGCT
qEXO1a-3R	*EXO1a*	TTTCATCCATGCGCATGTGC
qEXO1b_4095 F	*EXO1b*	GTCGTTGTTCTCGATGGTGGTA
qEXO1b_4287 R	*EXO1b*	CACCATTGCAGCGTCAAAGTT
qEXO1b_5728 F	*EXO1b*	CCTCTCCCACAACCTCCTGA
qEXO1b_6001 R	*EXO1b*	TACCTTCAGCAATTGCAACAGC

qHEI10-3F	*HEI10*	GGTGTCAGATGATGGAGCAAGA
qHEI10-3R	*HEI10*	TCTCATCAAGCTTCCTCTTCTGT
qHEI10-4F	*HEI10*	GGCACTGCTAATCCCCAGTC
qHEI10-4R	*HEI10*	TATAGCGTGAACAGCTGAGGG
qKUP9-368 F	*KUP9*	ATGGTCAAGGTGGGACTTTAGC
qKUP9-458 R	*KUP9*	TCCTCATCACTACGGTGCTGA
qMLH1-1351 F	*MLH1*	ACTGCTGATCTTTCTAGTGTCCAG
qMLH1-1441 R	*MLH1*	ATGTGCAATTCCTTACAGTCTCCA
qMLH1-3F	*MLH1*	TCCTCCATATCGACGTTACCC
qMLH1-3R	*MLH1*	CGCCATGCATCCTCCTCTTT
qMLH1-569 F	*MLH1*	CTGCTGATGATTACGGGAAAATCG
qMLH1-683 R	*MLH1*	ACTGAGTGAACATCAGCCTTAACA
qMLH3-1	*MLH3*	TTTCTCCACGAAAGGATGTATGGT
qMLH3-1	*MLH3*	GAGAATCAGTCCGGCTTTCAAATC
qMLH3-1F	*MLH3*	CAGGCGTTTCAAGAGATGATTTGG
qMLH3-1R	*MLH3*	AAGTTTCACTAGCTGTCTCCACG
qMLH3-2F	*MLH3*	GCAGATAAGAGACTGGGGTTGG
qMLH3-2R	*MLH3*	GATTGGTGTTGGTTTCCGCTG
qMLH3-345 F	*MLH3*	TATTGGGAGGCCTAATGGTTATCG
qMLH3-3F	*MLH3*	ACCCCCATAGATACTGCGGA
qMLH3-3R	*MLH3*	ACCAGTAACAGAGCTCCCCA
qMLH3-437 R	*MLH3*	GTCGTGCCAGAGTCTTTTCTATCA
qAct2-F	*Actin2*	TGCCAATCTACGAGGGTTTC
qAct2-R	*Actin2*	TTACAATTTCCCGCTCTGCT

8.7 Representation of the *MutL* genes mutant alleles

Supplemental figure 18. Representation of the MutL insertion and deletion mutants. Insertions are represented with red arrowheads and the deletion with red rectangles. A *MLH1* alleles: *mlh1-1* = insertion belzile, *mlh1-2* = GABI_067E10, *mlh1-3* = SK25975, and *mlh1-4* = 462bp deletion. B. *MLH3* alleles: *mlh3-1* = SALK_015849, *mlh3-2* = SALKseq_067953, and *mlh3-3* = SALseq_69853. C. PMS1 allele*: pms1-1* = SALK_124014C.