Charles University in Prague

Faculty of Science

Developmental and Cell Biology

Mgr. Václav Gergelits

Meiotic homologous recombination and hybrid sterility

Meiotická homologní rekombinace a hybridní sterilita

Doctoral thesis

Supervisor: Prof. MUDr. Jiří Forejt, DrSc.

Prague, 2020

Declaration

I declare that I wrote this work independently and that I did my best to acknowledge all people and literature. I did not use this work or a substantial part of it to obtain another academic degree or equivalent.

Prohlášení

Prohlašuji, že jsem závěrečnou práci zpracoval samostatně a že jsem uvedl všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

Prague, March 2020

Václav Gergelits

Acknowledgements

Hereby, I would like to express my thanks to my supervisor Jiří Forejt for discussions, sharing ideas, inspiration, and giving me an opportunity to observe how an excellent science can be done. Also, I appreciate to be stimulated and provoked as much as possible to become an independent and critically thinking scientist. I consider it an honour to study here.

I thank to my colleagues. In particular, to Petr Šimeček for initial introduction to the field of bioinformatics, valuable consultations later on, and his hospitality during my short visit of the Jackson Laboratory in USA. Also to Emil Parvanov for discussions we had, in particular about the noncrossover project.

I thank to my great friends for all the support during my Ph.D. studies and for being my friends, namely, to Olga, Ondřej, and František. Thanks a lot.

I am grateful to my family, especially my partner Klárka for all her continuous support, encouragement, patience, and love, and to my mother Šárka and my father Tomáš for supporting me during all my studies throughout whole my life. Thank you very much.

Contents

Abstract (English)

Meiotic homologous recombination, homologous chromosomes synapsis, and F1 hybrid sterility (enabling formation of species) are mutually interconnected phenomenons, one being the prerequisite to the latter. In the present thesis, these phenomenons were investigated on a genetic and mechanistic level using a mouse subspecies as a model.

Noncrossovers (NCOs, gene conversions), 90% prevalent resolution of *Prdm9*-determined meiotic double-strand breaks (DSBs), were uniquely identified and characterized on a chromosome-wide level. The mean gene conversion tract length, based on 94 NCOs events, was calculated to be 32 bp. On a local level, the NCOs overlapped the known hotspots of PRDM9-controlled histone trimethylation and DSB formation, indicating their origin in the standard meiotic DSB repair pathway. On chromosome-wide level, NCO and CO distributions differed, in particular COs being relatively preferred over NCOs in subtelomeric regions. A specific subset of nonparental/asymmetric NCOs and COs was underrepresented in our datasets, proposing their problematic repair, hypothetically enabled by sister chromatids, and thus not contributing to indispensable homologous synapsis.

Genome-wide crossover (CO) rates, genetically and mechanistically crucial ~10% of DSB repair, were proven to be genetically determined using crosses between laboratory strain representatives of *Mus musculus domesticus* (*M.m.m.*) (C57BL/6J, B6) and *Mus musculus domesticus* (*M.m.d.*) (PWD/PhJ, PWD) subspecies. A new genetic locus on Chr X, *Meiotic recombination Meir1,* was identified, having the largest effect on CO rate. Remarkably, *Meir1* genetically maps within the *Hstx2* locus, the known factor of F1 hybrid sterility. On the other hand the key speciation gene, *Prdm9*, did not affect CO rate, neither by adding, nor removing its copies, independently on genetic background.

While the *Prdm9* gene and *Hstx2* locus are the known necessary components controlling F1 hybrid sterility *in trans*, a likely *cis* component was still unknown. We found that randomly localized DNA homozygous stretches of at least 27 Mb were sufficient for proper synapsis. Importantly, when the threshold amounts of homozygosity were present in the four shortest, and thus most asynapsis-sensitive chromosomes, in an otherwise F1 hybrid sterility model, they were sufficient to rescue the fertility phenotype.

Furthermore, the X-localized *Hstx2* hybrid sterility locus, was narrowed down to the 2.7 Mb region, Chr X: 66.51-69.21 Mb. Importantly, it still overlapped with *Meir1* locus, controlling

genome-wide CO rate. The *Frm1nb* gene was excluded as a candidate for the second hybrid sterility gene, leaving now miRNA clusters, *miR-465* miRNA cluster in particular, a promising candidate.

Altogether, generally important meiotic recombination of NCO events of DSB repair were revealed chromosome-wide, suggesting possible consequences for homologous pairing. The concurrently studied F1 hybrid sterility architecture was established to have three components; 1) *trans*-acting *Prdm9* gene, 2) *cis*-acting intersubspecific homeology (as proposed asymmetric PRDM9-binding sites), and 3) *trans*-acting X-located *Hstx2* factor. Another group has suggested a mechanistic, molecular basis for F1 hybrid sterility (the *asymmetry of PRDM9 binding*).

Abstrakt (česky)

Meiotická homologní rekombinace, synapse homologních chromosomů a hybridní sterilita F1 kříženců jsou vzájemně provázané fenomény závisející jeden na druhém. V předložené dizertační práci jsme tyto fenomény zkoumali na genetické i mechanistické úrovni za využití modelu myších poddruhů.

Identifikovali a charakterizovali jsme noncrossovery (NCO, též genové konverze), které zprostředkovávají 90% všech oprav meiotických dvouřetězcových zlomů (DSB), determinovaných genem *Prdm9*, a to na celochromosomální úrovni. Na základě 94 detekovaných NCO jsme spočítali střední délku doprovodných úseků genové konverze jako 32 bp. Na lokální úrovni se NCO překrývaly s již známými *hotspoty* trimethylovaných histonů a utváření DSB. Lokalizace *hotspotů* v genomu je určená proteinem PRDM9. Z toho vyplývá, že NCO skutečně mají původ ve standardní dráze opravy meiotických DSB. Na celochromosomální úrovni jsme pozorovali rozdílné distribuce NCO a crossoverů (CO), projevující se zejména relativní preferencí CO před NCO v subtelomerických oblastech. Pozorovali jsme deficit specifické třídy tzv. *nonparentálních*, či *asymetrických* NCO a CO. Navrhujeme, že tyto NCO a CO jsou obtížně orpavitelné a hypoteticky by mohly být opraveny sesterskou chromatidou. Takováto oprava by nepřispěla k nezbytné homologní synapsi.

Alternativně, v 10 % případů, jsou meiotické DSB opraveny prostřednictvím crossoverů (CO), které jsou nezbytné pro genetickou rekombinaci i pro segregaci homologních chromosomů. S využitím křížení laboratorních myších kmenů reprezentujících poddruhy *Mus musculus domesticus* (*M.m.m.*) (C57BL/6J, zkráceně B6) a *Mus musculus domesticus* (*M.m.d.*) (PWD/PhJ, PWD) jsme ukázali, že celogenomový počet CO je geneticky podmíněný. Na chromozomu X jsme identifikovali nový genetický lokus, *Meiotic recombination Meir1*, který významně ovlivňuje frekvenci CO. Je pozoruhodné, že *Meir1* geneticky mapuje do lokusu *Hstx2*, již dříve identifikovaného faktoru hybridní sterility. Naproti tomu speciační gen *Prdm9* neovlivňoval frekvenci CO, a to ani při jeho odebrání, ani při přidání kopií.

Při studiích genetické architektury F1 hybridní sterility jsme již dříve identifikovali dvě její nutné komponenty, gen *Prdm9* a lokus *Hstx2*, obě operující *in trans*. Ukazovalo se však, že pravděpodobně existuje další neznámá komponenta operující *in cis*. Zjistili jsme, že náhodně rozmístěné homozygotní úseky o délce alespoň 27 Mb postačují ke správné homologní synapsi. Navíc jsme ukázali, že když jsou homozygotní úseky přítomny na nejkratších

a nejcitlivějších autosomech na jinak heterozygotním pozadí F1 kříženců, stačí to k odvrácení hybridní sterility.

Dále jsme zúžili lokus hybridní sterility *Hstx2* do oblasti o délce 2.7 Mb, Chr X: 66.51-69.21 Mb. Je důležité, že i takto zúžený lokus *Hstx2* stále překrývá lokus *Meir1*, který ovlivňuje míru CO. Dále jsme vyloučili gen *Fmr1nb* coby kandidáta na roli druhého genu hybridní sterility. Místo toho se nyní jeví jako slibný kandidát klastr miRNA *miR-465*.

Souhrně můžeme konstatovat, že jsme nalezli události meiotické rekombinace vedoucí k NCO v rozsahu celých chromosomů a navrhli jsme, jaké má oprava DSB možné důsledky pro párování homologů. Dále jsme ustanovili, že studovaná klasická hybridní sterilita F1 kříženců má tři komponenty: 1) gen *Prdm9* operující *in trans*, 2) mezipoddruhová homeologie operující *in cis* (pravděpodobně realizovaná prostřednictvím vazebných míst PRDM9) a 3) faktor *Hstx2* na chromozomu X operující *in trans*. Jinou vědeckou skupinou byla též navržena mechanistická, molekulární podstata hybridní sterility kříženců F1, tzv. *asymetrie vazby PRDM9*.

1. Introduction

1.1 Meiosis

Sexual reproduction gives species an advantage to evolve efficiently (in comparison to asexual reproduction), reacting to changing environmental conditions by the process of natural selection. This view is based on the evolutionary theory suggested by Charles Darwin in 1859 (DARWIN 1859). On the other hand, the sexual reproduction comes at costs, in particular the necessity for finding a sexual partner for reproduction requiring substantial energy and time.

Meiosis is a process of cell divisions characterizing all sexually reproducing organisms. The ultimate goal of meiosis is to form new combinations of alleles (more generally intervals of chromosomes), derived from their parents with the intent to provide a population with genetic variability. To enable this, meiosis has evolved to the form of two phases (Alberts et al. 2014).

Preceding meiosis itself, in S phase, both maternal and paternal chromosomes are replicated to give rise to two sister chromatids each, the similar way as in mitotic S phase. In the first meiotic phase (meiosis I), the replicated chromosomes are paired with their homologous maternal/paternal replicated chromosomes (*homologs*), forming chromosomal tetrads – bivalents. While being paired, maternal and paternal chromosomal pairs are recombined by crossovers, leading to new chromosomal combinations. Being recombined, originally completely maternal and paternal chromatid pairs (2n and 2n) are split into two different nuclei. Finally, in meiosis II, similarly as in mitosis the chromatid pairs are segregated once again into separate nuclei (without further recombination), giving both the nuclei the full haploid set of chromosomes (full set of autosomes and pair of sex chromosomes; X and Y in males, two Xs in females). Note that, in males the cells remain still partially connected forming a syncytium/coenocyte in males. Later, the male meiocytes, *spermatocytes*, undergo differentiation into sperms. On the other hand in females, in both meiosis I and meiosis II cytoplasm divides asymmetrically to produce two cells of greatly different sizes, the precursor of the egg and the small cells called *polar bodíes*, which eventually degenerate. The female meiocyte, *oocyte*, maturation becomes arrested in metaphase of meiosis II. At ovulation, the arrested *oocyte* is released from ovary and can be fertilized and the meiosis II is completed (ALBERTS *et al.* 2014).

While mitosis takes about one hour in mice, meiosis takes several days in males (but much prolonged in females due to the arrest). In males, meiosis II is a relatively straightforward process, which resembles mitosis, also time-wise. The majority of (male) meiotic time is occupied by meiosis I, which is more complex and in particular requires the non-trivial (and time-demanding) homologous chromosomes pairing in leptotene – zygotone stage of prophase I, which alone occupies ~90% of (male) meiotic time (ALBERTS *et al.* 2014). The deviations of standard meiotic homologous recombination and subsequent improper chromosomal pairing, synapsis and prophase I progression are the phenotypes we are trying to understand in the present thesis.

1.2 Meiotic prophase I

Unlike in mitotic prophase, in meiotic prophase I, the homologs are required to recognize each other and pair. This is a non-trivial, time-demanding process (in mice males meiotic prophase I takes 11 days), where the reason behind this is not yet fully understood (ALBERTS *et al.* 2014).

The meiotic prophase I is further divided into five substages: leptotene, zygotene, pachytene, diplotene, and diakinesis (ALBERTS *et al.* 2014).

In leptotene, the replicated sister chromatids are held together by cohesins (including REC8, STAG3) and chromatin starts to condense around the chromosomal axis. The main goal of prophase I is homologous recombination. To that end, there are several processes happening largely in parallel; a remarkable feature of meiosis is the formation of deliberate DNA double strand breaks (DSBs). The DSBs need to be subsequently repaired, however, in contrast to mitosis, the DSBs are to be repaired not by sister chromatid, but by the homologous chromatid (LAO AND HUNTER 2010). This leads to the recombination of DNA molecules. The sites of DSBs sites are genetically determined, most importantly by *Prdm9* gene. PRDM9 is a histone methyl-transferase having also a zinc-finger (ZnF) domain allowing a specific DNA binding. Estimated 4.700 genomic sites per meiocyte are localized by an affinity of PRDM9 ZnF to a specific DNA motif (BAKER *et al.* 2014) and the PRDM9 PR-SET domain tri-methylates surrounding histones at H3K4 (HAYASHI *et al.* 2005; SMAGULOVA *et al.* 2011; BAKER *et al.* 2014) and H3K36 (WU *et al.* 2013; ERAM *et al.* 2014; POWERS *et al.* 2016) lysine residues. The chromatids are in the form of loops protruding from chromosomal axis in leptotene and the sites of prospective DSBs are understood to be at the loops. Later on, still in leptotene, the PRDM9-marked sites are tethered to chromosomal axis and a subset of 200-300 of them are

cut by dimerised topoisomerase-like transesterases SPO11 and TOPVIBL (KEENEY *et al.* 1997; BAUDAT *et al.* 2013; ROBERT *et al.* 2016; VRIELYNCK *et al.* 2016). Each SPO11 monomer cuts the DNA at one strand, forming covalent bounds with ssDNA at 5' end. Afterwards, MRE catalyses second endonucleaotide cleavage on each strand tens of base pairs downstream leading to release of two SPO11-oligonucleotide complexes (NEALE *et al.* 2005). Then, the ssDNA is resected further leaving ssDNA overhangs (hundreds to a thousand of bp), which are coated and protected by RAD51 and meiosis-specific DMC1 proteins (NEALE AND KEENEY 2006).

The induction and repair of programmed DSBs is necessary for formation of *synaptonemal complex (SC)* along the whole lengths of homologous chromosomes (KAUPPI *et al.* 2013) and for homologous recombination (BAUDAT *et al.* 2013; BOLCUN-FILAS AND HANDEL 2018). The synaptonemal complex is a large and complex proteinaceous structure

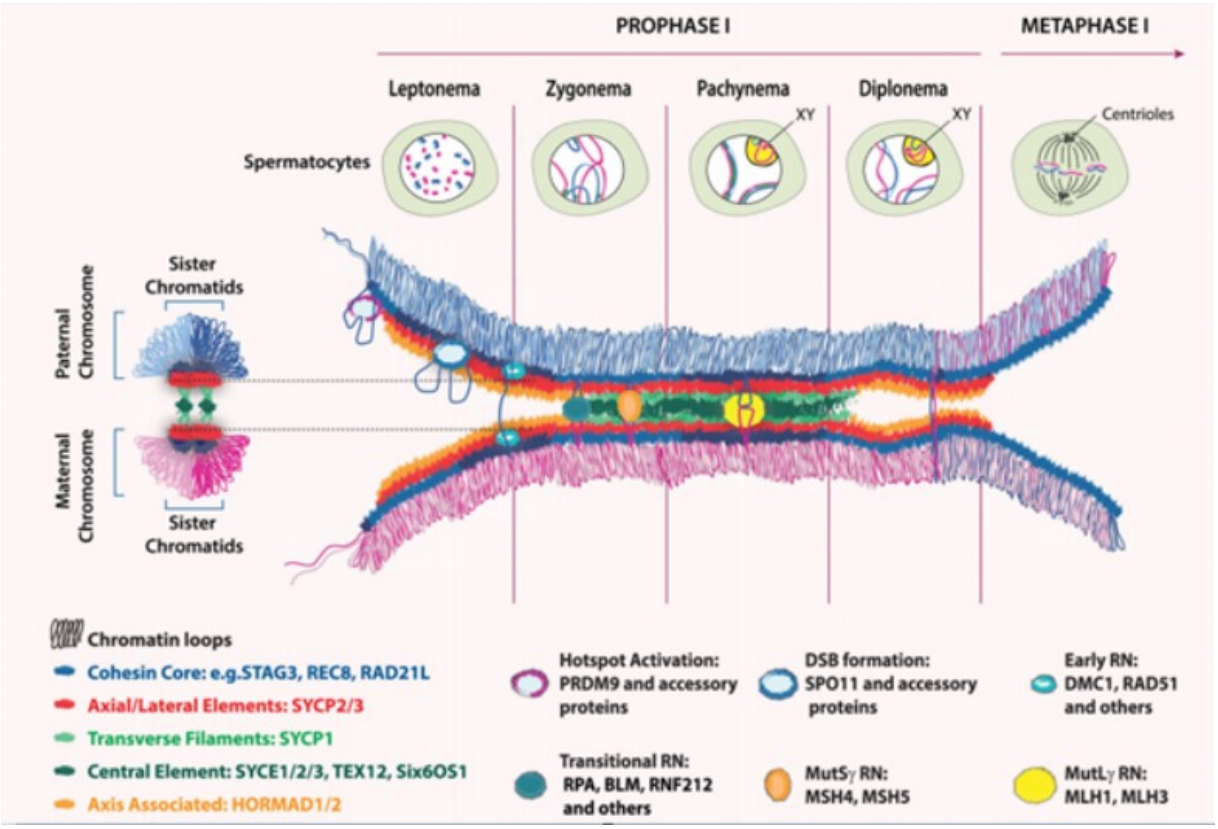

Figure 1. Homolog synapsis and desynapsis during stages of prophase I in males. Schematic spermatocyte is visualized with two pairs of homologous automosomes and the sex chromosome pair in upper panel. Cross section of the fully assembled synaptonemal complex is visualized in left panel. Longitudinal section of chromosomal axis and synaptonemal complex assembly including the recombination machinery is visualized in main panel. Adopted from (BOLCUN-FILAS AND HANDEL 2018)**.**

formed between two homologous chromosomes gradually assembled and disassembled during prophase I. It consists of three kinds of structures: a pair of parallel axial elements (AEs; later named lateral elements, LEs) and central element (CE), which are connected by transverse filaments (TFs). The axial elements are formed in leptotene together with cohesin complexes, the other SC parts are formed later in prophase I (BOLCUN-FILAS AND HANDEL 2018).

While, DNA homology search by ssDNA at DSB sites is crucial for forming of SC, initial pairing of homologs was shown to be achieved before meiotic leptotene (BOATENG *et al.* 2013). Also the formation of *bouquet* is observed at the nuclear envelope, where chromosomal telomeres are aggregated.

By the start of zygotene, the homologous (maternal/paternal) chromosomes have found their respective partner at the sites of programmed DSBs. So-called *presynaptic aligment* of homologs is followed by formation of transverse filaments (including SYCP1) and central element (including SYCE1/2/3, TEX12) between axial cores of homologs. I.e. the synaptonemal complex starts its assembly and the synapsis extends along the axial elements (consisting also of SYCP2 and SYCP3 proteins) which are then called lateral elements. Simultaneously, with the deposition of central elements, HORMAD1/2 proteins important for DSB creation and repair disappear from the chromosomal axes. The homologous pairing at the sites of DSB sites is achieved by homology search of ssDNA provided by DMC1 and RAD51 proteins. The sites of recombination, called *recombination nodules (RNs)*, gradually attract proteins necessary for a repair of DSBs at several stages.

When ssDNA finds its complementary DNA sequence in homolog, it invades into the homologous dsDNA and starts to form a D-loop. Subsequently several pathways of repair were discovered. Most frequent are synthesis-dependent strand annealing (SDSA) pathway (ALLERS AND LICHTEN 2001) leading to *noncrossover* (NCO) resolution, where the ssDNA is displaced, and double-strand break repair (DSBR) pathway (SZOSTAK *et al.* 1983), prevalently leading to *crossover* resolution (CO). In both pathways, a heteroduplex DNA (hDNA) is formed containing mispairing bases needed to be repaired. The reciprocal exchanges of homologous chromosomes, COs, are responsible for the genetical information reshuffling, necessary for a proper segregation of chromosomes later in anaphase I as well as earlier for homologous

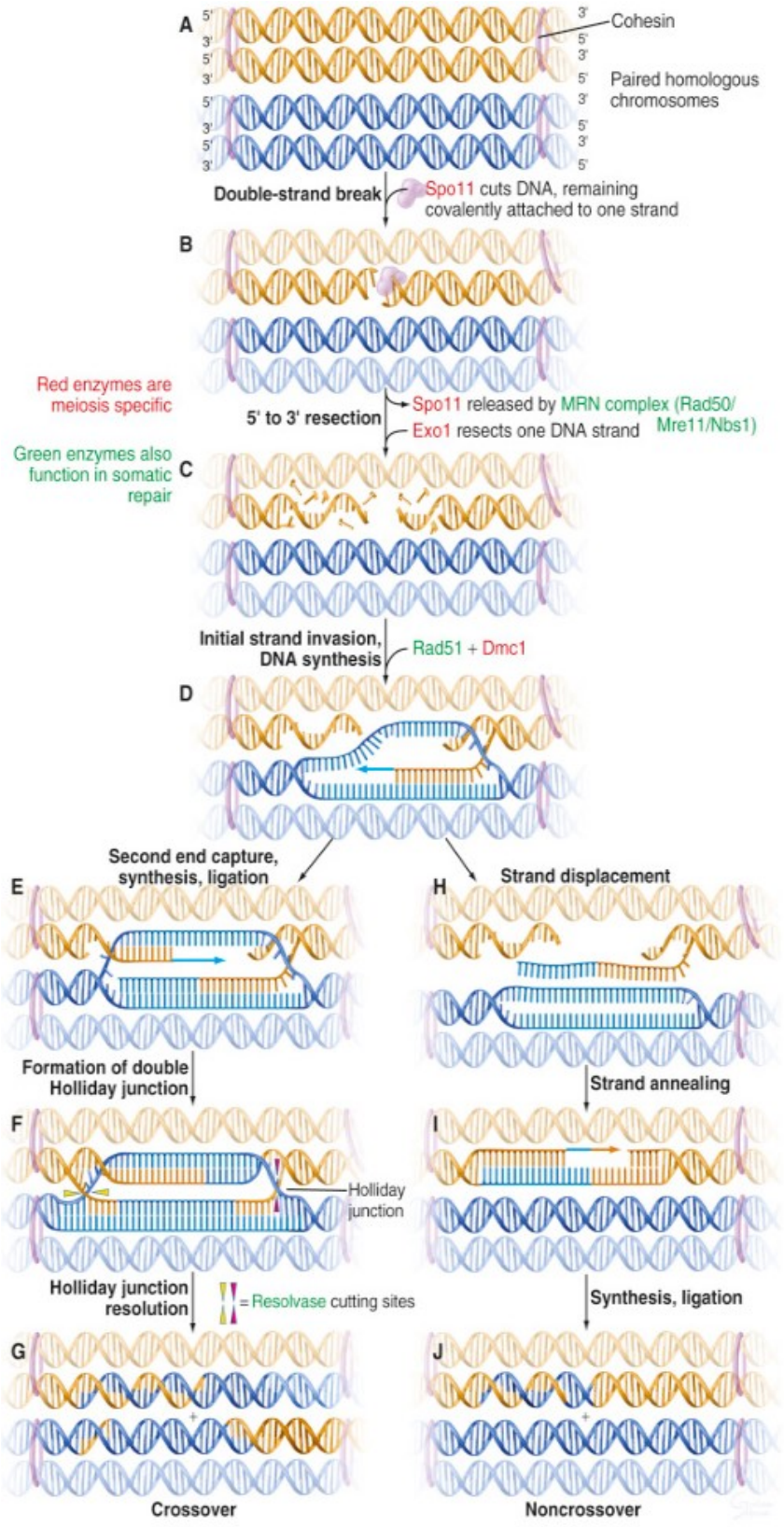

Figure 2. Events of meiotic recombination. Recombination occurs preferably between homologous chromosomes rather than between sister chromatids. In panel **D** one strand invades homologous chromatid and the pathway splits in two; **E-G**, leading to crossover and **H-J**, leading to noncrossover. Adopted from (POLLARD *et al.* 2016).

pairing and synapsis. However they represent only 10% of DSB repair (COLE *et al.* 2012). The remaining 90% are repaired by NCOs, events of non-reciprocal recombination, which are supposed to be important for homologous pairing and synapsis. In contrast to mitosis, the DSBs are repaired from the homologous chromosome with the responsible mechanisms being largely unknown. Also our understanding of mechanisms deciding whether CO or NCO resolution occurs is very limited.

The start of pachytene is standardly defined by completion of homologous synapsis along whole chromosomes. In recombination nodules the process of DSB repair by crossovers or noncrossovers continues.

The crossovers formation is regulated on several levels. At least one crossover is required in each chromosomal pair (Cole et al. 2012). On the other hand, the total number of crossovers per chromosomal pair is limited (typically just one crossover occurs, while as many as three have been rarely observed in mice) and co-occurrence of two crossovers in close proximity is highly unlikely, a phenomenon called crossover interference (Petkov et al. 2007). Also crossovers are more likely to occur at subtelomeric ends (Liu et al. 2014). Finally, on a whole nucleus scale the total number of COs is genetically determined and CO homeostasis is exhibited, leading to comparable counts of COs regardless of number of precursor DSBs (Cole et al. 2012). Note also, that the two pathways leading to CO repair were described. The 90% prevalent pathway requires proteins MLH1 (Baker et al. 1996; Edelmann et al. 1996) and MLH3 (Lipkin et al. 2002) and leads to CO interference (class I COs), while the alternative pathway is MUS81-dependent and does not cause CO interference (class II COs) (Holloway et al. 2008).

In contrast to COs, our understanding of NCO resolution is still limited and thus motivated my study of the NCOs as the main topic of my Ph.D. studies.

In diplotene, the desynapsis, i.e. disassembly of synaptonemal complex begins, namely transverse filaments and central elements are removed. At this stage, the results of crossovers, so-called chiasmata, inter-homolog connections can be observed, being the only sites where the homologs are still held together. Importantly, this enables to build a tension crucial for a meiotic plate formation, while the homologs are being gradually attached to opposite meiotic spindle poles in metaphase I.

In diakinesis (which is sometimes considered as a substage of prophase I), the meiotic spindle begins to form, similarly as in mitotic prometaphase, and nuclear membrane starts to break down.

1.3 Speciation and hybrid sterility

1.3.1 Speciation

The diversity of all living forms on our planet is arguably one of the most fascinating things we, people, can think about. The existence of diversity implies two fundamental questions:

1. Why are there different species on our planet?
   In other words, why there are species, rather than organisms of continuum of varying phenotypes, many of them freely mating together?
2. How did this diversity of species arise?

People were likely always asking these kind of questions. However, the first systematic studies on the topic are known from Charles Darwin, documented in his classics *On the Origin of Species* from 1859 (DARWIN 1859). While the first question could be even considered as philosophical, it can be hardly fully addressed and hypotheses can be hardly rigorously tested[1], the second question is a matter of research in the laboratory of Prof. Jiří Forejt, where I have been completing my Ph.D. studies.

To unambiguously define the research questions, it is necessary to define several terms, first. The above-stated term *species* is usually[2] understood as a group of organisms which can reproduce (i.e. produce fertile offspring) with one another, and which cannot reproduce with organisms outside this group (i.e. other species). This, "biological", definition was coined by evolutionary biologist Ernst Mayr (MAYR 1963) and is used also in the thesis. *Speciation* is the process leading to formation of new species from the parental species. Species are mutually separated by barriers called *mechanisms of reproductive isolations*. The barriers can be built either before mating and producing a zygote *(prezygotic barrier)*, or after forming a hybrid zygote, but not allowing it to develop into healthy and fertile adults *(postzygotic barrier)*. Finally, *subspecies* is a taxonomic rank below species, a geographic population diagnosable by

[1] Several explanations and theories were suggested. The well-accepted theories state having species is saving a lot of energy in nature. The organism within respective species are typically surviving and spreading their properties to next generations through their progeny. On the other hand, without species existence, a lot of energy would be lost as arguably many organisms would be less viable and would hardly reproduce. Thinking generally, the presence of species is rather an evolutionary necessity, considering the omnipresent stochasticity, consequent presence of errors, and evolution favorizing fitter individuals, together with geographic separations.
[2] Other definitions of *species* exist, and none of them is perfect, as is typical when describing a living matter. This problem is sometimes called "species problem"; Hey, J., 2001 The mind of the species problem. Trends in Ecology & Evolution 16**:** 326-329.

one or more phenotypic traits (which exhibit decreased fitness when intercrossed in nature), but see the species vs. subspecies definitions debate (BARTON AND HEWITT 1989; SILVER 1995).

At first sight, hybrid sterility is not consistent with the Darwin's evolutionary theory, which is expecting a gradual and continuous evolution of phenotypic traits. This puzzled Charles Darwin and other evolutionary biologists at the time. The critical question was, what could be the mechanism that seperate groups of organisms from one another to form species?

Later on, an elegant solution was suggested in *Bateson-Dobzhansky-Muller (BDM) model* of genetic incompatibility (BATESON 1909; MULLER AND PONTECORVO 1942; DOBZHANSKY 1951; ORR 1996). The BDM model suggests that in two subpopulations two (or more) epistatically interacting genes are evolving independently. When the subpopulations separate, in each of them, mutations in the two genetic loci become fixed (as an assumption). Within each population the genetic interacting loci are compatible because they have undergone a joint evolution. However, according to the BDM model, when organisms of the two subpopulation meet each other again, two genetic loci cannot successfully interact anymore epistatically, and thus the organism is either not able to reproduce or their progeny is sterile. I.e. the BDM model can potentially apply to construction of both prezygotic as well as postzygotic barriers (including as a special case *hybrid sterility* described in the next section).

1.3.2 Hybrid sterility

Hybrid sterility is one of the possible conditions leading to speciation. It restricts gene flow, belongs to postzygotic reproductive isolation barriers and was defined by Theodosius Dobzhansky as a situation where both parental species are fertile *inter se*, while their hybrid is sterile (DOBZHANSKY 1951).

Interestingly, if hybrid sterility affects only one sex, the sterile sex is almost always the heterogametic one. The condition is so prevalent across various species that it was denoted as *Haldane's rule* (HALDANE 1922). Note that the *heterogametic sex* is the one bearing two different sex chromosomes, i.e. a male with X and Y sex chromosomes in case of mammals and a female with Z and W sex chromosomes in case of birds, giving the most popular examples. The observed Haldane's rule led to the proposition of several theories aiming to explain it (the dominance theory (MULLER 1940; HOLLOCHER AND WU 1996; TRUE *et al.* 1996), the

theory of faster males (WU AND DAVIS 1993; PRESGRAVES AND ORR 1998), the theory of faster X chromosome (CHARLESWORTH *et al.* 1987)). However, none of them are able to explain all the examples observed in nature, while they are not mutually exclusive (MAHESHWARI AND BARBASH 2010).

While popular examples of sterile hybrids are the crosses between horse and donkey producing almost always sterile progeny (either mule, or hinny depending on mating direction), obviously other species were utilized for systematic hybrid sterility research.

Important findings were achieved with the "small fruit fly" *Drosophila* as a model. Apart from the generally appreciated properties (cheap to house and space-modest, short generation time, etc.), additionally the *Drosophila* genus encompasses many species (≥ 1500) including a classical model organism of genetics, *Drosophila melanogaster (D. melanogaster)*, and others like *D. mauritania*, *D. simulans* and *D. pseudoobscura*. Some hybrids between these species were shown to be sterile when being mutually crossed, allowing studies of interspecies barriers. First, candidate genetic loci of hybrid sterility were found in *Drosophila* crosses. Later on, by means of positional cloning, people were able to identify several hybrid sterility genes, *OdsH, Hmr, Nup96, JYAlpha, Ovd* (TING *et al.* 1998; BARBASH *et al.* 2003; PRESGRAVES *et al.* 2003; MASLY *et al.* 2006; PHADNIS AND ORR 2009). The physiological function of the genes varies. The first identified gene, for instance, Odysseus *(OdsH)* is a homeobox gene with a rapidly evolving DNA-binding homeodomain. Remarkably, an interacting BDM partner has not been identified for any of the genes yet.

1.3.3 Mouse model of hybrid sterility

In mammals, the classical model organism, "house mouse" *Mus musculus* has also been used successfully in speciation studies. Similar to *Drosophila melanogaster,* the *Mus musculus* model comes with benefits of a classical model organism (here, *relatively* low cost, *relatively* short generation time, availability of large amount of classical and *wild-derived* strains with fully sequenced genomes, commercially available antibodies for cytological investigation, etc.). In addition, there are specific reasons, why *Mus musculus* is a suitable model for hybrid sterility studies. First, narrow hybrid zones of subspecies secondary contact inhabited by mice with decreased fertility were indeed found in nature. Second, the divergence time between mouse subspecies is roughly 350 000 years (DIN *et al.* 1996; GERALDES *et al.* 2008). This is a

relatively short time, suggesting the investigated genetic architecture should be the cause of the hybrid sterility rather than its consequence.

Specifically, in the wild three house mouse subspecies have been described: *Mus musculus domesticus* (*M.m.domesticus, M.m.d.*), *Mus musculus musculus* (*M.m.musculus, M.m.m.*), *Mus musculus castaneus* (*M.m.castaneus*), each inhabiting their own part of Earth. One well-studied example of hybrid incompatibility is the European narrow hybrid zone between *M.m.domesticus* (western) and *M.m.musculus* (eastern). Evolutionary scientists were aiming to describe the responsible genetic architecture by sampling wild mice from the hybrid zone, mutually crossing them and correlating their phenotypes and genotypes (PAYSEUR *et al.* 2004; VYSKOCILOVA *et al.* 2005; MACHOLAN *et al.* 2007; TEETER *et al.* 2008; TEETER *et al.* 2010). This approach accurately describes the reality of both phenotypic and genomic variations in nature, on the other hand, the genetic resolution provided by the studies is typically low. As such it has not revealed any hybrid sterility gene, nor has it suggested molecular causes for hybrid incompatibilities. An alternative approach is to model the processes happening in nature using well-defined laboratory strains, allowing the increase of statistical power and reproducibility.

Back in 1969, Iványi first observed (*accidently* while studying polymorphisms of H2 histocompatibility systems) sterile hybrids when crossing local wild mice with laboratory strains including C57BL/10, representing *M.m.musculus* subspecies (IVANYI *et al.* 1969). The responsible locus was subsequently identified as hybrid sterility locus, mapped to Chr 17 and denoted *Hybrid sterility 1 (Hst1)* (FOREJT AND IVANYI 1974). A long-term effort led gradually to the identification of the *Hst1* locus with *Prdm9* gene encoding meiotic histone H3K4 trimethyl transferase (FOREJT 1985; FOREJT *et al.* 1991; GREGOROVA *et al.* 1996; TRACHTULEC *et al.* 2005; TRACHTULEC *et al.* 2008; MIHOLA *et al.* 2009).

One year later, the physiological function of *Prdm9* gene was discovered. *Prdm9* was shown to be responsible for the determination of meiotic programmed DSB positions by binding specific DNA motifs by PRDM9's ZnF domain and trimethylating surrounding histones H3 at K4 by PRDM9's PR/SET domain, consequently leading to chromatin opening (BAUDAT *et al.* 2010; MYERS *et al.* 2010; PARVANOV *et al.* 2010). This is known to be followed by other events leading to induction of DSBs there, CO formation in a subset of DSB sites, proper homologous chromosomes pairing, synapsis and segregation (see the Meiotic prophase I chapter of Introduction). However, many questions remained unresolved, e.g. what is/are the other

genomic partner(s) of *Prdm9* (assuming validity of BDM model) and why is DSB sites determination important for hybrid sterility, i.e. what is the explanation of the dual role of *Prdm9* gene, to name the most intriguing ones. And of course, the outstanding question is, how well the model recapitulates the speciation processes in nature in mice and in theory (and ambitiously) also in other species?

To that end, Prof. Jiří Forejt has chosen a classical F1 hybrid sterility model using two mouse strains, namely, representative of *M.m.musculus*, the wild-derived strain PWD/Ph (henceforth PWD) (GREGOROVA AND FOREJT 2000) and representative of *M.m.domesticus*, the classical laboratory strain C57BL/6J (B6). The male progeny of PWD female and B6 male ((PWD×B6)F1 cross) is sterile, while female progeny is fertile, consistent with Haldane's rule. Also, it was shown that the reciprocal cross (B6×PWD)F1 produces semifertile males and fertile females. The *asymmetry of male infertility* between the reciprocal hybrids is also a general phenomenon observed in many species (and should not be confused with *asymmetry theory (of PRDM9 binding)* described in Discussion, or with Haldane's rule). The long-term aim was defined as to unravel the (complete) genetic architecture of the (PWD×B6)F1 hybrid sterility model, the physiological function of the participating genomic elements and the molecular mechanisms responsible for the genomic incompatibility.

Importantly, throughout the investigation of F1 hybrid sterility, the model of consomic (chromosomal substitution) strains (CSs) has been useful. In each CS, a single B6 chromosome was replaced with its homolog from PWD. Altogether, the CS panel contains 18 consomic, nine subconsomic, and one conplastic strain (GREGOROVA *et al.* 2008). Specifically, the CS panel, contains subconsomics B6.PWD-ChrX1, B6.PWD-ChrX2, B6.PWD-ChrX3 strains with proximal, middle, and distal part of chromosomes, respectively. The fourth subconsomic strain B6.PWD-ChrX1s, carrying extended proximal part of Chr X from PWD, was prepared later (BHATTACHARYYA *et al.* 2013).

In the subsequent study, the cause of *asymmetry of male infertility* was resolved (DZUR-GEJDOSOVA *et al.* 2012). The PWD region on Chr X: 61.0–94.3 Mb interacting with heterozygous genome was found to be responsible for the sterility, because all other possible options (X-Y interaction, autosome-Y or mitochondrion incompatibility, or incompatibility of imprinted autosomal gene(s)) were excluded.

Also importantly, the presence of the two genetic factors was found necessary but not sufficient for the hybrid sterility as only 31 out of 70 (44%) males were sterile or semifertile. Subsequent QTL mapping using the 70 males, suggested one more genetic factor on Chr 14 contributing to sterility, provided being heterozygous. This would agree with the performed calculations; altogether 3-4 genes/loci should be responsible for the sterility. However, a significant correlation between whole-genome heterozygosity and sterility was observed (when all Chr 17, Chr X, and Chr 14 were excluded). As an alternative explanation, the correlation could be caused by presence of unknown noncoding DNA *cis*-components spread along the genome contributing somehow to the hybrid sterility.

The middle part of Chr X responsible for hybrid sterility was further narrowed down to 4.7 Mb region of Chr X (Chr X:64.88–69.51) and denoted as *Hybrid sterility 2* locus, *Hstx2* (BHATTACHARYYA *et al.* 2014) (**Publication #7**). It was also shown to contain the previously identified *Hstx1* locus (STORCHOVA *et al.* 2004). The *Hstx2* locus contained six known protein-coding genes, two of which, *Fmr1nb* and *4933436I01Rik*, being interesting candidates as both were expressed in the appropriate cell type and contained non-synonymous nucleotide substitutions. In addition to this, two microRNA clusters are localized in *Hstx2*, whose functions are generally hard to predict.

The sterility tests and sperm count in *ductus epididymis* are the defining criteria of male fertility, and testes weight utilized as a useful sterility approximation. However, for understanding the molecular mechanisms of investigated hybrid sterility in the (PWD×B6)F1 model, it was necessary to specifically define the first meiotic process showing significant disturbances.

Indeed, in laboratory of Prof. Jiří Forejt, it was shown that it was the (im)proper formation of synaptonemal complex (herefter (a)synapsis) in zygotene of meiotic prophase I, consistently with deficiency of mid-pachytene cells at 15.5 days post partum (dpp), that was first affected, when homeologous chromosomes (homologs from related (sub)species) need to be paired and leads to meiotic arrest (Bhattacharyya et al. 2013).

The asynapsis was found in 90% of meiocytes in (PWD×B6)F1 males, while only in 32% in the reciprocal (B6×PWD)F1 males, consistent with observed differences in sperm counts. Interestingly, in sterile (PWD×B6)F1 males, the asynapsis rate differed between autosomes with the Chr 19 (46.7%) being significantly more and Chr 2 (4.7%) and Chr 16 (6.0%) less asynapsed than randomly expected. In line with this, in crossing using consomic strains, the

synapsis was suggested to behave autonomously; 1) there were no asynapsis on Chr 19 (0/100 cells) in PWD×B6.PWD-Chr 19, and orthogonally, and 2) in B6.PWD-Chr X.1s × B6.PWD-Chr 17, the only asynapsis found (17%) was situated on Chr 17 (25/25).

Altogether, these observations specified the cellular phenotype responsible for F1 hybrid sterility and suggested a presence of cis-component controlling asynapsis rate and hybrid sterility.

2. Aims

The three main aims of the Thesis could be described as following:

1) To identify and characterize products of meiotic recombination, in particular the nonreciprocal events (noncrossovers, NCOs) occurring in meiotic prophase I. With the possibility of the unique detection of NCOs in consomic mice, many questions are arising, e.g.: Do all the NCOs result from the standard programmed meiotic DSBs pathway, e.g. is their position determined by PRDM9? What is chromosome-wide distribution of NCOs? Do NCO and CO chromosome-wide distributions differ? How long are the NCO gene conversion tracts? Could we get any insight about the importance of NCOs for homologous pairing, synapsis, and consequently hybrid sterility?

   We aimed to answer these questions with the use of a unique mouse model of chromosomal substitution strains (CSSs, also named consomics) B6.PWD-Chr#, by sequencing the whole genomes (WGS) of strains, identifying tiny NCO events from the WGS data and comparing them with the current data on meiotic homologous recombination.
2) In the mouse (PWD×B6)F1 hybrid sterility model we previously identified two *trans*-acting factors, *Prdm9* gene and *Hstx2* locus, necessary for male hybrid sterility. However, these two factors were not sufficient for the sterility. In theory, other two genes/loci could explain the remaining variability in sterility phenotype. Alternatively, also the heterozygosity of homologous chromosomes correlated well with sterility, suggesting a role of heterozygous *cis*-acting factor(s). Following this hypothesis we wanted to investigate sterility/fertility phenotypes (in particular formation of synaptonemal complex) using males having varying lengths of inter-subspecific heterozygosity on autosomes.
3) The long-term research aim of the Laboratory of Mouse Molecular Genetics is to understand the process of speciation using a mouse model. Apart from the two main projects in 1) and 2), I collaborated on other projects, particularly investigating the role of meiotic recombination, the role of *Prdm9* gene, and *Hstx2* locus.

3. List of Methods

Next-generation sequencing (NGS) data processing

Quality control of NGS data, Paired-end reads alignment to the reference genomes (bwa), variant detection (GATK), working on computional cluster

R, Python, Bash (Unix shell) scripting

Filtering the NCO candidate sites

Investigating colocalization of genomic features

Data visualization

Statistical analysis

QTL analysis

Generalized Linear Mixed Models, Generalized Linear Models (Logistic regression, Poisson regression), Linear Mixed Models, Robust Linear Mixed Models, Linear Models (Linear regression), ANOVA

Change-point detection regression model (estimating the amount of homology required for proper synapsis)

Bootstrap (confidence interval construction)

Permutation tests (comparing chromosome-wide distributions and evaluating of significance of genome-wide colocalizations of genomic features)

Power analysis

Probabilistic modelling

Expected amount of repaired meiotic DSBs

Data curation

Consistency of the achieved phenotype data

4. Research Papers and Manuscripts

4.1 List of the publications included in the thesis

During my Ph.D. studies I contributed to the following studies either as a (joint) first author, (joint) corresponding author, or co-author. Please note, that the studies #6, #7, #8 were published, while I was studying under formally different Ph.D. programme.

#1 Chromosome-wide distribution and characterization of meiotic noncrossovers in mice from intersubspecific crosses;

Gergelits V*, Parvanov E, Simecek P, Forejt J*; **submitted manuscript; *BioRxiv***, 2019, 792226.

* These authors are joint corresponding authors.

#2 Modulation of *Prdm9*-controlled meiotic chromosome asynapsis overrides hybrid sterility in mice;

Gregorova S[1], **Gergelits V[1]**, Chvatalova I, Bhattacharyya T, Valiskova B, Fotopulosova V, Jansa P, Wiatrowska D, Forejt J*; ***Elife***, 2018, 7: e34282.

[1] These authors contributed equally.

#3 Hybrid Sterility Locus on Chromosome X Controls Meiotic Recombination Rate in Mouse; Balcova M, Faltusova B, **Gergelits V**, Bhattacharyya T, Mihola O, Trachtulec Z, Knopf C, Fotopulosova V, Chatalova I, Gregorova S, Forejt J*; ***PLoS genetics***, 2016, 12.4: e1005906.

#4 Genomic structure of *Hstx2* modifier of *Prdm9*-dependent hybrid male sterility in mice;

Lustyk D, Kinsky K, Ullrich K K, Yancoskie M, Kasikova L, **Gergelits V**, Sedlacek R, Chan Y F, Odenthal-Hesse L, Forejt J*, Jansa P*; ***Genetics***, 2019

#5 Removal of the histone-methyltransferase PRDM9 guiding meiotic DNA breaks is mostly compatible with rat fertility;

Mihola O, Landa V, Pratto F, Brick K, Smagulova F, Flachs P, Kobets T, **Gergelits V**, Tresnak T, Camerini-Otero R D, Pravenec M, Petukhova G V*, Trachtulec Z*; **submitted manuscript**.

#6 On Gene Conversion Properties;

Gergelits V*; International Journal on Biomedicine and Healthcare, 2015

#7 X Chromosome Control of Meiotic Chromosome Synapsis in Mouse Inter-Subspecific Hybrids; Bhattacharyya T, Reifova R, Gregorova S, Simecek P, **Gergelits V**, Mistrik M, Martincova I, Pialek J, Forejt J*; ***PLoS genetics***, 2014, 10.2: e1004088.

#8 Maternal–foetal genomic conflict and speciation: no evidence for hybrid placental dysplasia in crosses between two house mouse subspecies; Kropackova L, Pialek J, **Gergelits V**, Forejt J, Reifova R*; ***Journal of evolutionary biology***, 2015, 28.3: 688-698.

4.2 Contribution

Publication #1:

Václav Gergelits as the first author and shared corresponding author co-designed the study, programmed some pipelines for the meiotic noncrossover detection and filtering from next-generation sequencing data, performed the statistical analysis, interpreted the results, wrote the manuscript and prepared the figures under the supervision of Prof. Jiří Forejt, participated in the review process, and co-founded the experiments from the student grant.

Publication #2:

Václav Gergelits as the shared first author performed data curation, statistical analysis (GLMM models, change point detection, power analysis, QTL analysis), probabilistic simulations (prediction of minimal amount of repaired DSBs, multivariate correlated distributions of asynapsis in autosomes), interpreted the results, co-wrote the manuscript, and prepared the figures under supervision of Prof. Jiří Forejt, and participated in the review process.

Publication #3:

Václav Gergelits as the co-author performed statistical analysis, interpreted the results, discussed the manuscript preparation, prepared the figures and wrote some parts of manuscripts and participated in the review process.

Publication #4:

Václav Gergelits as the co-author performed statistical analysis and discussed the manuscript preparation.

Publication #5:

Václav Gergelits as the co-author performed statistical analysis, shared the R scripting codes, and discussed the manuscript preparation.

Publication #6:

Václav Gergelits as the only author wrote the review publication, while the supervisor Prof. Jiří Forejt provided some corrections.

Publication #7:

Václav Gergelits as the co-author performed statistical analysis, wrote some parts of manuscript and prepared some figures, and discussed the manuscript preparation.

Publication #8:

Václav Gergelits as the co-author performed statistical analysis, wrote some parts of manuscript and prepared some figures, and discussed the manuscript preparation.

5. Discussion

This thesis summarizes my contribution to studies of meiotic homologous recombination, recombination related synapsis of homologous chromosomes, and the role of homologous recombination in hybrid sterility, using mouse (*Mus musculus*) as a model organism.

5.1 Genetic control of meiotic recombination rate in mouse

Induction of meiotic DSBs by SPO11 and subsequent repair by homologous recombination is crucial for homologous synapsis, the consequent homologs segregation, and proper progression of gametogenesis. The *Prdm9* gene is responsible for the positioning of meiotic DSBs as well as plays crucial role in hybrid sterility, in specific combination with the *Hstx2* locus leading to meiotic arrest in pachytene. The dual role of *Prdm9* motivated us to investigate possible connections between homologous recombination and hybrid sterility architecture in the intersubspecific crosses between B6 (prevalently *M.m.d.*) and PWD (derived from *M.m.m.*) strains.

With this aim, we systematically investigated the genetic determination of the genome-wide crossover (CO) recombination rate (BALCOVA *et al.* 2016) (**Publication #3**), by means of immunofluorescence microscopy we measured the cytological presence of MLH1 protein (marker of CO site) during the pachytene stage. We used the power of (sub)consomic strains B6.PWD-Chr# to assess a contribution of putative genetic loci in respective PWD chromosomes to the total amount of CO sites per cell (CO rate). In addition, we investigated the role of *Prdm9* gene on CO rate. In males, the two parental strains differed in their global CO rate. For B6, the CO rate mean was 24.9, while for the mean for PWD was 29.6. Importantly, we found the strongest modifier of CO rate in the region overlapping 4.7 Mb *Hstx2* locus, which we denoted as *Meiotic recombination 1, Meir1*. This finding was achieved using the power of subconsomic strains B6.PWD-ChrX1, B6.PWD-ChrX1s, and B6.PWD-ChrX2. Interestingly, other studies have reported the strongest effect in Chr X. In F2 crosses of PWD (representative of *M.m.d.* subspecies) and B6 (*M.m.c.*), the strongest QTL for CO rate was mapped to the 68.5 Mb – 87.3 Mb (GRCm38) region, which is overlaps the *Meir2* candidate locus (DUMONT AND PAYSEUR 2011). In F2 cross of CAST (*M.m.c.*) and B6 (*M.m.d.*) (i.e. the third pair of mouse subspecies trio) the strongest QTL on Chr X was found, but more distally from *Meir1* (MURDOCH *et al.* 2010).

Interestingly, the *Prdm9* gene was not found to interfere with the CO rate in mice with B6 or PWD genomic backgrounds. In mice with B6 genetic background, the CO rate was not affected either by removing one *Prdm9* allele (B6.*Prdm9*$^{tm1Ymat/wt}$, 24.3), or by adding two other allelic copies of *Prdm9*^{C3H} in two BAC transgenes (24.2). Similar findings were achieved in mice with a PWD genetic background. The heterozygous null Prdm9 allele PWD.*Prdm9*$^{wt/-}$ (29.6) had a statistically indistinguishable CO rate from wildtype PWD males (29.6). This contrasts with the effect of the *PRDM9* gene on human and cattle recombination rate (KONG *et al.* 2014; MA *et al.* 2015).

The colocalization of *Meir1* and *Hstx2* loci is remarkable, considering the similar dual role of *Prdm9* gene*,* in homologous recombination and hybrid sterility. On the other hand, while the parsimonious idea of one pleiotropic gene in *Meir1/Hstx2* locus controlling both phenomenons would be attractive, the 4.7 Mb long locus is still quite large, containing six known protein-coding genes and two miRNA clusters. Remarkably, we found that the *Meir1/Hstx2* locus behaves as a recombination coldspot, complicating further localization of the responsible gene(s) / miRNAs.

5.2 The theory of asymmetry of PRDM9 binding

The meiotic DSBs occur in the prespecified DNA sites, determined by the zinc finger domain of PRDM9 protein (BAUDAT *et al.* 2010; MYERS *et al.* 2010; PARVANOV *et al.* 2010). The repair of the meiotic double-strand breaks (DSBs) is provided canonically from the homologous chromosome. The SPO11 induces the double strand break and the subsequent ssDNA overhang searches for similarity in the homologous chromosome, creating a heteroduplex. The heteroduplex must be resolved, often leading to a change of the original genetic information (due to gene conversions), including the PRDM9 binding motif. During evolution, the binding affinity of PRDM9 to the DNA binding motif gradually decreases, and DSB activity at that site is either decreased or abolished completely. Altogether, this leads to the so-called "*erosion of Prdm9 binding sites"*, occurring in a *Prdm9* allele-specific way in each genome. This ongoing erosion of these sites creates a pressure on the change of the zinc finger domain of PRDM9, leading to a high variability of the *Prdm9* gene in wild populations.

This brought a question, if the hotspots are being constantly eroded during the evolution, how could there be any hotspots in the genome (*hotspot paradox*) (BOULTON *et al.* 1997)? There are several observations to explain the paradox:

1) The sequence coding for zinc the finger domain of PRDM9 is a segment of the genome undergoing rapid evolution in humans (MYERS *et al.* 2010). The reason is that the ZnF domain is encoded by a DNA microsatellite region, which is generally prone to mutations, insertions and deletions. The programmed DSBs landscape is highly sensitive to even a subtle change in the amino acids (SMAGULOVA *et al.* 2016), possibly bringing a completely new landscape of "uneroded" or "fresh" binding sites for the new PRDM9 allele.
2) It is likely that only the DNA bases in the PRDM9 binding motif are important for the binding, and not the surrounding DNA bases. In addition, the DNA heteroduplex migrates away from the PRDM9 binding motif and surprisingly only 20% of the programmed DSBs led to erosion of the PRDM9 binding site (COLE *et al.* 2014).
3) Even within the known PRDM9 binding sites, some DNA bases are less conserved and thus arguably less important for the PRDM9-DNA affinity, meaning their possible erosion would not necessarily affect the PRDM9 binding.
4) Not all the heteroduplex resolutions lead to actual substitution of the base on the initiating chromatid by the base from the homologous chromatid. Instead, it was shown this happens in a half of cases and it is governed by the preference to retain the G/C bases in the genome. This phenomenon leads to gene conversion GC bias (LI *et al.* 2019b).

The question now was how to understand the dual role of *Prdm9* gene. What is the relation between the positioning of DSBs, genome-wide erosion of PRDM9 binding sites and hybrid sterility?

In 2016, an Oxford group led by Peter Donnelly and Simon Myers proposed the *asymmetry of PRDM9 binding*, or *asymmetry theory*, as the answer (DAVIES *et al.* 2016). By replacing the B6 zinc finger with humanized zinc finger, they were able rescue mouse fertility and thus proved that, the zinc finger domain of PRDM9 is responsible for hybrid sterility.

Importantly, they used a ChIP-seq assay (chromatin immunoprecipitation (ChIP) followed by "next-generation" sequencing (NGS)) to measure the activity of DMC1 protein in

meiotic prophase I in respective homologs, and inferred the PRDM9 allele controlling the DSB activation based on the DNA motif. Using this approach, they observed a highly asymmetric distribution of DMC1 signal, showing that $PRDM9^{B6}$ preferably binds to PWD homolog and *vice versa* $PRDM9^{PWD}$ preferably binds to B6 homolog. They proposed, the observed "*asymmetry of PRDM9 binding*" to be main molecular basis of hybrid sterility. Other groups have reported similar observations of asymmetric binding (BAKER *et al.* 2015a; SMAGULOVA *et al.* 2016).

Note that, the asymmetry theory does not provide a full explanation for the hybrid sterility phenotype. While the (PWDxB6)F1 cross produces sterile males with an asynapsis rate as high as 90%, the reciprocal (B6xPWD)F1 males are semi-fertile with an asynapsis rate of 45%. Yet, the level of PRDM9 binding asymmetry is comparable. Apparently, the *Hstx2* locus identified by forward genetics approach plays a crucial and yet unknown role here.

In effort to explain the hybrid sterility in (PWDxB6)F1, alternative ideas were proposed. In particular a significant amount (12%) of PRDM9 binding occurs in so-called *default* DSB sites (i.e. detected in *Prdm9* null mutants, in open chromatin regions, especially promoters; these sites are evolutionary conserved in organisms lacking *Prdm9*) in (PWDxB6)F1 males, while a half of *default* sites in the reciprocal (B6xPWD)F1 males. As B6.$Prdm9^{-/-}$ are sterile (HAYASHI *et al.* 2005), the increased presence of *default* DSB (BRICK *et al.* 2012) could affect the hybrid sterility. On the other hand, on wild-derived mouse backgrounds, the null mutants PWD.$Prdm9^{-/-}$ were semifertile (MIHOLA *et al.* 2019). Similarly, in rats, *Prdm9*-deficient males (**Publication #5**) were semi-fertile. Also, a significant amount of DSB hotspots were identified in the repetitive sites (YAMADA *et al.* 2017). These could be harmful for homologous synapsis in intersubspecific hybrids because of the possibility of non-homologous interactions of non-homologous regions.

Furthermore, PRDM9 multimerizes *in vivo* (BAKER *et al.* 2015b), and was also reported to trimerize *in vitro* (SCHWARZ *et al.* 2019). Finally, one PRDM9 allele (Cast) was shown to be dominant over another (B6, or Hum), i.e. up to three times more binding sites were controlled by $PRDM9^{Cast}$, than by $PRDM9^{B6}$ in B6 x CAST heterozygotes (BAKER *et al.* 2015b) (or $PRDM9^{Hum}$ (HINCH *et al.* 2019), respectively). This observation shows further complexity, where the latter case shows the opposite of expectations of*asymmetry of PRDM9 binding*. It is not fully understood what role these features of PRDM9 could play in hybrid sterility.

5.3 *Cis*-acting component of F1 hybrid sterility architecture

In previous studies of F1 hybrid sterility architecture of the (PWD×B6)F1 model, two genetic factors (heterozygous *Prdm9*$^{PWD/B6}$ gene and *Hstx2*^{PWD} locus) were identified as necessary for the hybrid sterility. On the other hand, the combination was necessary but not sufficient, as only 44% of such males were sterile or semi-fertile (DZUR-GEJDOSOVA *et al.* 2012; BHATTACHARYYA *et al.* 2014). Apart from some undetected genes being possible missing pieces in the puzzle, overall heterozygosity of the genome was shown to negatively correlate with fertility level, measured by sperm count and testes weight. This suggested a presence of unknown heterozygous, possibly non-coding and *cis*-acting elements in autosomes are the culprit.

Thus, we systematically evaluated the effect of autosomal heterozygosity on the fertility level (GREGOROVA *et al.* 2018) (**Publication #2**). Here, we focused here solely on the situation with a fixed heterozygous *Prdm9*$^{PWD/B6}$ gene and *Hstx2*^{PWD} locus and prepared various crosses of 3rd and 4th generation between PWD females and B6.PWD-Chr# consomic males to gain mostly heterozygous males with varying lengths of homozygous PWD/PWD stretches on two to four autosomes. To that end, the power of chromosomal substitution strains B6.PWD-Chr# model was leveraged. In the males, we measured autosomal asynapsis rate by fluorescence *in situ* hybridization (FISH) and other fertility parameters.

We found that the autosomal synapsis is indeed *cis*-dependent and largely chromosomal autonomous. First, we found that the autosomal asynapsis rate in (PWD×B6)F1 hybrids was significantly higher in shorter chromosomes (Spearman's ρ=0.916). Importantly, we were able to rescue the male fertility when sufficient amounts (and randomly localized) homozygous stretches were introgressed to the shortest autosomes, proving the causal role of *cis*-component for hybrid sterility. While the maximal asynapsis rate varied per autosome, a universal presence of 27 Mb of PWD/PWD homozygous stretch was sufficient for the autosomal synapsis. In line with these observations, we suggested a probabilistic model of a minimum of two repaired DSBs necessary for proper autosomal synapsis.

We propose that the observed *cis*-component of genetic control of hybrid sterility are in fact the PRDM9 binding sites throughout the whole genome. As explained in the previous chapter, the DNA binding sites were detected to be bound largely asymmetrically by PRDM9^{B6} and PRDM9^{PWD} alleles in (PWD×B6)F1 hybrids (DAVIES *et al.* 2016) and were suggested to severely affect fertility. Indeed, we observed a gradual, continuous change of autosomal asynapsis rate depending on the length of PWD/PWD stretches. Thus, the expected amount

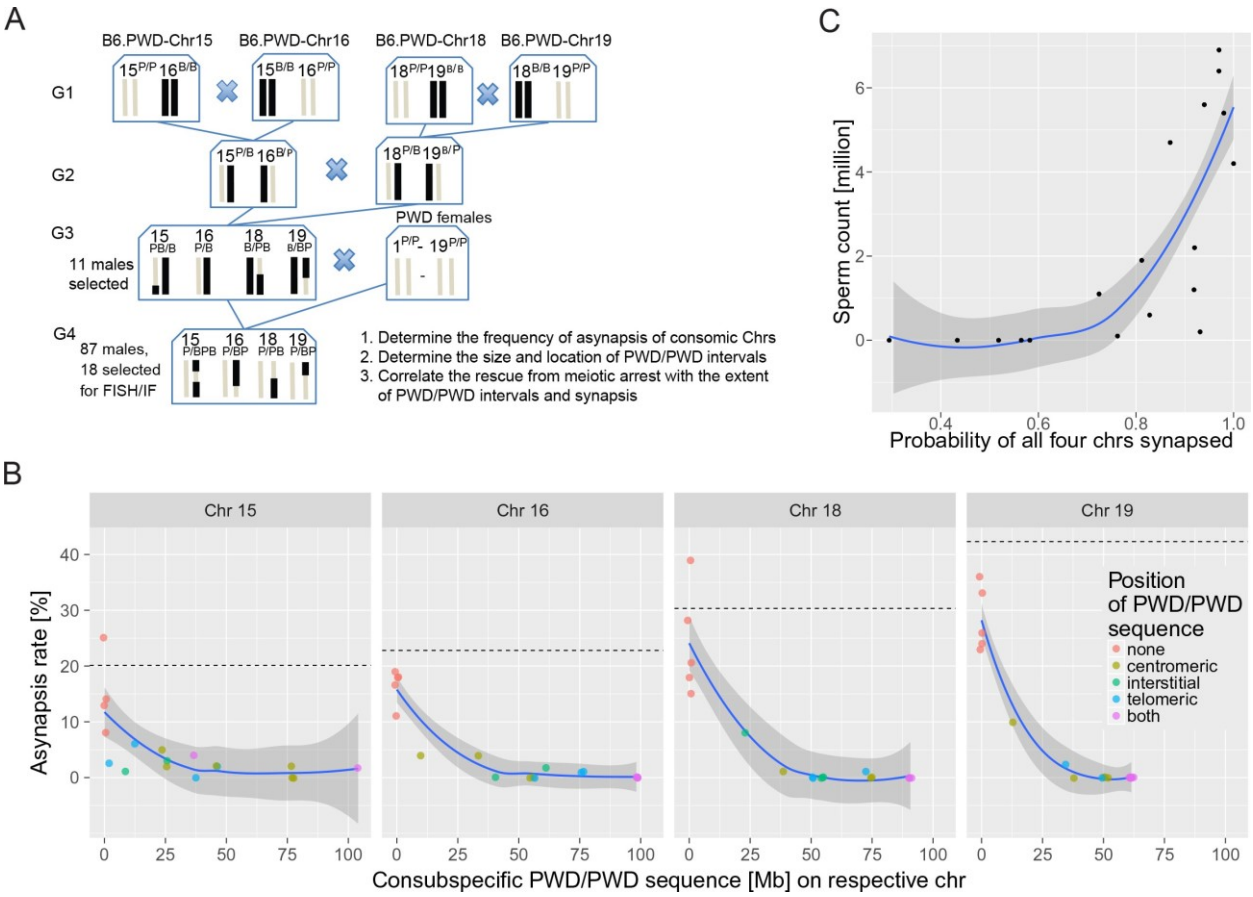

Figure 3. The effect of consubspecific stretches on pachytene synapsis and meiotic progression. A Scheme of preperation of F1 hybrids with consubspecific stretches on the four shortest chromosomes; Chr 15, Chr 16, Chr 18, Chr 19. **B** Autosomal asynapsis is reversed when 27 Mb of consubspecific sequence is present. **C** Hybrid sterility is reversed when probability of synapsis of four shortest chromosomes exceeds 0.7. Adopted from (GREGOROVA *et al.* 2018).

of symmetric PRDM9 binding sites is fully consistent with the *asymmetry theory* (DAVIES *et al.* 2016), giving the first experimental evidence of its validity.

At the time when the DMI model was introduced, the understanding of molecular foundations of genetics was limited, so interaction of *genes* (and not anything more general) was considered in the definition.

Admitting the DMI definition might be applied not only to genes but also to other genomic elements (MAHESHWARI AND BARBASH 2010), then the *Prdm9* gene and the whole set of genomic PRDM9 binding sites would represent the searched BDM pair fulfilling the (extended) description of BDM incompatibility model.

At a subset of ~250 PRDM9-specified sites per cell, meiotic topoisomerase-like protein SPO11 induces DSBs, creating SPO11-DNA covalent bonds. The creation of DSBs is essential

and a lack leads to pachytene arrest of spermatogenesis and subsequent apoptosis. The *Spo11* null mouse mutant has no DSBs, but artificially induced of DSBs can rescue the meiotic progression in the *Spo11*$^{-/-}$ mouse (CAROFIGLIO *et al.* 2018), as well as improve the insufficient synapsis of chromosomes in prophase I in the (PWD×B6)F1 laboratory mouse hybrid (WANG *et al.* 2018). These findings support the idea of a necessity of proper DSB repair for the progression of spermatogenesis.

While in some lower species, such as *Drosophila melanogaster* and *Caenorhabditis elegans*, the programmed DSBs appear after chromosomal synapsis, in mammals the programmed DSBs precede chromosomal pairing and play a crucial role to enable the proper synapsis of chromosomes. This goes in hand with the amount of the NCOs as they are supposed to play a crucial role in meiotic pairing (TIEMANN-BOEGE *et al.* 2017).

All ~250 programmed meiotic DSBs (positionally determined by PRDM9 and induced by SPO11) need to be repaired in meiosis I, either by crossovers (COs, ~10%), or noncrossovers (NCOs, ~90%). The persistence of unrepaired DSBs in the pachytene or later stage is cell-lethal. In theory, one can imagine the DSB repair performed using the sister chromatids as templates like in sex chromosomes in males (LU AND YU 2015), though an evidence is missing yet.

5.4 Identification and characterization of non-reciprocal products of meiotic recombination

The NCOs, while being a prevalent repair of DSB, are poorly understood. It is due in particular, to a lack of any available NCO-specific cytological markers and the fact they are hard to detect in the genome, due to the small size of their converted tracts that must carry a sequence polymorphism to be visible.

In our study (GERGELITS *et al.* 2019) (**Publication #1**) we identified 94 chromosome-wide sites repaired by meiotic noncrossover (NCO) resolution. This was achieved with the use of a unique mouse model of consomic strains B6.PWD-Chr#. We sequenced (whole genome sequencing, WGS) 10 strains and filtered for candidate sites by comparing to the progenitor strains B6 (*M.m.d.*) and PWD (*M.m.m.*). Assuming an exponential distribution, we determined the mean length of a conversion tract to be 32 bp. The NCOs were detected almost exclusively in the PRDM9-determined hotspots known from previous studies using DMC1 and H3K4me3 ChIP-seq assays (DAVIES *et al.* 2016; SMAGULOVA *et al.* 2016). Furthermore, we identified ~1600

meiotic crossover (CO) sites using a related backcross model. While we confirmed the already known bimodal chromosome-wide distribution of COs with a significant abundance at subtelomeric ends (Cox *et al.* 2009; Liu *et al.* 2014), a significant deficit of NCOs relative to COs was observed at subtelomeric ends.

Remarkably, we observed a significant deficit of NCO and CO events overlapping *nonparental* (i.e. events not found in none of parental strains, neither B6, nor PWD) DSB sites. According to the *PRDM9 binding asymmetry theory*, the *aymmetric/nonparental* sites are suggested to be more problematic for repair in prophase I. Hypothetically they are repaired later and/or by sister chromatid as a template, or not at all (Davies *et al.* 2016; Li *et al.* 2019b), leading to asynapsis of homologous chromosomes in pachytene and later in extreme cases, to sterility (Gregorova *et al.* 2018). Here, we used *nonparentality* as a proxy for *asymmetry* and indeed observed a deficiency of *nonparental* NCO/CO sites. We speculate that, in these cases, the meiocytes could be eliminated. This suggests a further step in the explanation of the molecular mechanisms of hybrid sterility between *M.m.d.* and *M.m.m.* subspecies. These findings are consistent with those in an extensive genome-wide NCO study, where a deficit of both NCO and CO events was inferred based on expected vs. observed H3K4me3 and DMC1 ChIP-seq signal (Li *et al.* 2019b), and the same suggestion of DBSs repair using sister chromatids as templates was offered (Li *et al.* 2019b).

The DNA heteroduplex that arises during DSB repair needs to be resolved to include the base pair that is situated on the initiating or template chromatid. In NCO sites we observed a significant GC bias, indicating the heteroduplex resolution is driven by the mismatching bases. We also observed GC bias surrounding NCO sites, as well as in other known DSBs sites, indicating an already present, i.e. historical GC bias. Together, these two observations represent both just-arising and historical GC bias. We speculated whether the eroded genome could affect the observable gene conversion tract lengths, and if this could explain the differences between $Prdm^{B6}$- (32 bp), $Prdm^{Cast}$- (30 bp) vs $Prdm9^{Hum}$- (41 bp) controlled NCO gene conversion tract lengths (Li *et al.* 2019b).

This observed GC bias can be explained thanks to a large dataset of 1575 NCO events which gave rise to a probable model suggesting the that G/C bases are resistant to conversion (Li *et al.* 2019b). This would have important consequences for surveillence of specific kinds of PRDM9 binding motifs and, consequently, slower evolution of *Prdm9* gene.

The observed deficiency of NCOs relative to COs in subtelomeric regions is consistent with findings in a recent study exploiting strengths of single sperm DNA sequencing, which reported an increased probability of CO repair at the sites proximal to telomeres (HINCH *et al.* 2019). In contradiction, the recent extensive genome-wide study of NCOs did not find a deficiency of NCOs at subtelomeric ends (LI *et al.* 2019b).

Also a higher probability of CO resolution was reported for sites where PRDM9 is bound to both chromatids (HINCH *et al.* 2019; LI *et al.* 2019b), while at the same sites the repair should be faster (HINCH *et al.* 2019). To further support this finding, the preferred repair of earlier formed DSBs by CO resolution was shown in a study investigating comprehensive epigenomic profiles of mouse spermatogenic cells (CHEN *et al.* 2020).

5.5 Further understanding of meiotic homologous recombination

In recent years, the democratization of DNA sequencing and CRISPR/Cas9 technology for null mutant preparation has significantly accelerated the pace of the discovery process, including that of homologous recombination.

5.5.1 Multiple levels of regulations of meiotic homologous recombination

One interesting feature of the recombination is the multiple levels of regulation eventually leading to determination of CO and NCO positions on chromosomes. To summarize the above mentioned pathway in terms of figures: First, ~32 000 PRDM9 binding sites were detected in the genome on so-called "*naked*" DNA *in vitro* (shown for B6 mouse (WALKER *et al.* 2015)). Second, ~15 000-20 000 of these sites are accessible to the PRDM9 protein *in vivo*, based on their detection in genome-wide studies of DMC1 and H3K4me3 sites using various ChIP-seq assays. What is worth noting, the frequency of these sites range within four orders of magnitude. Third, in one study it was calculated that in every spermatocyte, a subset of ~4700 sites per cell is expected to be bound by PRDM9 (BAKER *et al.* 2014). Fourth, out of these PRDM9 bound sites some 200-400 lead to actual DSBs. Finally, in a normal spermatocyte, all of these DSBs need to be repaired, but only 10% by CO, whilst 90% is by NCO. Remarkably, none of the decision making processes between the respective stages of homologous recombination progression is fully understood.

The sites that are bound *in vivo* are more likely to be positioned in open chromatin (WALKER *et al.* 2015) in the chromosomal loops. A recent study also investigated the presence of the seventeen most common histone modifications, here specifically for meiotic prophase I. They also showed a specific presence of H3K9Ac marks associated with DNA DSBs (LAM *et al.* 2019).

Surprisingly, one group mapped origins of premeiotic DNA replication and found that premeiotic replication timing in mouse meiosis appeared to regulate meiotic recombination (PRATTO *et al.* 2019, personal communication). This points to a further level of complexity in DSB position determination, preceding the occurrence of PRDM9.

5.5.2 PRDM9 interactors

Two pull-down studies showed the interaction of PRDM9 with four other proteins EWSR1, CDYL, EHMT2 (PARVANOV *et al.* 2017), and CXXC1 (IMAI *et al.* 2017; PARVANOV *et al.* 2017). It is hypothesized that some PRDM9-interacting proteins tether such sites to the chromosomal axis and later in leptotene, the prespecified sites are found by topo-isomerase like protein SPO11. The first candidate, PRDM9 interactor CXXC1, was shown to be redundant for normal DSB formation and recombination, and was inferred not to tether DSBs to the chromosomal axis (TIAN *et al.* 2018).

Recent studies showed two other PRDM9 interactors of key importance. First, the ATP-dependent chromatin remodeler HELLS interacts with PRDM9 and together form the so-called *pioneer complex* (to have a *pioneer function* means to determine genomic position, deposit histone marks, and move the nucleosomes), which opens the chromatin at sites bound by PRDM9 (SPRUCE *et al.* 2020). The phenotype of a mouse conditional knockout CKO *Hells*$^{-/-}$ is similar to B6.*Prdm9*$^{-/-}$, importantly with the DSB sites repositioned to the *default* hotspots. Co-immunoprecipitation of PRDM9 and HELLS showed interaction of the proteins.

Second, ZCWPW1 was discovered to be recruited to recombination sites by PRDM9 and also to be a dual histone methyl reader of H3K4me3 and H3K36me3, which facilitates the repair of meiotic double strand breaks (HUANG *et al.* 2019; MAHGOUB *et al.* 2019; WELLS *et al.* 2019).

Remarkably, ZCWPW1, the reader of dual PRDM9 histone marks, is required for meiosis prophase I in males but not females, a feature analogous to PRDM9 (LI *et al.* 2019a).

5.6 Current understanding of genetic architecture of (PWD × B6)F1 mouse hybrid sterility model

The X-linked factor *Hstx2* remains the least understood component of (PWD × B6)F1 hybrid sterility architecture. In the recent study, the group of Prof. Jiří Forejt managed to shorten the 4.7 Mb hybrid sterility region, first to 4.3 Mb by NGS sequencing and then to even narrower region of 2.70 Mb. Now *Hstx2* is localized in ChrX: 66.51-69.21 Mb interval (LUSTYK *et al.* 2019) (**Publication #4**).

The original *Hstx2* region, inherently resistant to recombination (recombination coldspot), was challenged to be recombined by specifically tailored techniques of double transgenes bearing Hstx2-CRISPR and SPO11/Cas9 (Cas9 under SPO11 promoter). This effort produced one homologous recombination inside the original *Hstx2* locus and led to a preparation of new congenic strain B6.PWD-Chr X.66-69.

By phenotyping, all three related genetic loci, *Hstx1*, *Hstx2* and *Meir1*, were shown to be localized in the same 2.70 Mb region. Importantly, the region contains two protein genes, *Fmr1nb* and *Fmr1* and two miRNA clusters. The *Fmr1nb* null mutant males were prepared, however the phenotype was not consistent with hybrid sterility, thus excluding the *Fmr1nb* gene from the list of candidates. It is suggested that microRNA *Mir465* cluster could be the plausible candidate for the hybrid sterility *Hstx2* factor, although the responsible sequence remains to be identified.

To summarize, our contemporary understanding of the (PWD × B6)F1 males hybrid sterility architecture encompasses three components:

1) The *Prdm9* gene (MIHOLA *et al.* 2009),
2) Heterozygosity of the genome (GREGOROVA *et al.* 2018), in particular heterozygosity of PRDM9 binding sites, leading to PRDM9 asymmetrical binding (DAVIES *et al.* 2016), and
3) The *Hstx2* locus (DZUR-GEJDOSOVA *et al.* 2012; BHATTACHARYYA *et al.* 2014; LUSTYK *et al.* 2019).

Together the *Prdm9* gene and PRDM9 binding sites represent the first discovered example of a pair of genetic/genomic components fulfilling the (extended) description of Bateson-Dobzhansky-Muller incompatibility model.

6. Conclusions

My thesis was focused on repair of programmed DSBs by crossovers and noncrossovers during meiotic recombination and its role in formation of synaptonemal complex (synapsis) in meiotic prophase I. These processes are necessary for proper progression of spermatogenesis and their defects play important role in hybrid male sterility.

I detected 94 NCOs chromosome-wide and estimated mean length of a conversion tract to be 32 bp. Locally, almost all NCOs were positionally determined by PRDM9, while chromosome-wide, there was a relative deficiency compared to COs at subtelomeric ends. Importantly, a significant deficit of NCOs descending from asymmetric DSBs pointed to their harmful effect on meiotic recombination, suggesting sister chromatids as alternative templates for repair.

Global amount of crossovers per cell (CO rate) is genetically determined. Using methods of immunofluorescense microscopy on panel of chromosome substitution strains, we localized a major modifier of meiotic recombination rate to the 4.7 Mb region in the middle part of Chr X, and denoted it as *Meiotic recombination 1*, *Meir1*. Importantly, *Meir1* locus overlaps the previously identified *Hstx2* locus, the known hybrid sterility factor. Rather unexpectedly, hybrid sterility gene *Prdm9* did not affect CO rate.

We found that chromosomal asynapsis in intersubspecific sterile hybrid males is chromsomes-autonomous and is higher in shorter chromosomes (up to 42%) in intersubspecific hybrids. We identified a threshold amount of 27 Mb of homology per autosomal pair as sufficient for synapsis. Importantly, when the four shortest autosomes were rescued from asynapsis, consequently the male hybrid sterility was rescued, unraveling the *cis*-acting factor (most likely PRDM9 binding sites), necessary for hybrid sterility.

The *Hstx2* hybrid sterility locus, localized in recombination coldspot interval, was narrowed down to 2.7 Mb, still acting as CO rate modifier, *Meir1* locus. The *Fmr1nb* gene was excluded as hybrid sterility candidate, leaving microRNA *Mir465* cluster as a promising candidate for *Hstx2*.

We identified three major components of the investigated F1 hybrid sterility genetics architecture; *Prdm9* gene, intersubspecific F1 heterozygosity of genome likely exhibiting asymmetric PRDM9 binding sites, and *Hstx2* locus. The first two can be considered as the

incompatible pair within Bateson-Dobzhansky-Muller model. In future, DNA sequencing assays, in particular positioning meiotic DSB sites, could answer the remaining questions.

7. List of Abbreviations

ATM	ataxia telangiectasia mutated kinase
ATR	ataxia telangiectasisa and Rad3 related kinase
B6	C57BL/6J (mouse “classical”/”laboratory” strain)
(B6×PWD)F1	F1 hybrid from cross between B6 female and PWD male
B6.PWD-Chr#	Chromosome substitution (consomic) strain with introgressed PWD chromosomes on B6 background
BER	base excision repair
bp; kb; Mb	base pair; kilobase pair; megabase pair
SC	synaptonemal complex
Chr	chromosome
ChIP-seq	chromatin immunoprecipitation followed by sequencing
CI	confidence interval
CO	crossover
DMC1	DNA meiotic recombinase 1
DSB	double-strand break
DSBR	double-strand break repair
F1 hybrid	filial generation 1 hybrid; the first filial generation of offspring
Hstx1	X-linked hybrid sterility locus 1
Hstx2	X-linked hybrid sterility locus 2
Meir1	meiotic recombination 1
MLH1	mutL homolog 1 (*E. coli*)
M.m.c. / M.m.castaneus	*Mus musculus castaneus* (house mouse subspecies)
M.m.d. / M.m.domesticus	*Mus musculus domesticus* (house mouse subspecies)
M.m.m. / M.m.musculus	*Mus musculus musculus* (house mouse subspecies)
MMR	mismatch repair
MSCI	meiotic sex chromosome inactivation
NCO	noncrossover
NGS	“next-generation” sequencing
PRDM9	PR domain containing 9
PWD	PWD/Ph (mouse wild-derived strain)

(PWD×B6)F1	F1 hybrid from cross between PWD female and B6 male
RAD51	RAD51 recombinase
REC8	REC8 meiotic recombination protein
RN	recombination nodule
SDSA	synthesis-dependent strand annealing
SNP	single nucleotide polymorphism
SPO11	SPO11 meiotic protein covalently bound to DSB
SYCE1	synaptonemal complex central element protein 1
SYCE2	synaptonemal complex central element protein 2
SYCP1	synaptonemal complex protein 1
SYCP2	synaptonemal complex protein 2
SYCP3	synaptonemal complex protein 3
ZnF	zinc finger
WGS	whole genome sequencing

8. References

Alberts, B., A. D. Johnson, J. Lewis, D. Morgan, M. Raff *et al.*, 2014 *Molecular Biology of the Cell: Sixth International Student Edition*. Garland Science, Taylor and Francis Group.

Allers, T., and M. Lichten, 2001 Differential timing and control of noncrossover and crossover recombination during meiosis. Cell 106: 47-57.

Baker, C. L., S. Kajita, M. Walker, R. L. Saxl, N. Raghupathy *et al.*, 2015a PRDM9 drives evolutionary erosion of hotspots in Mus musculus through haplotype-specific initiation of meiotic recombination. PLoS Genet 11: e1004916.

Baker, C. L., P. Petkova, M. Walker, P. Flachs, O. Mihola *et al.*, 2015b Multimer Formation Explains Allelic Suppression of PRDM9 Recombination Hotspots. PLoS Genet 11: e1005512.

Baker, C. L., M. Walker, S. Kajita, P. M. Petkov and K. Paigen, 2014 PRDM9 binding organizes hotspot nucleosomes and limits Holliday junction migration. Genome Res 24: 724-732.

Baker, S. M., A. W. Plug, T. A. Prolla, C. E. Bronner, A. C. Harris *et al.*, 1996 Involvement of mouse Mlh1 in DNA mismatch repair and meiotic crossing over. Nat Genet 13: 336-342.

Balcova, M., B. Faltusova, V. Gergelits, T. Bhattacharyya, O. Mihola *et al.*, 2016 Hybrid Sterility Locus on Chromosome X Controls Meiotic Recombination Rate in Mouse. Plos Genetics 12.

Barbash, D. A., D. F. Siino, A. M. Tarone and J. Roote, 2003 A rapidly evolving MYB-related protein causes species isolation in Drosophila. Proc Natl Acad Sci U S A 100: 5302-5307.

Barton, N. H., and G. M. Hewitt, 1989 Adaptation, speciation and hybrid zones. Nature 341: 497-503.

Bateson, W., 1909 Heredity and variation in modern lights. Darwin and modern science.

Baudat, F., J. Buard, C. Grey, A. Fledel-Alon, C. Ober *et al.*, 2010 PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice. Science 327: 836-840.

Baudat, F., Y. Imai and B. de Massy, 2013 Meiotic recombination in mammals: localization and regulation. Nat Rev Genet 14: 794-806.

Bhattacharyya, T., S. Gregorova, O. Mihola, M. Anger, J. Sebestova *et al.*, 2013 Mechanistic basis of infertility of mouse intersubspecific hybrids. Proc Natl Acad Sci U S A 110: E468-477.

Bhattacharyya, T., R. Reifova, S. Gregorova, P. Simecek, V. Gergelits *et al.*, 2014 X chromosome control of meiotic chromosome synapsis in mouse inter-subspecific hybrids. PLoS Genet 10: e1004088.

Boateng, K. A., M. A. Bellani, I. V. Gregoretti, F. Pratto and R. D. Camerini-Otero, 2013 Homologous pairing preceding SPO11-mediated double-strand breaks in mice. Dev Cell 24: 196-205.

Bolcun-Filas, E., and M. A. Handel, 2018 Meiosis: the chromosomal foundation of reproduction. Biology of reproduction 99: 112-126.

Boulton, A., R. S. Myers and R. J. Redfield, 1997 The hotspot conversion paradox and the evolution of meiotic recombination. Proc Natl Acad Sci U S A 94: 8058-8063.

Brick, K., F. Smagulova, P. Khil, R. D. Camerini-Otero and G. V. Petukhova, 2012 Genetic recombination is directed away from functional genomic elements in mice. Nature 485: 642-645.

Carofiglio, F., E. Sleddens-Linkels, E. Wassenaar, A. Inagaki, W. A. van Cappellen *et al.*, 2018 Repair of exogenous DNA double-strand breaks promotes chromosome synapsis in SPO11-mutant mouse meiocytes, and is altered in the absence of HORMAD1. DNA repair 63: 25-38.

Charlesworth, B., J. A. Coyne and N. H. Barton, 1987 The relative rates of evolution of sex chromosomes and autosomes. The American Naturalist 130: 113-146.

Chen, Y., R. Lyu, B. Rong, Y. Zheng, Z. Lin *et al.*, 2020 Refined spatial temporal epigenomic profiling reveals intrinsic connection between PRDM9-mediated H3K4me3 and the fate of double-stranded breaks. Cell Research: 1-13.

Cole, F., F. Baudat, C. Grey, S. Keeney, B. de Massy *et al.*, 2014 Mouse tetrad analysis provides insights into recombination mechanisms and hotspot evolutionary dynamics. Nat Genet 46: 1072-1080.

Cole, F., L. Kauppi, J. Lange, I. Roig, R. Wang *et al.*, 2012 Homeostatic control of recombination is implemented progressively in mouse meiosis. Nat Cell Biol 14: 424-430.

Cox, A., C. L. Ackert-Bicknell, B. L. Dumont, Y. Ding, J. T. Bell *et al.*, 2009 A new standard genetic map for the laboratory mouse. Genetics 182: 1335-1344.

Darwin, C., 1859 On the Origin of Species by Means of Natural Selection or the Preservation of Favored Races in the Struggle for Life. (London:Murray).

Davies, B., E. Hatton, N. Altemose, J. G. Hussin, F. Pratto *et al.*, 2016 Re-engineering the zinc fingers of PRDM9 reverses hybrid sterility in mice. Nature 530: 171-176.

Din, W., R. Anand, P. Boursot, D. Darviche, B. Dod *et al.*, 1996 Origin and radiation of the house mouse: Clues from nuclear genes. Journal of Evolutionary Biology 9: 519-539.

Dobzhansky, T., 1951 in *Genetics and the origin of Species*. Columbia University, New York.

Dumont, B. L., and B. A. Payseur, 2011 Genetic analysis of genome-scale recombination rate evolution in house mice. PLoS Genet 7: e1002116.

Dzur-Gejdosova, M., P. Simecek, S. Gregorova, T. Bhattacharyya and J. Forejt, 2012 Dissecting the genetic architecture of f(1) hybrid sterility in house mice. Evolution 66: 3321-3335.

Edelmann, W., P. E. Cohen, M. Kane, K. Lau, B. Morrow *et al.*, 1996 Meiotic pachytene arrest in MLH1-deficient mice. Cell 85: 1125-1134.

Eram, M. S., S. P. Bustos, E. Lima-Fernandes, A. Siarheyeva, G. Senisterra *et al.*, 2014 Trimethylation of histone H3 lysine 36 by human methyltransferase PRDM9 protein. Journal of Biological Chemistry 289: 12177-12188.

Forejt, J., 1985 Chromosomal and genic sterility of hybrid type in mice and men. Experimental and Clinical Immunogenetics 2: 106-119.

Forejt, J., and P. Ivanyi, 1974 Genetic studies on male sterility of hybrids between laboratory and wild mice (Mus musculus L.). Genet Res 24: 189-206.

Forejt, J., V. Vincek, J. Klein, H. Lehrach and M. Loudova-Mickova, 1991 Genetic mapping of the t-complex region on mouse chromosome 17 including the Hybrid sterility-1 gene. Mamm Genome 1: 84-91.

Geraldes, A., P. Basset, B. Gibson, K. L. Smith, B. Harr *et al.*, 2008 Inferring the history of speciation in house mice from autosomal, X-linked, Y-linked and mitochondrial genes. Molecular ecology 17: 5349-5363.

Gergelits, V., E. Parvanov, P. Simecek and J. Forejt, 2019 Chromosome-wide distribution and characterization of intersubspecific meiotic noncrossovers in mice. BioRxiv: 792226.

Gregorova, S., P. Divina, R. Storchova, Z. Trachtulec, V. Fotopulosova *et al.*, 2008 Mouse consomic strains: exploiting genetic divergence between Mus m. musculus and Mus m. domesticus subspecies. Genome Res 18: 509-515.

Gregorova, S., and J. Forejt, 2000 PWD/Ph and PWK/Ph inbred mouse strains of Mus m. musculus subspecies--a valuable resource of phenotypic variations and genomic polymorphisms. Folia Biol (Praha) 46: 31-41.

Gregorova, S., V. Gergelits, I. Chvatalova, T. Bhattacharyya, B. Valiskova *et al.*, 2018 Modulation of Prdm9-controlled meiotic chromosome asynapsis overrides hybrid sterility in mice. Elife 7.

Gregorova, S., M. Mnukova-Fajdelova, Z. Trachtulec, J. Capkova, M. Loudova *et al.*, 1996 Sub-milliMorgan map of the proximal part of mouse Chromosome 17 including the hybrid sterility 1 gene. Mamm Genome 7: 107-113.

Haldane, J., 1922 Sex ratio and unisexual sterility in hybrid animals. J. Genet. 12: 101-109.

Hayashi, K., K. Yoshida and Y. Matsui, 2005 A histone H3 methyltransferase controls epigenetic events required for meiotic prophase. Nature 438: 374-378.

Hey, J., 2001 The mind of the species problem. Trends in Ecology & Evolution 16: 326-329.

Hinch, A. G., G. Zhang, P. W. Becker, D. Moralli, R. Hinch *et al.*, 2019 Factors influencing meiotic recombination revealed by whole-genome sequencing of single sperm. Science 363: eaau8861.

Hollocher, H., and C.-I. Wu, 1996 The genetics of reproductive isolation in the Drosophila simulans clade: X vs. autosomal effects and male vs. female effects. Genetics 143: 1243-1255.

Holloway, J. K., J. Booth, W. Edelmann, C. H. McGowan and P. E. Cohen, 2008 MUS81 generates a subset of MLH1-MLH3–independent crossovers in mammalian meiosis. PLoS genetics 4.

Huang, T., S. Yuan, M. Li, X. Yu, J. Zhang *et al.*, 2019 The histone modification reader ZCWPW1 links histone methylation to repair of PRDM9-induced meiotic double stand breaks. BioRxiv: 836023.

Imai, Y., F. Baudat, M. Taillepierre, M. Stanzione, A. Toth *et al.*, 2017 The PRDM9 KRAB domain is required for meiosis and involved in protein interactions. Chromosoma.

Ivanyi, P., M. Vojtiskova, P. Demant and M. Mickova, 1969 Genetic factors in the ninth linkage group influencing reproductive performance in male mice. Folia Biol (Praha) 15: 401-421.

Kauppi, L., M. Barchi, J. Lange, F. Baudat, M. Jasin *et al.*, 2013 Numerical constraints and feedback control of double-strand breaks in mouse meiosis. Genes Dev 27: 873-886.

Keeney, S., C. N. Giroux and N. Kleckner, 1997 Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family. Cell 88: 375-384.

Kong, A., G. Thorleifsson, M. L. Frigge, G. Masson, D. F. Gudbjartsson *et al.*, 2014 Common and low-frequency variants associated with genome-wide recombination rate. Nat Genet 46: 11-16.

Lam, K.-W. G., K. Brick, G. Cheng, F. Pratto and R. D. Camerini-Otero, 2019 Cell-type-specific genomics reveals histone modification dynamics in mammalian meiosis. Nature communications 10: 1-11.

Lao, J. P., and N. Hunter, 2010 Trying to avoid your sister. PLoS biology 8.

Li, M., T. Huang, M.-J. Li, C.-X. Zhang, X.-C. Yu *et al.*, 2019a The histone modification reader ZCWPW1 is required for meiosis prophase I in male but not in female mice. Science advances 5: eaax1101.

Li, R., E. Bitoun, N. Altemose, R. W. Davies, B. Davies *et al.*, 2019b A high-resolution map of non-crossover events reveals impacts of genetic diversity on mammalian meiotic recombination. Nature communications 10: 1-15.

Lipkin, S. M., P. B. Moens, V. Wang, M. Lenzi, D. Shanmugarajah *et al.*, 2002 Meiotic arrest and aneuploidy in MLH3-deficient mice. Nature genetics 31: 385-390.

Liu, E. Y., A. P. Morgan, E. J. Chesler, W. Wang, G. A. Churchill *et al.*, 2014 High-resolution sex-specific linkage maps of the mouse reveal polarized distribution of crossovers in male germline. Genetics 197: 91-106.

Lu, L.-Y., and X. Yu, 2015 Double-strand break repair on sex chromosomes: challenges during male meiotic prophase. Cell cycle 14: 516-525.

Lustyk, D., S. Kinský, K. K. Ullrich, M. Yancoskie, L. Kašíková *et al.*, 2019 Genomic structure of Hstx2 modifier of Prdm9-dependent hybrid male sterility in mice. Genetics 213: 1047-1063.

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

Macholan, M., P. Munclinger, M. Sugerkova, P. Dufkova, B. Bimova *et al.*, 2007 Genetic analysis of autosomal and X-linked markers across a mouse hybrid zone. Evolution 61: 746-771.

Maheshwari, S., and D. A. Barbash, 2010 The Genetics of Hybrid Incompatibilities. Annu Rev Genet 45: 331-355.

Mahgoub, M., J. Paiano, M. Bruno, W. Wu, S. Pathuri *et al.*, 2019 Dual Histone Methyl Reader ZCWPW1 Facilitates Repair of Meiotic Double Strand Breaks. BioRxiv: 821603.

Masly, J. P., C. D. Jones, M. A. Noor, J. Locke and H. A. Orr, 2006 Gene transposition as a cause of hybrid sterility in Drosophila. Science 313: 1448-1450.

Mayr, E., 1963 Animal species and evolution. Animal species and evolution.

Mihola, O., F. Pratto, K. Brick, E. Linhartova, T. Kobets *et al.*, 2019 Histone methyltransferase PRDM9 is not essential for meiosis in male mice. Genome Research 29: 1078-1086.

Mihola, O., Z. Trachtulec, C. Vlcek, J. C. Schimenti and J. Forejt, 2009 A mouse speciation gene encodes a meiotic histone H3 methyltransferase. Science 323: 373-375.

Muller, H., and G. Pontecorvo, 1942 Recessive genes causing interspecific sterility and other disharmonies between Drosophila melanogaster and simulans. Genetics 27: 157.

Muller, H. J., 1940 Bearing of the Drosophila work on systematics. The new systematics: 185-268.

Murdoch, B., N. Owen, S. Shirley, S. Crumb, K. W. Broman *et al.*, 2010 Multiple loci contribute to genome-wide recombination levels in male mice. Mamm Genome 21: 550-555.

Myers, S., R. Bowden, A. Tumian, R. E. Bontrop, C. Freeman *et al.*, 2010 Drive against hotspot motifs in primates implicates the PRDM9 gene in meiotic recombination. Science 327: 876-879.

Neale, M. J., and S. Keeney, 2006 Clarifying the mechanics of DNA strand exchange in meiotic recombination. Nature 442: 153-158.

Neale, M. J., J. Pan and S. Keeney, 2005 Endonucleolytic processing of covalent protein-linked DNA double-strand breaks. Nature 436: 1053-1057.

Orr, H. A., 1996 Dobzhansky, Bateson, and the genetics of speciation. Genetics 144: 1331-1335.

Parvanov, E. D., P. M. Petkov and K. Paigen, 2010 Prdm9 controls activation of mammalian recombination hotspots. Science 327: 835.

Parvanov, E. D., H. Tian, T. Billings, R. L. Saxl, C. Spruce *et al.*, 2017 PRDM9 interactions with other proteins provide a link between recombination hotspots and the chromosomal axis in meiosis. Molecular biology of the cell 28: 488-499.

Payseur, B. A., J. G. Krenz and M. W. Nachman, 2004 Differential patterns of introgression across the X chromosome in a hybrid zone between two species of house mice. Evolution 58: 2064-2078.
Petkov, P. M., K. W. Broman, J. P. Szatkiewicz and K. Paigen, 2007 Crossover interference underlies sex differences in recombination rates. Trends in Genetics 23: 539-542.
Phadnis, N., and H. A. Orr, 2009 A single gene causes both male sterility and segregation distortion in Drosophila hybrids. Science 323: 376-379.
Pollard, T. D., W. C. Earnshaw, J. Lippincott-Schwartz and G. Johnson, 2016 *Cell Biology E-Book*. Elsevier Health Sciences.
Powers, N. R., E. D. Parvanov, C. L. Baker, M. Walker, P. M. Petkov *et al.*, 2016 The Meiotic Recombination Activator PRDM9 Trimethylates Both H3K36 and H3K4 at Recombination Hotspots In Vivo. PLoS Genet 12: e1006146.
Presgraves, D. C., L. Balagopalan, S. M. Abmayr and H. A. Orr, 2003 Adaptive evolution drives divergence of a hybrid inviability gene between two species of Drosophila. Nature 423: 715-719.
Presgraves, D. C., and H. A. Orr, 1998 Haldane's rule in taxa lacking a hemizygous X. Science 282: 952-954.
Robert, T., N. Vrielynck, C. Mezard, B. de Massy and M. Grelon, 2016 A new light on the meiotic DSB catalytic complex. Semin Cell Dev Biol 54: 165-176.
Schwarz, T., Y. Striedner, A. Horner, K. Haase, J. Kemptner *et al.*, 2019 PRDM9 forms a trimer by interactions within the zinc finger array. Life science alliance 2.
Silver, L., 1995 Mouse Genetics. Concepts and Applications. Oxford University Press, Oxford, New York.
Smagulova, F., K. Brick, Y. M. Pu, R. D. Camerini-Otero and G. V. Petukhova, 2016 The evolutionary turnover of recombination hot spots contributes to speciation in mice. Genes & Development 30: 266-280.
Smagulova, F., I. V. Gregoretti, K. Brick, P. Khil, R. D. Camerini-Otero *et al.*, 2011 Genome-wide analysis reveals novel molecular features of mouse recombination hotspots. Nature 472: 375-378.
Spruce, C., S. Dlamini, G. Ananda, N. Bronkema, H. Tian *et al.*, 2020 HELLS and PRDM9 form a pioneer complex to open chromatin at meiotic recombination hot spots. Genes & Development 34: 398-412.
Storchova, R., S. Gregorova, D. Buckiova, V. Kyselova, P. Divina *et al.*, 2004 Genetic analysis of X-linked hybrid sterility in the house mouse. Mamm Genome 15: 515-524.
Szostak, J. W., T. L. Orr-Weaver, R. J. Rothstein and F. W. Stahl, 1983 The double-strand-break repair model for recombination. Cell 33: 25-35.
Teeter, K. C., B. A. Payseur, L. W. Harris, M. A. Bakewell, L. M. Thibodeau *et al.*, 2008 Genome-wide patterns of gene flow across a house mouse hybrid zone. Genome Res 18: 67-76.
Teeter, K. C., L. M. Thibodeau, Z. Gompert, C. A. Buerkle, M. W. Nachman *et al.*, 2010 The variable genomic architecture of isolation between hybridizing species of house mice. Evolution 64: 472-485.
Tian, H., T. Billings and P. M. Petkov, 2018 CXXC1 is not essential for normal DNA double-strand break formation and meiotic recombination in mouse. PLoS genetics 14: e1007657.
Tiemann-Boege, I., T. Schwarz, Y. Striedner and A. Heissl, 2017 The consequences of sequence erosion in the evolution of recombination hotspots. Philosophical Transactions of the Royal Society B: Biological Sciences 372: 20160462.

Ting, C.-T., S.-C. Tsaur, M.-L. Wu and C.-I. Wu, 1998 A rapidly evolving homeobox at the site of a hybrid sterility gene. Science 282: 1501-1504.

Trachtulec, Z., O. Mihola, C. Vlcek, H. Himmelbauer, V. Paces *et al.*, 2005 Positional cloning of the hybrid sterility 1 gene: fine genetic mapping and evaluation of two candidate genes. Biological Journal of the Linnean Society 84: 637-641.

Trachtulec, Z., C. Vlcek, O. Mihola, S. Gregorova, V. Fotopulosova *et al.*, 2008 Fine haplotype structure of a chromosome 17 region in the laboratory and wild mouse. Genetics 178: 1777-1784.

True, J. R., B. S. Weir and C. C. Laurie, 1996 A genome-wide survey of hybrid incompatibility factors by the introgression of marked segments of Drosophila mauritiana chromosomes into Drosophila simulans. Genetics 142: 819-837.

Vrielynck, N., A. Chambon, D. Vezon, L. Pereira, L. Chelysheva *et al.*, 2016 A DNA topoisomerase VI–like complex initiates meiotic recombination. Science 351: 939-943.

Vyskocilova, M., Z. Trachtulec, J. Forejt and J. Pialek, 2005 Does gegography matter in hybrid sterility in house mouse? Biological Journal of the Linnean Society 64: 663-674.

Walker, M., T. Billings, C. L. Baker, N. Powers, H. Tian *et al.*, 2015 Affinity-seq detects genome-wide PRDM9 binding sites and reveals the impact of prior chromatin modifications on mammalian recombination hotspot usage. Epigenetics Chromatin 8: 31.

Wang, L., B. Valiskova and J. Forejt, 2018 Cisplatin-induced DNA double-strand breaks promote meiotic chromosome synapsis in PRDM9-controlled mouse hybrid sterility. Elife 7: e42511.

Wells, D., E. Bitoun, D. Moralli, G. Zhang, A. G. Hinch *et al.*, 2019 ZCWPW1 is recruited to recombination hotspots by PRDM9, and is essential for meiotic double strand break repair. BioRxiv: 821678.

Wu, C.-I., and A. W. Davis, 1993 Evolution of postmating reproductive isolation: the composite nature of Haldane's rule and its genetic bases. The American Naturalist 142: 187-212.

Wu, H., N. Mathioudakis, B. Diagouraga, A. Dong, L. Dombrovski *et al.*, 2013 Molecular basis for the regulation of the H3K4 methyltransferase activity of PRDM9. Cell Rep 5: 13-20.

Yamada, S., S. Kim, S. E. Tischfield, M. Jasin, J. Lange *et al.*, 2017 Genomic and chromatin features shaping meiotic double-strand break formation and repair in mice. Cell Cycle: 1-15.

9. Reprints of Publications