

Univerzita Karlova
Přírodovědecká fakulta

Studijní program: Biologie

Studijní obor: Biologie



Lucie Stočesová

Rekombinační hot-spots u různých organismů
Recombination hot-spots in various organisms

Typ závěrečné práce:

Bakalářská práce

Vedoucí práce/Školitel: doc. RNDr. Dana Holá, Ph.D.

Praha, 2020

Poděkování

Chtěla bych upřímně poděkovat vedoucí mé bakalářské práce doc. RNDr. Daně Holé, Ph.D. za ochotu, trpělivost, cenné rady a za čas, který mi věnovala při přípravě mé bakalářské práce. Dále bych ráda poděkovala své rodině a přátelům, kteří mě při studiu hmotně i duševně podpořili.

Prohlášení:

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

V Praze, 13. 5. 2020

Podpis

Abstrakt

Meiotická rekombinace je významný proces, který probíhá v pohlavně se rozmnožujících organismech a vytváří nové kombinace alel. Frekvence i distribuce crossing-overů je ovlivněna mnoha vnitřními a vnějšími faktory. Místa s vyšší frekvencí crossing-overů se nazývají rekombinační hot-spoty a v této práci jsou rozděleny na ancestrální a odvozené. Ancestrální hot-spoty jsou původnější a určují je především histonové modifikace, oblasti bez nukleozomů, promotory a sekvenční motivy. Tento typ hot-spotů mají například kvasinky, ptáci a rostliny. Odvozené hot-spoty jsou určeny proteinem PRDM9, který vyhledává sekvenční motivy a vytváří u nich histonové modifikace. Odvozené hot-spoty jsou typické pro většinu savců, kromě psovitých. Psovití o funkční protein PRDM9 přišli kvůli mutacím genu *Prdm9*. Aktivita proteinu PRDM9 ničí cestou genové konverze původní umístění hot-spotů. Tento proces se nazývá „hot-spot paradox“ a jeho řešením je velmi rychlá evoluce alel genu *Prdm9*.

Tato práce shrnuje základní poznatky o distribuci a určení hot-spotů mezi eukaryotickými organismy. Blíže se zaměřuje na protein PRDM9, jeho stavbu, funkci a evoluci. V závěru práce je nastíněno určení polohy rekombinace u druhů bez hot-spotů a porovnání míry rekombinace mezi různými skupinami organismů.

Klíčová slova: crossing-over, dvouřetězcový zlom, evoluce, histonové modifikace, PRDM9 protein, preferenční místo, rekombinace

Abstract

Meiotic recombination is an important process which occurs in sexually reproducing organisms and creates new allelic combinations. Frequency and distribution of crossing-overs (COs) are affected by many internal and external factors. Regions with higher frequency of COs are called recombination hot-spots and in this thesis, they are divided into ancestral and derived hot-spots. Ancestral hot-spots are the more original ones and they are determined by histone modifications, nucleosome-depleted regions, promoters and sequence motifs. This type of hot-spots occurs, e.g., in yeasts, birds and plants. Derived hot-spots are determined by the PRDM9 protein, which searches for specific sequence motifs and creates histone modifications. These hot-spots are typical for most mammals except the canines which lost functional PRDM9 due to the mutation of the Prdm9 gene. Activity of PRDM9 destroys primary locations of hot-spots via gene conversion. This process is called “hot-spot paradox” and is solved by the rapid evolution of alleles of the Prdm9 gene.

This thesis summarizes basic information on distribution and determination of hot-spots among various eukaryotes. It particularly focuses on the PRDM9 protein, its structure, function and evolution. A determination of location of recombination in species without hot-spots and comparison of recombination rates among different groups of organisms is presented at the end of this thesis.

Key words: crossing-over, double-strand break, evolution, histone modifications, PRDM9 protein, hot-spot, recombination

Obsah

1. Úvod	1
2. Faktory ovlivňující frekvenci a distribuci crossing-overu	2
2.1 Vnitřní faktory	2
2.2 Faktory jedince	3
2.3 Vliv prostředí	3
3. Rekombinační hot-spoty	5
3.1 Ancestrální hot-spoty	6
3.1.1 Houby	6
3.1.1.1 Rod <i>Saccharomyces</i>	6
3.1.1.2 <i>Schizosaccharomyces pombe</i>	7
3.1.1.3 <i>Zyoseptoria tritici</i>	9
3.1.2 Ptáci	9
3.1.3 Rostliny	10
3.2 Odvozené hot-spoty	11
3.2.1 Savci	11
3.2.1.1 PRDM9	12
Stavba PRDM9 a funkce jednotlivých domén	13
Evoluce PRDM9	15
3.2.1.2 Člověk a další primáti	17
3.2.1.3 Psovití	19
3.3 Druhy bez hot-spotů	20
3.3.1 Rod <i>Drosophila</i>	20
3.3.2 <i>Caenorhabditis elegans</i>	21
3.4 Porovnání míry rekombinace skupin a druhů organismů	21
4. Závěr	26
Seznam použité literatury	28

Seznam použitých zkratek

A	adenin
<i>ade6</i>	gen <i>S. pombe</i> , kóduje protein fosforibosylaminoimidazol karboxylázu
Atf1-Pcr1	heterodimerní komplex transkripčních faktorů, který se váže na <i>M26</i>
bp	páry bazí (base pair)
C	cytosin
C2H2	typ zinc finger domény Cys ₂ His ₂ (cystein, histidin) (Cys ₂ His ₂ -like fold group)
CGI	CpG ostrůvky (CpG islands)
cM	centimorgan, jednotka užívaná k měření frekvence rekombinace
CO	crossing-over
CXXC1	CXXC domain containing 1
DSB	dvouřetězcový zlom (double-strand break)
G	guanin
gBGC	genová konverze vytvářející přednostně GC (GC-biased gene conversion)
H2A.Z	varianta histonu H2A
H3K4me3	trimethylace lysinu 4 na histonu H3
H3K9ac	acetylace lysinu 9 na histonu H3
<i>HIS4</i>	gen pro protein HIS4 (histidine biosynthesis trifunctional protein) u <i>S. cerevisiae</i>
kb	kilobáze, 1 000 bp
KRAB	Krüppel-associated box
<i>M26</i>	sekvence vznikající změnou G za T u genu <i>ade6</i> , může fungovat jako hot-spot
Mb	megabáze, 1 000 000 bp
MER3	protein patřící do skupiny ZMM proteinů, známý také jako HFM1 (Helicase Family Member 3)
MMS4	endonukleáza, která vytváří CO třídy II a vytváří komplex s Mus81 (Methyl MethaneSulfonate sensitivity)
MSH4	protein patřící do skupiny ZMM proteinů (MutS Homolog 4)
MSH5	protein patřící do skupiny ZMM proteinů (MutS Homolog 5)
MUS81	endonukleáza, která vytváří CO třídy II a vytváří komplex s Mms4 (MMS and UV-sensitive protein 81)
NAHR	nealeická homologní rekombinace (non allelic homologous recombination)
NCO	non-cross over
ncRNA	nekódující molekuly RNA (non-coding RNA)
NDR	oblast bez nukleozomů (nucleosome depleted region)
NLS	nukleolární lokalizační signál (nucleolar localization signal)
Obr.	obrázek
PRDM9	PR/SET domain containing protein 9
<i>prl</i>	poly-A-bearing RNA without long open reading frames

PR/SET	centrální doména proteinu PRDM9; patří do podtřídy SET domén (PRD1-BF1 and RIZ homologous region)
RR	míra rekombinace uváděná v cM/Mb (recombination rate)
Set1	histon-lysin methyltransferáza H3 lysin-4 specifická u <i>S. cerevisiae</i> (histone-lysine N-methyltransferase, H3 lysine-4 specific)
SNP	jednonukleotidový polymorfismus (single-nucleotid polymorphism)
Spp1	protein nacházející se u <i>S. cerevisiae</i> ; suppressor of PRP protein 1 (complex proteins associated with SET1 protein SPP1)
SSXRD	motiv v proteinu PRDM9 (SSX repression domain)
T	thymin
Tab.	tabulka
TSS	transkripční start (transcription start site)
TTS	místo ukončení transkripce (transcription termination sites)
UV	ultrafialové záření (ultraviolet)
ZIP1	protein patřící do skupiny ZMM proteinů (molecular ZIPer)
ZIP2	protein patřící do skupiny ZMM proteinů (ZIPing up meiotic chromosomes)
ZIP3	protein patřící do skupiny ZMM proteinů, známý je také jako CST9 (Chromosome STability)
ZMM	skupina proteinů, které vytváří CO třídy I, patří tam především ZIP1, ZIP2, ZIP3, MER3, MSH4, MSH5, ZMM znamená ZIP-MSH-MER
ZnF	zinkový prst (zinc finger)

1. Úvod

Rekombinace je velmi významným procesem u pohlavně se rozmnožujících organismů (Villeneuve and Hillers, 2001; de Massy, 2013). Proces rekombinace zajišťuje správnou segregaci homologních chromozomů (Villeneuve a Hillers, 2001). Pro bezchybný průběh rekombinace je nutný alespoň 1 CO (crossing-over), tzv. obligátní CO. Pokud by nedošlo k obligátnímu CO, může dojít k nondisjunkci a aneuploidii (Ritz et al., 2017). Rekombinace vytváří nové kombinace alel a může tak zvýšit fitness potomka (Rice, 2002) i genovou diverzitu v populaci (de Massy, 2013). Tento proces zároveň zvyšuje účinnost selekce (Otto a Lenormand, 2002) a také urychluje adaptaci (Colegrave, 2002; Morran et al., 2009).

Proces rekombinace probíhá při meiosi v profázi I. Dochází při něm k reciproké výměně části genetické informace mezi otcovským a mateřským chromozomem jedince (tedy mezi homologními chromozomy) a to uvnitř prekurzorů gamet (San Filippo et al., 2008; Phadnis et al., 2011; de Massy, 2013).

K rekombinaci ale v různých místech chromozomů nedochází stejně často. V některých úsecích (centromerách, telomerách) by rekombinace mohla mít fatální důsledky, protože by ohrozila bezchybnou segregaci chromozomů (Gerton et al., 2000; Rockmill et al., 2006). Naopak některé části chromozomů jsou pro rekombinaci vhodnější a místům, kde k rekombinacím dochází velmi často, říkáme rekombinační hot-spoty (Lichten a Goldman, 1995). Tyto hot-spoty se mezi různými organismy hodně liší, a to i v případě, že jsou si příslušné organismy blízce příbuzné. Rozdíly mohou být jak v umístění hot-spotů, tak v počtu či v mechanismu jejich určení.

Právě rekombinační hot-spoty a rozdíly v nich mezi různými druhy jsou náplní mé bakalářské práce. Jsou zde shrnuty současné informace o hot-spotech u různých skupin eukaryotických organismů, základní srovnání frekvence rekombinace mezi skupinami organismů a také nastínění mechanismu určení polohy rekombinace u organismů bez hot-spotů.

Cíle této bakalářské práce jsou:

- I.** Představit základní faktory mající vliv na rekombinaci.
- II.** Shrnout současné poznatky o rekombinačních hot-spotech u různých skupin či druhů eukaryotických organismů a více se zaměřit na protein PRDM9, který určuje rekombinační hot-spoty u savců.
- III.** Porovnat míru rekombinace mezi různými skupinami organismů.

2. Faktory ovlivňující frekvenci a distribuci crossing-overu

Na umístění a počet CO má vliv velké množství faktorů. V úvodní kapitole mé práce jsem je rozdělila do 3 hlavních skupin: vnitřní faktory, faktory jedince a vliv prostředí. Nejprve stručně shrnu esenciální mechanismus určování polohy CO a DSB (dvouřetězcových zlomů), tj. CO a DSB interferenci. Následně přejdu k faktorům majícím vliv na polohu a frekvenci CO.

Byly popsány 2 třídy CO (a to poprvé u kvasinky pивní (*Saccharomyces cerevisiae* (Meyen ex E. C. Hansen))) (**de los Santos et al., 2003**). V první třídě CO (značené v literatuře jako „*class I.*“) působí CO interference (**Zalevsky et al., 1999**). CO interference je proces, kdy při vytvoření CO v jednom regionu dochází k výraznému snížení pravděpodobnosti, že se v okolí vytvoří další CO. Míra interference klesá se vzdáleností od CO (**Jones a Franklin, 2006**). Podobný proces je spojeny i s tvorbou DSBs. Dochází buď k cis-interferenci (na tom samém homologu) nebo trans-interferenci (na druhém homologu v analogické pozici) (**Zhang et al., 2011; Garcia et al., 2015**). U většiny studovaných organismů tvoří 85 % ze všech CO (**Börner et al., 2004; Higgins, 2004**). Druhá skupina (značená v literatuře jako „*class II.*“) pod vlivem interference není. Obě tyto skupiny se liší i biochemicky (**de los Santos et al., 2003**). První třída CO vyžaduje skupinu proteinů, které se značí jako ZMM proteiny (mezi ně patří Zip1, Zip2, Zip3, Msh4, Msh5 a Mer3) (**Börner et al., 2004**). Druhá třída CO je určena dimerem endonukleáz Mus81-Mms4 (**de los Santos et al., 2003**).

2.1 Vnitřní faktory

Jedním z vnitřních faktorů, které souvisí s mírou rekombinace, je velikost genomu. U velkých genomů rostlin byla objevena negativní korelace s mírou rekombinace. Pravděpodobné vysvětlení je, že velké genomy rostlin obsahují hodně repetitivních sekvencí, ve kterých rekombinace neprobíhá tak často. U živočichů a hub nebyl nalezen jednotný vztah míry rekombinace s velikostí genomu (**Tiley a Burleigh, 2015; Stapley et al., 2017**).

Rozdíly byly nalezeny i v rámci polyploidie. Allopolyploidní jedinci mají vyšší míru rekombinace než jejich diploidní předchůdci (**Leflon et al., 2010**). Naopak autopolyploidní jedinci mají nižší míru rekombinace než jejich diploidní předchůdci (**Yant et al., 2013**). Navíc nově vzniklé polyploidní organismy mají vysokou míru rekombinace, která s postupem času klesá (**Le Comber et al., 2010**).

Stav chromatinu je druhým důležitým faktorem. V heterochromatinu, který obsahuje více repetitivních sekvencí, k rekombinaci dochází velmi málo. Naopak v genově bohatém

euchromatinu je četnost rekombinace vyšší (**Termolino et al., 2016**). Rozvolněnosti euchromatinu (tzv. open chromatin) napomáhají určité histonové modifikace a tím je umožněn lepší přístup rekombinačního komplexu k vláknu DNA. Typické histonové modifikace spojené se zvýšením četnosti rekombinací jsou uvedeny v dalších částech mé práce (například v části 3.2.1.1).

2.2 Faktory jedince

Větší rozdíly v míře rekombinace i v umístění CO jsou mezi pohlavími. Záleží na druhu organismu, zda mají frekvenci CO vyšší samci či samice. V tomto ohledu se liší i blízké příbuzné druhy. Příkladem jsou druhy myš domácí (*Mus musculus castaneus* (Waterhouse)) a myš západoevropská (*Mus musculus domesticus* (Schwarz & Schwarz)). U prvního z druhů je vyšší míra rekombinace u samic, zatímco u myši západoevropské je vyšší míra rekombinace u samců (**Dumont a Payseur, 2011**).

Frekvence CO se může měnit i s věkem. Například u člověka se u žen s rostoucím věkem zvyšuje i míra rekombinace (**Kong et al., 2004**). Důvodem tohoto vztahu by mohla být například klesající CO interference. I CO interference totiž s věkem matky klesá (**Campbell et al., 2015**).

2.3 Vliv prostředí

Výrazný vliv má především teplota. Extrémní hodnoty teploty mohou porušit synaptonemální komplex a kvůli tomu dojde ke snížení četnosti rekombinací a snížení fertility. Vliv vysoké teploty na rekombinaci byl ve studii z roku 1948 zkoumán na několika rostlinách, mezi které patří například pšenice setá (*Triticum aestivum* (L.)), vikev ptačí (*Vicia cracca* (L.)) a žito seté (*Secale cereale* (L.)) (**Pao a Li, 1948**). Novější studie provedená u ječmene prokázala nelineární vztah (kde tvar křivky připomíná obrácené písmeno U) mezi mírou rekombinace a teplotou. Stejnou závislost objevili vědci i mezi mírou rekombinace a roční solární radiací. Positivní lineární vztah byl nalezen také s úhrnem ročních srážek, vliv měla i UV radiace a stav živin (**Dreissig et al., 2019**).

Dalším faktorem je infekce patogenním organismem. Infikovaní jedinci vykazují vyšší míru rekombinace než ti zdraví. Je to nejspíše způsobeno koevolučním soubojem hostitele a patogena. Vyšší míra rekombinace infikovaného jedince byla objevena například u tabáku virginského (*Nicotiana tabacum* (L.)) (**Kovalchuk et al., 2003**) nebo u potměníka hnědého (*Tribolium castaneum* (Herbst)) (**Kerstes et al., 2012**).

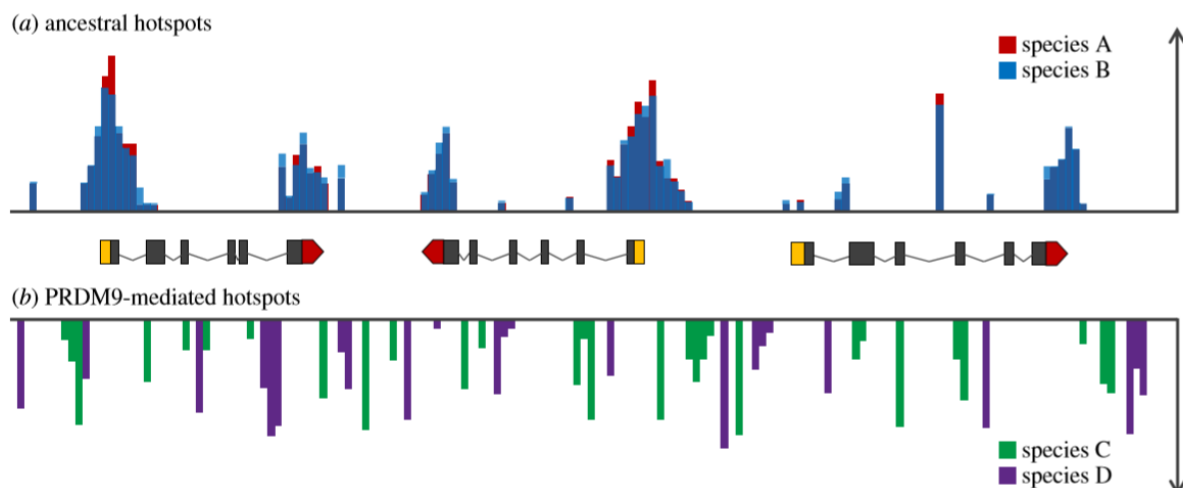
Prokazatelně významný vliv na četnost rekombinace má eusocialita. U eusociálního hmyzu byla naměřena mnohem vyšší míra rekombinace než u jejich solitérních příbuzných či předků. Příkladem je včela medonosná (*Apis mellifera* (L.)) (Beye et al., 2006) nebo vosa obecná (*Vespula vulgaris* (L.)) (Sirviö et al., 2011).

V minulosti se některé studie zaměřily na vliv domestikace. Domestikovaní jedinci vykazovali vyšší míru rekombinace než jejich divocí předci. Je tomu tak například mezi původnějším křížencem kura bankivského (*Gallus gallus* (L.)) a leghornky bílé (*Gallus gallus* f. *domestica* „White leghorn“), a domestikovanými potomky plymutky bílé (*Gallus gallus* f. *domestica* „Plymouth Rock White“) (Groenen et al., 2008). Vyšší míra rekombinace byla naměřena i u domestikovaného druhu ječmene setého (*Hordeum vulgare* spp. *vulgare* (L.)) ve srovnání s ancestrálnějším ječmenem (*Hordeum spontaneum* (K. Koch)) (Dreissig et al., 2019). Avšak v studii o domestikovaných savcích byla naopak popsána míra rekombinace vyšší u jejich divokých předků. Ve studii z roku 2015 je porovnána míra rekombinace těchto druhů (vždy v pořadí domestikovaný druh a původnější druh): pes domácí (*Canis familiaris* (L.)) a vlk obecný (*Canis lupus* (L.)), koza domácí (*Capra hircus* (L.)) a kozorožec iberský (*Capra pyrenaica* (Schinz)), ovce domácí (*Ovis aries* (L.)) a muflon (*Ovis musimon* (Pallas)) (Muñoz-Fuentes et al., 2015). Je tedy otázkou, zda domestikace má skutečně vliv na míru rekombinace a u jakých skupin organismů.

3. Rekombinační hot-spoty

Oblasti chromosomů, kde dochází k rekombinaci častěji, než je celogenomový průměr, se označují jako rekombinační hot-spoty (Lichten a Goldman, 1995). V hot-spotech může výskyt CO narůst o 2 – 10 řádů než je průměrný počet CO v celém genomu (Stapley et al., 2017). Thomas D. Petes ve svém článku z roku 2001 rozdělil hot-spoty do 3 kategorií. α hot-spoty jsou takové, které vyžadují navázání transkripčního faktoru pro svou aktivaci. Mezi ně patří například *ade6-M26* (White et al., 1993; Petes, 2001). β hot-spoty jsou charakterizovány specifickou sekvencí (Kirkpatrick et al., 1999; Petes, 2001). A poslední γ hot-spoty jsou spojeny s vyšším obsahem GC (Gerton et al., 2000; Petes, 2001).

V práci z roku 2016 jsou hot-spoty rozděleny jinak, a to na ancestrální, tedy původnější, a odvozené (Lenormand et al., 2016). Takto je rozdělují i v mé bakalářské práci a srovnání umístění rekombinací u obou typů je ukázáno na Obr. 1. Jejich bližší charakteristika je v následujících kapitolách 3.1 a 3.2.



Obrázek 1: Umístění CO u ancestrálních a odvozených hot-spotů. (a) ancestrální hot-spoty – červeně druh A, modře druh B; (b) odvozené hot-spoty (působení proteinu PRDM9) – zeleně druh C, fialově druh D; osa x: genomová sekvence obsahující 3 geny (žlutě – transkripční počátek (TSS), červeně – místo ukončení transkripce (TTS), tmavo-šedě – exony, linie mezi nimi – introny), osa y: míra rekombinace v cM/Mb

(a) V ancestrálních hot-spotech je u obou druhů nejvyšší míra rekombinace v transkripčních počátcích genů a o něco menší v místech ukončení transkripce. Vybrané druhy mají velmi podobné umístění rekombinací. (b) Naopak u druhů s odvozenými hot-spoty je vidět velká variabilita rozložení rekombinací, která je dána rozdíly v jedné z domén proteinu PRDM9. Není zde vidět žádná větší souvislost s TSSs nebo TTSS (Lenormand et al., 2016).

3.1 Ancestrální hot-spoty

Jako ancestrální hot-spoty označujeme původnější typy hot-spotů. Jsou velmi stabilní v čase (milióny let) a nacházejí se především v oblastech blízko promotorů a transkripčních počátků (Obr. 1). Takovéto hot-spoty mají např. houby (prozkoumané jsou nejvíce kvasinky), rostliny a ptáci (**Lenormand et al., 2016**).

3.1.1 Houby

3.1.1.1 Rod *Saccharomyces*

Kvasinky rodu *Saccharomyces* mají vysoce konzervované hot-spoty. Tato místa se znatelně vyšší mírou rekombinace jsou dlouhá 2 – 5 kb (kilobází) u kvasinky *S. paradoxus* (Bach.-Raich.) a kolem 3 kb u kvasinky *S. cerevisiae*. Oba druhy mají velmi podobné hot-spoty, a to ve smyslu obsahu GC párů a v rozmístění hot-spotů napříč genomem. Tyto společné znaky si pravděpodobně uchovaly od společného předka. Možnou variantou však může být, že hot-spoty vznikly v historii vícekrát, což by takto malý genom umožňoval (**Tsai et al., 2010**). Podle nedávných studií jsou hot-spoty kvasinek zachované už 15 milionů let (**Lam a Keeney, 2015**). Míra zachování častých cílů rekombinace je způsobena sníženým působením gBGC (genová konverze přednostně vytvářející GC). Ke genové konverzi dochází velmi často právě při rekombinaci. Tomuto jevu předchází chybné párování bází, ke kterému může během rekombinace dojít (**Dressler a Potter, 1982**). Chyby v párování bází se opravují tzv. „*mismatch*“ opravou. Při procesu gBGC dochází preferenčně k opravě chybného páru na pár GC a tato dvojice nukleotidů se následně hromadí v místech, kde probíhá rekombinace často (**Brown a Jiricny, 1988**).

Ke snížení vlivu gBGC u kvasinky *S. cerevisiae* došlo zřejmě kvůli menší četnosti sexuálního rozmnožování a také kvůli omezení mezidruhového křížení. Oba tyto jevy byly v historii častější a umožnily tak silnější působení gBGC (**Tsai et al., 2010**). Ačkoliv hot-spoty kvasinek stále obsahují zvýšený obsah GC párů (právě kvůli dřívějšímu působení gBGC) (**Gerton et al., 2000**), v recentnějším časovém měřítku byl pozorován zcela opačný trend. V hot-spotech dochází častěji ke změně GC na AT. K tomuto obratu mohlo dojít kvůli zmírnění tlaku gBGC a v budoucnu se korelace mezi hot-spoty a obsahem GC může ještě snížit (**Tsai et al., 2010**).

Hot-spoty kvasinek se nachází nejčastěji v intergenických oblastech, respektive v promotorech (**Pan et al., 2011**). Promotory jsou funkční elementy genomu, které jsou nezbytné pro správnou funkci genů, a proto je u nich evoluce omezená. I to je zřejmě důvod

stálosti těchto hot-spotů (**Lam a Keeney, 2015**). Pro častou rekombinaci a s tím související umístění hot-spotů je velmi důležitý volný chromatin bez nukleozomů (jinak také nazýván open chromatin nebo NDR) (**Petes, 2001; Pan et al., 2011**). Tyto volné regiony se nachází často právě v promotorech a jsou u kvasinek průměrně 147 bp (bazických párů) dlouhé. V jejich blízkosti se nachází +1 nukleozom, který je počátečním prvkem transkripčního počátku (TSS) a od NDR je dál ve směru transkripce (tzv. downstream) (**Kaplan et al., 2009; Radman-Livaja a Rando, 2010; Pan et al., 2011**).

Důležitým faktorem pro výskyt hot-spotů je také trimethylace lysinu 4 na histonu H3 (H3K4me3). Tato histonová modifikace je u kvasinky *S. cerevisiae* vytvářena methyltransferázou Set1 a při jejím odstranění klesne počet DSB (a tedy i výsledných počet CO) o 84 %. Takto silná redukce značí významnost histonové modifikace H3K4me3 (**Borde et al., 2009**). Kvasinky rodu *Saccharomyces* navíc nejsou zdaleka jedinou skupinou organismů, kde H3K4me3 hraje zásadní roli (viz část 3.2.1. Savci).

Vliv na rekombinační hot-spoty mohou mít i transkripční faktory. Uvažuje se, že by mohly měnit strukturu chromatinu, popřípadě by mohly interagovat s komplexem vytvářejícím DSBs. Příkladem je lokus *HIS4*, kde byla naměřena vyšší míra rekombinace při navázání transkripčních faktorů. Důvodem této stimulace je pravděpodobně to, že navázání transkripčních faktorů může vytvořit otevřený chromatin, který je vhodný jak pro rekombinaci, tak pro transkripci (**White et al., 1993**). Pozdější výzkumy ale spíše naznačují, že větší vliv na rekombinaci mají jiné faktory. Klíčovou roli má rozhodně struktura chromatinu a přítomnost nukleozomů v daném regionu (**Petes, 2001; Pan et al., 2011**).

Rekombinace (primárně tvorba DSBs) je u kvasinek rodu *Saccharomyces* naopak potlačena v repetitivních sekvencích jako jsou telomery a centromery (**Gerton et al., 2000; Blitzblau et al., 2007**). Dochází k tomu proto, aby tyto sekvence nepodléhaly nealelické homologní rekombinaci (NAHR) a nedocházelo k přestavbám, nebo aby nedocházelo k chybám při segregaci sesterských chromatid (to platí především pro centromery) (**Rockmill et al., 2006; Sasaki et al., 2010**).

3.1.1.2 *Schizosaccharomyces pombe*

Stejně jako kvasinky rodu *Saccharomyces*, i kvasinky poltivé (*Schizosaccharomyces pombe* (Lindner)) mají hot-spoty. Jejich průměrná délka se pohybuje v rozmezí 1 – 2 kb a hlavním společným znakem většiny jejich hot-spotů jsou NDRs, tedy úseky otevřeného chromatinu bez nukleozomů. Vyskytují se ve velkých intergenických regionech (větších než

3 kb), které jsou od sebe vzdálené přibližně 65 kb (**Cromie et al., 2007; de Castro et al., 2012**). Na rozdíl od kvasinky *S. cerevisiae* nejsou hot-spoty kvasinky *S. pombe* příliš často umístěny v promotorech (**Cromie et al., 2007; Fowler et al., 2014**). Zvláštností tohoto druhu je tzv. CO invariance. Je to proces, který udržuje celkovou četnost COs velice uniformní. Počet DSBs se v různých místech genomu rapidně mění, zatímco počet COs na jednotku délky (nejčastěji na kb DNA) zůstává stále stejný. Tento mechanismus je založen na volbě způsobu opravy DSB, a to buď pomocí homologního chromozomu nebo pomocí sesterské chromatidy. V silných DSB hot-spotech dochází nejčastěji k opravě sesterskou chromatidou, zatímco v místech s malým počtem DSBs dochází k opravě pomocí homologního chromozomu a tedy k CO (**Hyppa a Smith, 2010**).

Významnou roli mají i zde histonové modifikace, a to konkrétně acetylace lysinu 9 na histonu H3 (zkráceně H3K9ac) (**Yamada et al., 2004**). Naopak v hot-spotech kvasinky *S. pombe* nebylo nalezeno zvýšené množství H3K4me3, jako je tomu u jiných druhů. Nicméně, přestože spojitost mezi hot-spoty kvasinky *S. pombe* a H3K4me3 nebyla prokázána, při delecii methyltransferázy Set1 dochází k redukci počtu DSBs. Set1 se tedy jeví jako možný faktor, který ovlivňuje hot-spoty u tohoto druhu kvasinek (**Yamada et al., 2013**).

U kvasinky *S. pombe* byla objevena možná korelace vazby transkripčních faktorů a transkripce s některými hot-spoty. Jedná se zejména o hot-spot *M26*, který vznikl záměnou 1 nukleotidu G za T v genu *ade6*. Výsledkem je sekvence ATGACGT (kde podtržený nukleotid je ten změněný) (**Ponticelli et al., 1988**), která na sebe váže komplex transkripčních faktorů Atf1-Pcr1 (**Kon et al., 1997**). Sekvence *M26* a jí podobné sekvence, obecně zapsané jako (C/G/T)TGACGT(C/A), po navázání transkripčních faktorů fungují jako hot-spoty (**Fox et al., 2000**). Atf1 stimuluje přepis krátkých transkriptů, které podporují tvorbu H3K9ac a lokálně pak dochází k dekonduzaci okolního chromatinu. Vytvořením NDR se v dané lokalitě ustanoví vhodné prostředí pro tvorbu DSBs. Malá část hot-spotů tedy koreluje s transkripcí a přepisovanými regiony (**Yamada et al., 2017**).

Zajímavým faktorem, majícím vliv na hot-spoty a jejich umístění, jsou polyadenylované nekódující molekuly RNA (ncRNA) patřící do třídy *pri*. V místech, kde se tyto ncRNA exprimují, jsou velmi často i DSB hot-spoty, a tedy i čtenější rekombinace. Spekuluje se, zda exprese ncRNA napomáhá vytvoření otevřeného chromatinu nebo zda tyto molekuly fungují jako průvodce, který navádí homologní DNA do hot-spotů k následné rekombinaci (**Wahls et al., 2008**).

3.1.1.3 *Zymoseptoria tritici*

Braničnatka pšeničná (*Zymoseptoria tritici* ((Desm.) Quaedvl. & Crous)) je houbový patogen pšenice a z hlediska rekombinací je velice zajímavý. Stejně jako některé druhy živočichů (například šimpanzi, lidé – více v části 3.2.1.2), má nejvyšší míru rekombinace v subtelomerické oblasti. Prochází častou pohlavní reprodukcí, která způsobuje, že hot-spoty této patogenní houby nejsou tolik konzervované, jako je tomu například u kvasinky *S. cerevisiae*. Velké množství hot-spotů je uloženo v okolí genů, které kódují patogenní proteiny. Častá rekombinace zajišťuje rychlou evoluci a vytváření nových forem toxinů (Croll et al., 2015). Odlišností od jiných organismů je, že v centroměře nebyla naměřena téměř žádná suprese rekombinace. Významnou roli zde nehraje ani proces gBGC, který se u braničnatky jeví jako slabý či s žádným efektem na výskyt hot-spotů (Stukenbrock a Dutheil, 2018).

3.1.2 Ptáci

Ptáci mají hot-spoty, avšak jejich genom neobsahuje *Prdm9* gen, jako je tomu u savců (Oliver et al., 2009; Singhal et al., 2015). Při srovnání 2 druhů zebřičky pestré (*Taeniopygia guttata* (Vieillot)) a pásovníka dlouhoocasého (*Poephila acuticauda* (Gould)) (oba z řádu pěvců) bylo objeveno, že mají 73 % hot-spotů stejných (a to především co se týče jejich umístění). Podobně tomu bylo i u dalších druhů. Analýza ukazuje, že rozložení ptačích hot-spotů zůstalo velmi podobné přes 10 milionů let. To značí, že jsou vysoce konzervované (Singhal et al., 2015). Shodně je tomu i u kvasinky *S. cerevisiae*, kde je míra zachování hot-spotů ještě vyšší a jsou stálé už 15 milionů let (Lam a Keeney, 2015).

Ptačí hot-spoty se v polovině případů nacházejí blízko TSS, vedle místa ukončení transkripce (TTS) nebo blízko CpG ostrůvků (CGIs). Hot-spoty blízko CGIs jsou mezi druhy sdílené nejvíce a jsou také hodně ovlivněny procesem gBGC. Míra rekombinace je u TSS a TTS 2krát větší a u CGIs 3krát větší, než je průměr míry rekombinace ve zbytku genomu (Singhal et al., 2015). Všechny tyto 3 regiony jsou funkčními elementy a jsou spojeny s nestabilitou okolních nukleozomů. Tím se v nich vytváří místo vhodné pro přístup rekombinačního komplexu. Hot-spoty ptáků mají vysoký obsah adeninu, který by mohl podporovat posun nukleozomů a vytvoření NDRs (Kaplan et al., 2009; Singhal et al., 2015). Celkově dochází častěji k rekombinaci v exonech než v intronech a míra rekombinace stoupá směrem ke koncům genů (tedy uprostřed genů je nejmenší). Stejně tak DSB se přednostně

vytváří v exonech poblíž TSS a TTS, ovšem mechanismus, který by vysvětloval tuto preferenci, je zatím neznámý (Singhal et al., 2015).

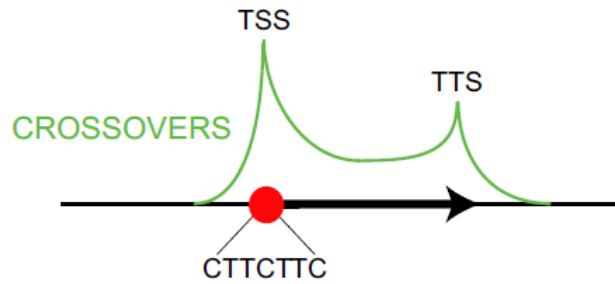
3.1.3 Rostliny

Hot-spoty rostlin byly zatím nejvíce prozkoumány u huseníčku rolního (*Arabidopsis thaliana* ((L.) Heynh)), kde vykazují až 50krát více COs než je genomový průměr (Drouaud et al., 2013). Nalezeny byly ale i v dalších rostlinách jako je kukuřice (Liu et al., 2009), rýže (Marand et al., 2019) a další rostliny uvedené ve článku Wang a Copenhaver (2018). U huseníčku vykazují hot-spoty silnou asociaci s TSSs a o něco méně pak s TTSs. V pozici +1 u TSS se nachází varianta histonu H2A.Z, která se překrývá s nejvyšší mírou rekombinace (Obr. 2). Pro hot-spoty tohoto druhu rostlin je také typický A-bohatý motiv, který je dlouhý 6 – 30 bp a nachází se nad TSS, tzn. upstream (Choi et al., 2013; Wijnker et al., 2013). Zvýšená přítomnost A lépe umožňuje vznik NDR, což napomáhá tvorbě zlomů a následné rekombinaci (Field et al., 2008). Druhým významným motivem je trimer CTT, který se opakuje v délce 6 – 21 bp a nachází se především za TSSs, tzn. downstream, takže se překrývá s H2A.Z (Obr. 2) (Choi et al., 2013).

Pozdější studie ukázaly, že častým motivem je také CCN, kde N značí jakýkoliv nukleotid (z výběru A, T, C, G). Tento motiv se vyskytuje převážně uvnitř genů, na rozdíl od předchozích motivů, které se vyskytovaly především v oblastech promotorů, TSSs a TTSs. V genech se podle této analýzy nalézá 42 % CO, což je více než v promotorech (tam je 31 % CO). Motiv CCN je spojen s přítomností nukleozomů, které nesou epigenetickou značku H3K4me3 (Shilo et al., 2015).

Typickým znakem je také hypomethylace DNA v TSSs a TTSs, v nichž se nachází hot-spoty. Naopak v transkripčních počátcích a v místech ukončení transkripce, ve kterých se hot-spoty nevyskytují, nebyla snížená míra methylace DNA pozorována. Celkově je methylace DNA naopak zvýšená uvnitř genů (směrem od TSS stoupá) a zejména v exonech (Cokus et al., 2008).

Snížený počet COs se vyskytuje ve většině repetitivních sekvencí, ale byly popsány i výjimky. U rýže seté (*Oryza sativa* (L.)) existují 2 transpozony *Stowaway* a *PIF/Harbinger*, které se často nacházejí právě v místech se zvýšenou mírou rekombinace (nad a pod TSSs) a zvyšují aktivitu hot-spotu. U kukuřice je v hot-spotech vyšší obsah GC, způsobený procesem gBGC. Uvažuje se, že transpozony bohaté na AT se vkládají do úseků bohatých na GC a tím vyrovnávají poměr nukleotidů AT:CG (Marand et al., 2019).



Obrázek 2: Schéma rekombinačních hot-spotů u *A. thaliana*. TSS – transkripční počátek, TTS – místo ukončení transkripce, CTT – hlavní motiv v TSS, červený bod – nukleozom +1 obsahující H2A.Z a H3K4me3 (Choi a Henderson, 2015)

3.2 Odvozené hot-spoty

Tento typ se vyvinul z ancestrálních hot-spotů a je tedy mladší. Jeho pozice je určena proteinem PRDM9 nebo jeho orthologem a vyskytuje se u většiny savců (Lenormand et al., 2016). Tyto hot-spoty byly prozkoumané především u myši a člověka.

3.2.1 Savci

Rekombinační hot-spoty savců jsou dlouhé průměrně 1 – 2 kb (Kauppi et al., 2004). Z celkového počtu se 80 % rekombinací stane na 10 – 20 % genomu. Vzdálenost mezi jednotlivými hot-spoty se mezi druhy liší. U lidí je to v průměru 50 – 100 kb (Myers et al., 2005), zatímco u myši jsou od sebe hot-spoty vzdálené 60 – 330 kb (Smagulova et al., 2011).

Umístění savčích hot-spotů určuje protein PRDM9 (Baudat et al., 2010; Myers et al., 2010; Parvanov et al., 2010). Hot-spoty savců mají dvojitou značku H3K4me3 a H3K36me3, která je pouze v hot-spotech a vytváří ji pouze protein PRDM9 (Powers et al., 2016). Histonová modifikace H3K4me3 se jinak běžně nachází v aktivních promotorech (Heintzman et al., 2007) a enhancerech (Pekowska et al., 2011). Tato značka je spojena s vytvořením NDRs, protože se nachází vždy z obou stran regionu NDR (Baker et al., 2014). Druhá z modifikací H3K36me3 je značkou transkripční elongace (Li et al., 2007).

3.2.1.1 PRDM9

Protein PRDM9 (zkratka z anglického názvu *PR/SET domain containing protein 9*) určuje umístění DSBs a hot-spotů svojí přímou vazbou na vlákno DNA. Jako determinant lokalizace hot-spotů byl popsán v roce 2010 třemi skupinami zároveň (**Baudat et al., 2010; Myers et al., 2010; Parvanov et al., 2010**). V případě *knockoutu* (tedy vypnutí) genu *Prdm9* u myši se pozice hot-spotů výrazně změní. Myš s vypnutým genem *Prdm9* stále má hot-spoty, ale téměř polovina z nich (44 %) se nachází ve funkčních elementech genomu jako jsou promotory a enhancery. Oba tyto elementy nesou značku H3K4me3, a proto se tam hot-spoty v případě vypnutí *Prdm9* přemístí. Hot-spoty těchto myši připomínají rozložením ancestrální hot-spoty kvasinky *S. cerevisiae*. PRDM9 je tedy zodpovědný za umístění hot-spotů a jeví se jako odvozenější a mladší způsob určování rekombinačních hot-spotů (**Brick et al., 2012**).

Protein PRDM9 se nachází pouze ve varlatech a vaječnicích a k expresi jeho genu dochází u obou pohlaví v meiotické profázi I., konkrétně je přítomen od stádia leptotene do stádia zygotene (**Hayashi et al., 2005; Sun et al., 2015**). Je označován jako gen hybridní sterility obratlovců. Hybridní sterilita je jedním z postzygotických mechanismů, které reprodukčně izolují odlišné druhy a mají roli ve speciaci. Popsání této funkce genu *Prdm9* bylo provedeno na myši střeoevropské (*Mus musculus musculus* (L.)) a myši západoevropské. Z fertilních rodičovských myši vznikl sterilní hybrid. Sterilita se zde prokázala pouze u samčího pohlaví, u něhož došlo k selhání ve spermatogenezi (**Mihola et al., 2009**).

Ačkoliv PRDM9 se přiřazuje převážně k savcům, protože většina savců tento gen kóduje, může být přítomný i u dalších tříd. U psovitých se z *Prdm9* stal pseudogen a ztratil funkci určování pozice rekombinací (**Oliver et al. 2009, Muñoz-Fuentes et al. 2011**). Je vhodné zmínit, že nebyli prozkoumáni všichni savci a není tak zcela jasné, že nefunkční gen *Prdm9* se u savců nachází pouze u psovitých. V rámci savců byl gen *Prdm9* nalezen například u slona afrického (*Loxodonta africana* (Blumenbach)), kočky domácí (*Felis catus* (L.)), tura domácího (*Bos taurus* (L.)) či netopýra hnědavého (*Myotis lucifugus* (Le Conte)). Mimo savce byl nalezen u lososa obecného (*Salmo salar* (L.)) a pstruha duhového (*Oncorhynchus mykiss* (Walbaum)) (**Oliver et al., 2009**). V nedávné analýze z roku 2017 bylo nalezeno 227 orthologů proteinu PRDM9. V tomto výzkumu byl ortholog PRDM9 nalezen u bezčelistnatců (*Agnatha*), paryb, ryb, želv, hadů, ještěřů a latimérií. Naopak chybí u obojživelníků, ptáků a také u některých plazů (například krokodýlů) (**Oliver et al., 2009; Singhal et al., 2015; Baker et al., 2017**).

Stavba PRDM9 a funkce jednotlivých domén

Protein PRDM9 se skládá ze 3 hlavních funkčních domén a několika dalších menších částí. Hlavními funkčními doménami jsou N-terminální KRAB doména, centrální PR/SET doména a C-terminální doména C2H2 typu zinc finger (ZnF). Mezi KRAB a PR/SET doménou se nachází SSXRD motiv a před doménou C2H2 je umístěn 1 oddělený ZnF (Obr. 3) (Parvanov et al., 2010; Baudat et al., 2013). Většina domén genu *Prdm9* je vysoce konzervována napříč savci. Výjimkou je ZnF doména, která je naopak vysoce polymorfni (Thomas et al., 2009)

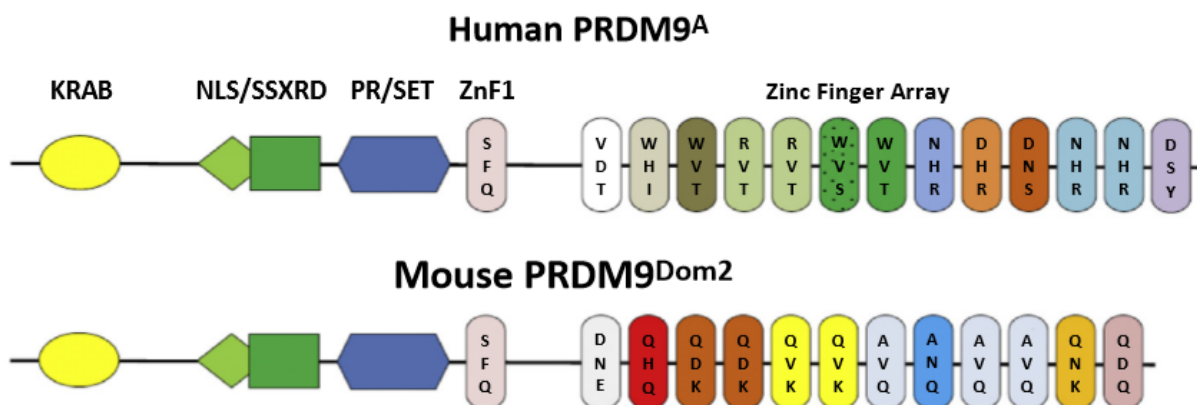
Hlavní funkční domény a další součásti PRDM9 vykazují různou aktivitu. KRAB doména zajišťuje především připojení rekombinačního hot-spotu, který se nachází na chromatinové smyčce, k chromozómové ose (tzv. axis či core) v místech, kde se nachází proteiny zodpovědné za tvorbu DSBs. Toto propojení se děje pomocí interakce s CXXC1 proteinem (což je savčí homolog kvasinkového proteinu Spp1) (Parvanov et al., 2016; Imai et al., 2017).

Centrální PR/SET doména vykazuje methyltransferázovou aktivitu. V místě nasednutí PRDM9 tato doména trimethyluje histon H3 na lysinu 4 (Hayashi et al., 2005) a stejný histon trimethyluje i na lysinu 36. Vytvoří tak dvojitou značku, která se vyskytuje pouze v hotspotech a umí jí vytvořit jen protein PRDM9 (Powers et al., 2016). PRDM9 trimethyluje 1 – 2 nukleozómy na obou stranách od místa nasednutí proteinu (zda 1 nebo 2 záleží pravděpodobně na míře stočení chromatinu, tedy na vzdálenosti nukleozómů od PRDM9 v prostoru). Značka H3K4me3 určuje rozpětí migrace Hollidayova spoje při rekombinaci. Tím určuje i délku vyměněné genetické informace. Takový úsek může být dlouhý 1 000 – 2 500 bp (Baker et al., 2014). Naopak délka genové konverze (spojené s non-crossoverem (NCO)) v hot-spotu je výrazně kratší (přibližně 10 – 120 bp, maximální délka úseku může být i větší, ale i přesto bývá menší než délka genové konverze spojené s CO). Tento poznatek se využívá v analýzách při rozlišování CO a NCO, kdy úsek genové konverze spojené s CO je delší než úsek genové konverze spojené s NCO (Cole et al., 2010).

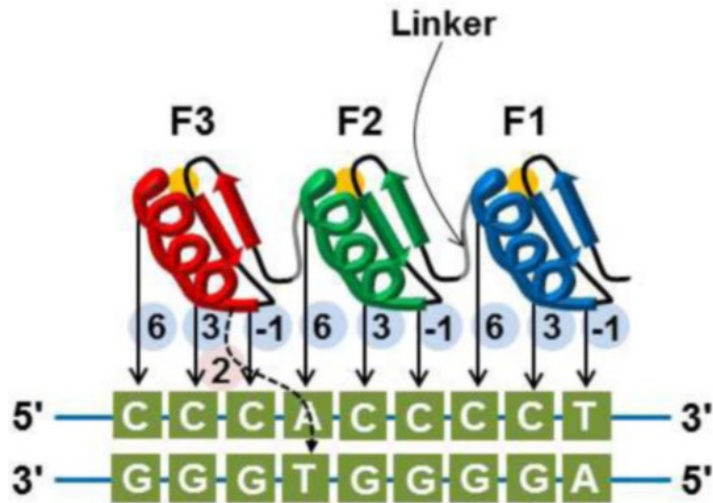
Poslední z hlavních funkčních domén je C2H2 doména, která obsahuje tandemové repetice, kdy 1 repetice je 1 ZnF (Parvanov et al., 2010). Těchto ZnFs může být v různých populacích různý počet. U lidí bylo zatím objeveno 8 – 18 ZnFs (Berg et al., 2010). Některé aminokyseliny jsou často substituovány a to především ty v pozici -1, 3 a 6 (Obr. 4) (Ponting, 2011). U člověka se pouze 5 – 6 ZnFs významně podílí na vazbě hledaného motivu DNA a to konkrétně ZnFs 8 – 12 u alely A (která má celkem 13 ZnFs) a ZnFs 8 – 13 u alely C (tato

alela má dohromady 14 ZnFs) (Baudat et al., 2010; Berg et al., 2011; Patel et al., 2017). Zbylé ZnFs mají jiné funkce, například při stabilizaci vazby PRDM9 s DNA (Billings et al., 2013). Doména C2H2 je vysoce polymorfní, liší se i mezi jedinci 1 populace a způsobuje tak různou distribuci hot-spotů (Thibault-Sennett et al., 2018). Rozpoznává cílový DNA motiv a určuje místo následné rekombinace (Grey et al., 2011).

Role SSXRD motivu a osamoceného ZnF zatím není zcela jasná. Jednou z možností je, že interagují s dalšími proteiny (Baudat et al., 2013). V nedávných pokusech bylo zjištěno, že při úplném odstranění SSXRD ztrácí PRDM9 svou methyltransferázovou aktivitu. Když vědci odstranili z tohoto proteinu pouze 24 ze 32 aminokyselin tvořících SSXRD, methylace mohla proběhnout. Zdá se tedy, že alespoň 8 aminokyselin tohoto motivu je potřebných pro správnou funkci PRDM9 (Thibault-Sennett et al., 2018). Role osamoceného ZnF by mohla být autoregulační. V případě této autoregulační teorie by docházelo k blokaci substrátu pomocí tohoto ZnF a reguloval by tím methyltransferázovou aktivitu (Wu et al., 2013).



Obrázek 3: Struktura PRDM9 u člověka a myši. V horní polovině je schematický obrázek nejčastějšího lidského PRDM9, značeného jako alela A (která váže motiv CCN CCN TNN CCN C (Myers et al., 2008)). Ve spodní polovině je myší (*Mus musculus domesticus*) ortholog PRDM9. Opět je zde nejčastější alela, která je u myši označena jako alela Dom2 (ta váže motiv GN(T/A) GNT (G/A)CT (G/A)(C/T)(C/T) (Sun et al., 2015)). U obou je při popisu zleva KRAB doména, následuje motiv SSXRD (někdy také značený jako NLS – nukleolární lokalizační signál), za ním doména s methyltransferázovou aktivitou tzv. PR/SET doména, poté osamocené zinkové prsty a poslední je zinková prst doména (jinak také C2H2 doména), která může obsahovat různý počet zinkových prstů. Právě tímto počtem se myší a lidský PRDM9 liší (zde u lidí 13 ZnFs a u myši 12 ZnFs). Písmena v každém ze ZnF značí aminokyseliny (Paigen a Petkov, 2018).



Obrázek 4: Schématická stavba domény C2H2 a hlavní vazebné pozice. Na obrázku jsou vyznačeny 3 ZnFs (červeně, zeleně a modře). Každý ze zinkových prstů se skládá ze 2 β -skládaných listů (šipky) a 1 α -helixu. Pozice -1, 3 a 6 tohoto α -helixu se přímo váží na vlákno DNA a jsou velmi důležité pro vazebnou specifitu proteinu PRDM9. Pozice 2 v α -helixu se váže na druhé vlákno DNA, což naznačuje důležitost nejen 3 hlavních pozic, ale i dalších aminokyselin v ZnF. Šedé vlákno označené jako „Linker“ je spojovací část mezi jednotlivými ZnFs (Hossain et al., 2015).

Evoluce PRDM9

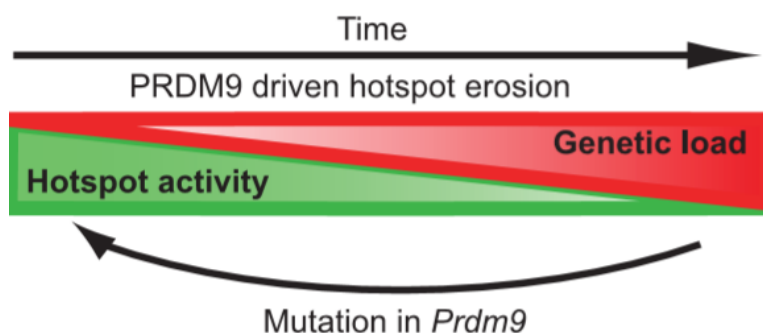
Evoluce hot-spotů a *Prdm9* vytváří obrovskou diverzitu alel ZnF domény PRDM9. Touto diverzitou vzniká rozsáhlá diverzita variant alel a jejich vazebných cílů (Baudat et al., 2013). U genu *Prdm9* často dochází k delecím a duplikacím. Výsledkem jsou již zmiňované různé počty ZnFs mezi druhy. Blízce příbuzné druhy se díky evoluci liší jak počtem ZnFs a substitucemi v nich, tak i pozicí (Oliver et al., 2009; Myers et al., 2010).

Evoluce PRDM9 i DNA motivu, na který se tento protein váže, je velice rychlá a vysvětluje rozdíly v tomto proteinu a v umístění hot-spotů mezi druhy. ZnF doména je repetitivní sekvence, kvůli čemuž by mohla být méně stabilní a v tom případě by evoluce probíhala rychleji (Myers et al., 2010). Velkou roli v procesu evoluce hraje pozitivní selekce a to zejména na pozice -1, 3, 6 v α -helixu (tedy v každém ZnF) (Obr. 4) (Oliver et al., 2009). Právě tyto tři pozice se váží na DNA (Choo a Klug, 1997) a nejvíce se podílí na sekvenční specifitě (Choo a Klug, 1994).

Naprostou zásadním jevem u hot-spotů je tzv. „hot-spot paradox“. Tento jev je v podstatě řízenou inaktivací hot-spotů, a to pomocí procesu genové konverze. Sekvence kolem vzniklého DSB jsou nahrazeny sekvencemi neporušeného homologa, což vede ke ztrátě alel náchylných ke zlomům. Naopak jsou tyto alely nahrazeny alelami, které vytvoření DSBs

nepodporují (**Boulton et al., 1997; Baudat et al., 2013**). Právě místo, na které se váže PRDM9, je vždy blízko DSBs a velmi často prochází procesem genové konverze. V důsledku toho dochází k odstranění vazebného místa pro PRDM9. Jak je tedy možné, že hot-spoty jsou v genomu stále přítomné? Vysvětlením „hot-spot paradoxu“ je rapidní evoluce alel PRDM9 a vznik nových hot-spotů dříve, než začne míra rekombinace znatelně klesat (Obr. 5) (**Baudat et al., 2013**). Hot-spoty jsou uchráněny před zmizením vznikem nového rekombinačního hot-spotu, a tedy nového cílového motivu, což je zapříčiněno změnou aminokyseliny v ZnF doméně PRDM9 (**Oliver et al., 2009; Baudat et al., 2010; Myers et al., 2010; Ponting, 2011**). Hot-spot paradox je tak v podstatě řešen mutacemi a genovým driftem (**Coop a Myers, 2007**). Míra rekombinace je udržována pozitivní selekcí, která cílí na zvýšení frekvence alel výhodných pro fitness jedince (a tedy i na udržení dostatečné míry rekombinace). Výhodné mohou být v tomto případě nové alely *Prdm9* genu, z nichž se exprimují nové formy PRDM9 proteinu, které se váží na motiv, který je u nevýhodných alel, které jsou ve vazbě. Nová alela určí nové místo rekombinace a rozruší tím vazbu nevýhodných alel a tedy napomůže zvýšení fitness jedince (**Ponting, 2011**). Obecně je platné, že silnější hot-spoty jsou mnohem více náchylné k extinkci vlivem genové konverze, než slabší hot-spoty (**Coop a Myers, 2007**).

Dosud zde byla zmiňována silná pozitivní selekce. V nedávném výzkumu se ale autoři zaměřili i na působení slabé selekce. Ta nefixuje silně nějaký specifický haplotyp a naopak udržuje všechny haplotypy v podobném podílu, popřípadě udržuje podíly každého z haplotypů cyklicky v čase. V prvním případě (udržování haplotypů ve stejném podílu v populaci) vykazují jednotlivé hot-spoty velmi mírnou aktivitu a téměř nezanikají. Ve druhém případě (udržení podílu určitého haplotypu cyklicky v čase) vykazují hot-spoty nízkou a vysokou aktivitu ve střídavých cyklech, takže mizí a obnovují se v čase (**Úbeda et al., 2019**). Zatímco podle starších modelů vysvětlujících „hot-spot paradox“ silnou selekcí se předchází extinkci hot-spotů vznikem nových rekombinačních míst (**Baudat et al., 2010; Myers et al., 2010**), v této novější teorii slabá selekce nepředchází zániku jednotlivých hot-spotů. Místo toho je chrání před vymizením jejich řízeným vznikem v homozygotních cílových místech, kde je vliv genové konverze mnohem menší. Tato homozygotní cílová místa jsou přitom poblíž původního cílového místa (**Úbeda et al., 2019**).



Obrázek 5: Cyklus zániku a obnovy hot-spotů. Nahoře je zleva doprava označen čas. Zelené pole značí aktivitu hot-spotu. Červené pole označuje takzvanou genetickou zátěž, tedy postupnou inaktivaci hot-spotů procesem genové konverze kvůli vazbě PRDM9. Na tomto schématu jsou vyznačeny 2 procesy, které působí proti sobě. Prvním (zde značeným jako „PRDM9 driven hotspot erosion“ a působícím od aktivního hot-spotu k inaktivaci) je eroze hot-spotů pomocí genové konverze a jejich extinkce. Druhým procesem, který řeší hot-spot paradox, je mutace v *Prdm9* (zde značeno jako „Mutation in *Prdm9*“). Mutace probíhá především v ZnF doméně a vede k posunu umístění hot-spotů na jiné místo (Baker et al., 2015).

3.2.1.2 Člověk a další primáti

Pro lidské hot-spoty platí vše, co bylo uvedeno v sekci 3.2.1.1. Jsou určeny, jako u většiny ostatních savců, proteinem PRDM9 (Baudat et al., 2010; Myers et al., 2010; Parvanov et al., 2010). Alely *Prdm9* genu lidí byly prozkoumány detailněji a zatím jich bylo objeveno přes 40. Různé etnické skupiny (Afroameričané, Chanové, mexičtí Američané, Evropané) mají různé alely s odlišným počtem ZnFs (Baudat et al., 2010; Berg et al., 2010, 2011; Parvanov et al., 2010). Z africké populace nese asi 50 % lidí alelu A, 13 % alelu C a další lidé mají mnoho méně zastoupených alel PRDM9 (Berg et al., 2010, 2011; Paigen a Petkov, 2018). Naopak jiné zkoumané populace a etnické skupiny (převážně Američané a Evropané) nesou z 90 % alely A a B (alelu A ve větší míře). Tyto 2 alely se přitom liší pouze o 1 aminokyselinu v té samé pozici (alela A – serin, alela B – threonin), což neovlivňuje vazebnou specifitu. Tato aminokyselina se totiž nachází na 6. ZnF a neúčastní se tak přímo vazby (viz výše) (Baudat et al., 2010; Paigen a Petkov, 2018).

S alelami souvisí i motivy, které se nejčastěji nachází v hot-spotech. První objevené motivy byly 2. První z nich je dlouhý 7 bp a jeho sekvence je CCTCCCT. Druhý z motivů je dlouhý 9 bp a má sekvenci CCCACCCC. Oba tyto motivy se vyskytují přibližně v 10 % všech lidských hot-spotů (Myers et al., 2005). Dosud nejvýznamnějším objeveným motivem je 13mer, který rozpoznávají formy proteinu PRDM9 kódované právě alelami A a B. Je přítomný ve 41 % hot-spotů (a to i na chromozomu X, takže je společný pro obě pohlaví). Sekvence tohoto motivu je následující CCNCCNTNNCCNC (kde N značí jakýkoliv

nukleotid z možností A, T, C, G). První část tohoto motivu je zároveň velmi podobná prvnímu ze zmíněných motivů CCTCCCT (Myers et al., 2008). Alela C, kterou nesou především Afričané a Afroameričané, má afinitu k 13mernímu motivu velmi nízkou. Váže se na odlišný motiv dlouhý 17 bp, jehož sekvence je CCCCaGTGAGCGTtgCc (malá písmena značí, že na dané pozici může být i jiná báze, ale nejčastěji je tam právě ta zapsaná) (Hinch et al., 2011).

Pokud srovnáme motivy hot-spotů člověka a blízkce příbuzného druhu šimpanze, jsou zcela odlišné a jsou i naprosto jinak rozmístěné. Genom člověka a šimpanze vykazuje cca 99% sekvenční homologii. I přes tak vysokou sekvenční shodu však nesdílí téměř žádné společné hot-spoty (Ptak et al., 2005; Winckler et al., 2005). Pro hot-spoty šimpanzů jsou typické 2 motivy CGCG a CCCGGC (Auton et al., 2012). Mezi různými druhy primátů existuje i značná variabilita v počtu ZnF PRDM9 proteinu. Příkladem jsou šimpanzi s 15 ZnFs, orangutani s 10 ZnFs, kosmani s 9 ZnFs a lidé se 13 ZnFs (pokud tedy zmiňujeme alelu A) (Oliver et al., 2009).

Počet hot-spotů v genomu člověka byl stanoven na více než 25 000, ovšem doopravdy by jich mohlo být až 50 000. Jsou od sebe vzdálené 50 – 100 kb (Winckler et al., 2005; Arnheim et al., 2007). Kromě centromer nebylo v lidském genomu nalezeno místo delší než 200 kb, které by bylo bez hot-spotu. Obecně se lidské hot-spoty nachází blízko genů, avšak mnohem méně často přímo v genech (Myers et al., 2005). Z hlediska míry rekombinace nejvíce CO probíhá v subtelomerických oblastech (Myers et al., 2005). Stejně tak je tomu i u šimpanzů (Auton et al., 2012).

S rekombinačními hot-spoty a chybami v nich je u člověka spojeno i několik onemocnění. Zlomy v DNA, které jsou opravovány prostřednictvím NAHR, se často nachází právě v rekombinačních hot-spotech. Při takové opravě často dochází k deleci nebo duplikaci a s tím jsou pak spojené problémy. Mezi takto způsobené nemoci patří Charcot-Marie-Tooth syndrom (delece na chromozomu 17, vrozená neuropatie), Hunterův syndrom, X-vázaná ichtyóza, Potocki-Lupski/Smith-Magenis syndrom, hemofilie A, azoospermie a další onemocnění (Pratto et al., 2014). Spojitost se našla i s Downovým syndromem. Matky, které nesly minoritní alely *Prdm9* genu (tedy jiné než je alela A), měly potomky s Downovým syndromem častěji než matky nesoucí majoritní alelu A. Důvodem je zřejmě menší četnost vazebných míst pro PRDM9 na chromozomu 21 (Oliver et al., 2016). U skupiny japonských pacientů s azoospermií byla objevena 3 místa v *Prdm9* genu s SNP (single-nucleotid polymorphism, tedy jednonukleotidový polymorfismus). U těchto pacientů se našel vztah SNPs v PRDM9 a azoospermie. Mechanismus zastavení meiose zatím nebyl popsán. Počet

zkoumaných pacientů byl ale malý a složený pouze z Japonců. V tomto případě by bylo vhodné analyzovat i další etnické skupiny s větším počtem jedinců (**Miyamoto et al., 2008**). Naprosto odlišným a zajímavým případem je pákistánská žena, která měla homozygotní nulovou mutaci v genu *Prdm9* (tedy mutaci způsobující naprostou ztrátu funkce tohoto genu). I přesto byla fertilní a měla potomka. Gen *Prdm9* této ženy měl defektní doménu Set, takže v jejím genomu nebyl typický H3K4me3 vzor. Pouze 5,9 % COs se překrývalo s místy, kde se hot-spoty běžně vyskytují u lidí s alelou A (**Narasimhan et al., 2016**).

Uvažuje se i možné korelaci proteinu PRDM9 a hot-spotů s různými typy rakoviny. Tento vztah zatím ale není úplně jasný (**Paigen a Petkov, 2018**).

3.2.1.3 Psovití

Čeď *Canidae* je dosud jedinou objevenou čeledí savců, která ztratila funkční gen *Prdm9*. (**Oliver et al., 2009; Muñoz-Fuentes et al., 2011**). Celá tato skupina živočichů má proto hot-spoty podobné svým umístěním spíše těm ancestrálním. U této čeledi došlo k 5 funkčně-ztrátovým mutacím (typu *nonsense* a *missense*), k posunovým mutacím a k vytvoření předčasného stop kodonu (**Oliver et al., 2009; Axelsson et al., 2012**). Ztráta funkce genu *Prdm9* se odehrála ještě před domestikací, protože stejné mutace byly nalezeny i u vlků, od nichž se psi odštěpili. Takto nekompletní a nefunkční gen mají téměř všichni psovití, což naznačuje, že se tyto mutace musely vytvořit už před diverzifikací této čeledi živočichů. K inaktivaci genu *Prdm9* časově došlo nejspíše až po oddělení od pandy (přibližně před 49 miliony let) a zároveň před oddělením, vlků, lišek, šakalů a psů od jejich společného předka (přibližně před 7,8 miliony let) (**Eizirik et al., 2010; Axelsson et al., 2012**).

Ačkoliv *Canidae* nemají funkční gen *Prdm9*, jejich míra rekombinace je velmi podobná jako u lidí. V rekombinačních hot-spotech mají o 40 – 50 % vyšší obsah GC párů než se průměrně vyskytuje v celém genomu (**Axelsson et al., 2012**), zatímco např. lidské hot-spoty mají pouze o 1 – 2 % více GC (**Spencer et al., 2006**). Příčinou korelace vyššího obsahu GC a hot-spotů u psových je nejspíše gBGC (**Duret a Arndt, 2008**). Je ale možné, že vysoký obsah GC není zapříčiněn pouze gBGC a další faktory, které mají vliv na rekombinaci u psů, ještě nebyly objeveny (**Axelsson et al., 2012**). Jakákoliv spojitost výskytu hot-spotů s vyšším obsahem H3K4me3 nebyla u psových nalezena. Celkově se hot-spoty jeví jako velmi stabilní v umístění. V porovnání s lidmi a myši jsou GC-bohaté oblasti, kde se hot-spoty psových často vyskytují, mnohem méně v regionech kolem promotorů (jako je tomu například u kvasinek) a poblíž velmi důležitých a vysoce exprimovaných genů (**Han a Zhao,**

2009). Důvodem by mohla být snaha o snížení rekombinace v důležitých genech, aby nedocházelo k jejich případné inaktivaci (Auton et al., 2013).

3.3 Druhy bez hot-spotů

3.3.1 Rod *Drosophila*

V mnoha studiích bylo potvrzeno, že rod *Drosophila* nemá hot-spoty (Cirulli et al., 2006; Stevison and Noor, 2010; Comeron et al., 2012; Howie et al., 2019). Části genomu se mezi sebou liší v míře rekombinace, ale velmi mírně. U všech zkoumaných druhů rodu *Drosophila* existuje asociace četností rekombinace s repetitivním motivem, který lze souhrnně zapsat jako (CA)_n motiv (kde n je počet opakování) (Howie et al., 2019).

U octomilky obecné (*Drosophila melanogaster* (Meigen)) se krátké repetitivní motivy vyskytují nejčastěji v euchromatinu. Typické pro druh octomilky *D. melanogaster* jsou 2 motivy. První z nich je CACAC a druhý je CCTCCCT. Míra rekombinace je velmi malá blízko centromer a telomer. Žádné COs nebyly nalezeny v malém achiasmatickém chromozomu. Genová konverze u tohoto druhu octomilky probíhá, ale oprava je náhodná (není přednostně na GC, jako tomu bylo u mnoha výše zmíněných druhů živočichů) (Comeron et al., 2012; Howie et al., 2019).

Dalším detailněji zkoumaným druhem je *Drosophila pseudoobscura* (Frolova & Astaurov). U tohoto druhu bylo zjištěno, že se 80 % rekombinací stane v 50 % sekvencí. Nižší míra rekombinace byla objevena v TSS (Smukowski Heil et al., 2015). I tento druh má sekvenční motivy spojené s vyšší mírou rekombinace. Jsou to 3 motivy: CACAC (Cirulli et al., 2006), CCCCACCCC a CCTCCCT (Kulathinal et al., 2008).

Posledním z více zkoumaných druhů je *Drosophila persimilis* (Dobzhansky & Epling). U tohoto druhu pravděpodobně funguje genová konverze, která přednostně opravuje DNA na GC. Positivní korelace byla ve studii objevena s motivem CCNCCNTNNCCNC (Stevison a Noor, 2010). V novějších studiích ale tento motiv nebyl nalezen. Je možné, že roli v určování pozice rekombinace mají i další, doposud neobjevené, mechanismy (Howie et al., 2019).

3.3.2 *Caenorhabditis elegans*

Hád'átka obecné (*Caenorhabditis elegans* (Maupas)) je další druh, u kterého dosud nebyly nalezeny hot-spoty.

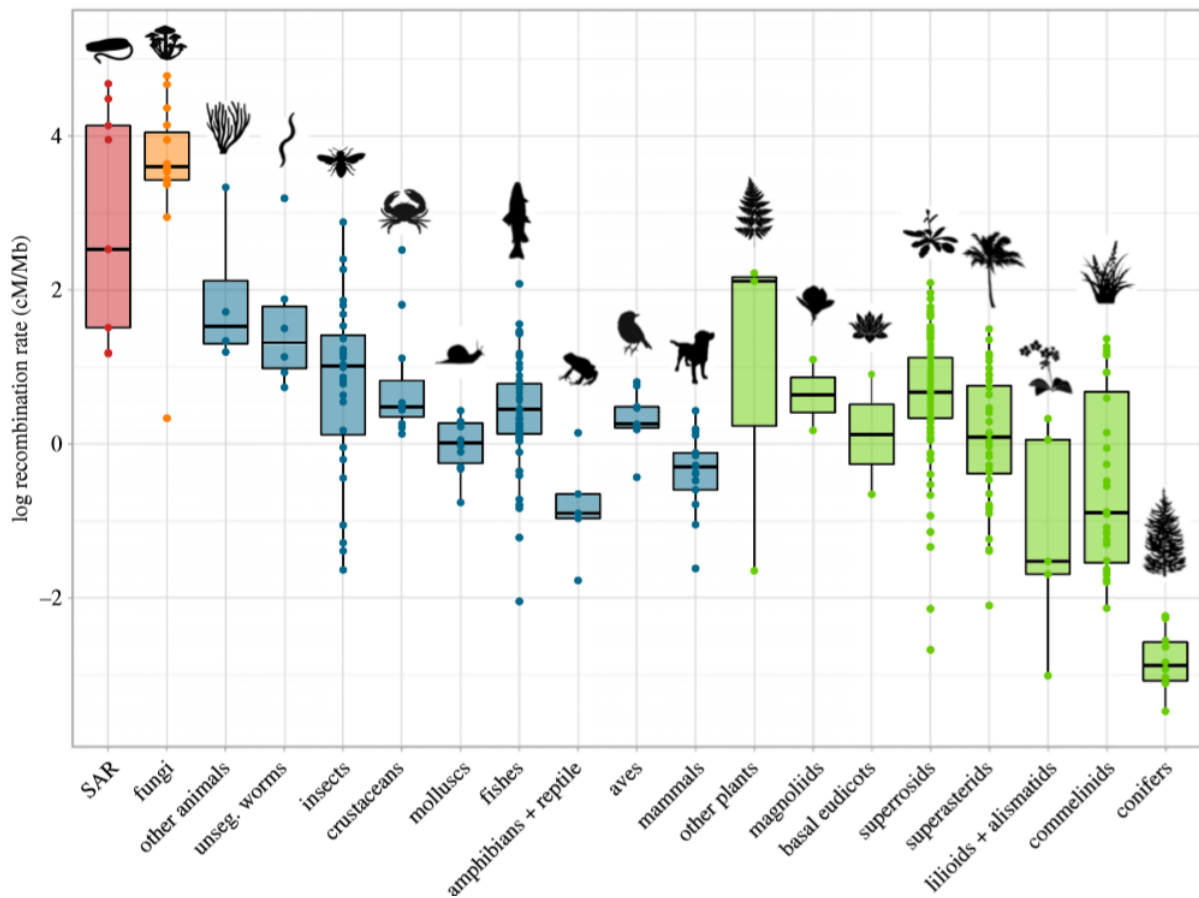
Studií zabývajících se absencí hot-spotů a determinanty míst rekombinací u hád'átka je však velice málo. Obecně lze říct, že v centrálním regionu chromozómů je míra rekombinace nižší a zároveň je tam vysoká genová denzita i vyšší míra exprese genů. Naopak raménka chromozómů vykazují vyšší míru rekombinace a nachází se na nich větší počet repetitivních sekvencí. Na raménkách chromozómů jsou geny s nižší mírou exprese a celkově je tam menší genová denzita. Míra rekombinace se ve jednotlivých oblastech liší, ovšem velmi mírně (jako je tomu u *Drosophila*). Nebyla objevena žádná korelace s histonovými modifikacemi. Zatím jediná výraznější spojitost s rekombinačními hot-spoty jsou nejspíše sekvenční motivy. Žádný konkrétní nebyl popsán, ale v místech s vyšší mírou rekombinace se nachází větší množství AT-bohatých sekvencí (Kaur a Rockman, 2014). Efekt genové konverze je zde velice slabý, což je pro samooplodňovací organismy časté, protože jsou ve zvýšené míře homozygotní (Marais et al., 2004; Kaur and Rockman, 2014). V chromozómech hád'átka působí velmi silná CO interference, která umožňuje vznik pouze 1 CO (Kaur a Rockman, 2014).

3.4 Porovnání míry rekombinace skupin a druhů organismů

V této části práce se zaměřím na srovnání míry rekombinace různých skupin a druhů organismů. Pro vyjádření síly rekombinačních hot-spotů i míst bez hot-spotů se používá takzvaná míra rekombinace (v angličtině značená jako RR – *recombination rate*). Tato veličina vyjadřuje frekvenci rekombinace na jednotku fyzické vzdálenosti za 1 generaci a její jednotkou je tedy cM/Mb (cM = centimorgan, Mb = megabáze).

Nejprve se zaměřím na porovnání skupin v obecnější rovině (Obr. 6). Houby a endoparazitické mikroorganismy mají celkovou míru rekombinace mnohem vyšší než živočichové a rostliny. Číselně lze vyjádřit průměrnou míru rekombinace následovně: SAR (*Stramenopila*, *Alveolata*, *Rhizaria*) 38,67 cM/Mb, houby 48,68 cM/Mb, živočichové 2,52 cM/Mb a rostliny 1,85 cM/Mb (Stapley et al., 2017). Míra rekombinace u člověka je 1,13 cM/Mb. Shrnující tabulka s příklady organismů a jejich mírou rekombinace je níže (Tab. 1). Z hmyzu mají nejvyšší míru rekombinace sociální *Hymenoptera*, naopak nejnižší *Diptera* (Wilfert et al., 2007). Pokud se přesuneme k obratlovcům, u nich nejvyšší míra rekombinace náleží rybám (Stapley et al., 2017). Mezi savci naopak skupinu s nejnižší mírou

rekombinace tvoří vačnatci (Dumont a Payseur, 2008). Ze všech zkoumaných rostlin mají výrazně nejnižší míru rekombinace jehličnany (Jaramillo-Correa et al., 2010).



Obrázek 6: Porovnání míry rekombinace mezi jednotlivými skupinami organismů. Osa y – logaritmus z hodnoty míry rekombinace (cM/Mb); osa x – jednotlivé skupiny organismů. Osa x zleva: SAR = *Stramenopila*, *Alveolata*, *Rhizaria*; houby, další nezařazení živočichové (korálnatci, sumýši, sumky); ploštěnci a hlístice; hmyz; koryši; měkkýši; ryby; obojživelníci a plazi; ptáci; savci; nezařazené rostliny (kaprad'orosty, mechorosty a zelené řasy); *Magnoliidae*; *Magnolipsida* (nižší dvouděložné rostliny); skupina „Superrosids“ (obsahuje 18 řádů včetně *Rosales*, *Fabales*, *Geraniales*, *Brassicales* a *Cucurbitales*); skupina „Superasterids“ (obsahuje 20 řádů včetně *Santales*, *Aquafoliales*, *Asterales*, *Solanales* a *Lamiales*); skupiny „Lilioids“ a „Alismatids“ (obsahují dohromady 7 řádů včetně *Acorales*, *Alismatales*, *Lilliales* a *Asparagales*); skupina „Commenilids“ (obsahuje 4 řády: *Arecales*, *Zingiberales*, *Commelinales* a *Poales*); jehličnany (Stapley et al., 2017)

Tabulka 1: Porovnání míry rekombinace mezi jednotlivými druhy. Tabulka byla vytvořena na základě souhrmných článků Wilfert et al. 2007, Dumont a Payseur 2008 a Jaramillo-Correa et al. 2010. Odkazy na primární zdroje, ze kterých pocházejí uvedená data, lze najít ve zmíněných 3 článcích. Míru rekombinace pro některé druhy jsem spočítala z informací uvedených ve článcích (Croll et al. 2015) pro *Z. tritici*, (Barton et al. 2008) pro *S. cerevisiae* a (Rockman a Kruglyak 2009) pro *C. elegans* a (Brion et al., 2017) pro *S. pombe*. V některých případech však raději uvádím míru rekombinace v původních jednotkách uvedených v příslušném článku.

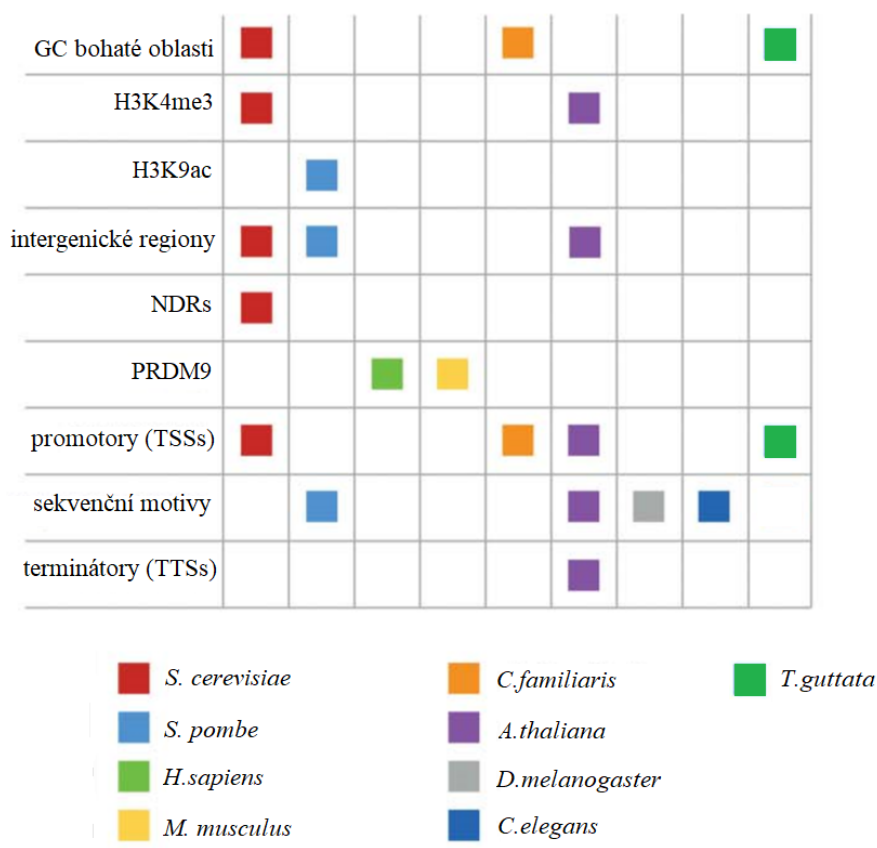
Druh	Třída	Řád	Čeleď	Míra rekombinace (cM/Mb)
<i>Asparagus officinalis</i> (L.)	Liliopsida	Asparagales	Asparagaceae	0,55
<i>Cocos nucifera</i> (L.)	Liliopsida	Arecales	Areaceae	1,14
<i>Oryza sativa</i> (L.)	Liliopsida	Poales	Poaceae	3,22
<i>Zea mays</i> (L.)	Liliopsida	Poales	Poaceae	0,67
<i>Daucus carota</i> (L.)	Rosopsida	Apiales	Apiaceae	0,57
<i>Helianthus petiolaris</i> (Nutt)	Rosopsida	Asterales	Asteraceae	0,35
<i>Arabidopsis thaliana</i> (L.) Heynh.)	Rosopsida	Brassicales	Brassicaceae	4,21
<i>Brassica oleracea</i> (L.)	Rosopsida	Brassicales	Brassicaceae	2,11
<i>Carica papaya</i> (L.)	Rosopsida	Brassicales	Caricaceae	1,53
<i>Spinacia oleracea</i> (L.)	Rosopsida	Caryophyllales	Amaranthaceae	0,59
<i>Cucurbita pepo</i> (L.)	Rosopsida	Cucurbitales	Cucurbitaceae	3,6
<i>Camellia sinensis</i> (L.) Kuntze)	Rosopsida	Ericales	Theaceae	0,35
<i>Acacia mangium</i> (Willd)	Rosopsida	Fabales	Fabaceae	1,02
<i>Vicia faba</i> (L.)	Rosopsida	Fabales	Fabaceae	0,12
<i>Betula pendula</i> (Roth)	Rosopsida	Fagales	Betulaceae	1,01
<i>Castanea sativa</i> (Mill.)	Rosopsida	Fagales	Fagaceae	2,09
<i>Coffea liberica</i> (W. Bull ex Hiern.)	Rosopsida	Gentianales	Rubiaceae	2,11
<i>Olea europaea</i> (L.)	Rosopsida	Lamiales	Oleaceae	0,06
<i>Manihot esculenta</i> (Crantz)	Rosopsida	Malpighiales	Euphorbiaceae	0,76
<i>Passiflora edulis f. flavicarpa</i> (O. Deg.)	Rosopsida	Malpighiales	Passifloraceae	2,09
<i>Populus nigra</i> (L.)	Rosopsida	Malpighiales	Salicaceae	3,76

Druh	Třída	Řád	Čeleď	Míra rekombinace (cM/Mb)
<i>Theobroma cacao</i> (L.)	Rosopsida	Malvales	Malvaceae	2,62
<i>Eucalyptus globulus</i> (Labill.)	Rosopsida	Myrtales	Myrtaceae	1,29
<i>Macadamia integrifolia</i> (Maiden & Betche)	Rosopsida	Proteales	Proteaceae	2,37
<i>Humulus lupulus</i> (L.)	Rosopsida	Rosales	Cannabidaceae	0,25
<i>Prunus armeniaca</i> (L.)	Rosopsida	Rosales	Rosaceae	1,12
<i>Rosa wichuriana</i> (Crépin)	Rosopsida	Rosales	Rosaceae	0,78
<i>Rubus idaeus</i> (L.)	Rosopsida	Rosales	Rosaceae	1,53
<i>Citrus sinensis</i> (L.) Osbeck)	Rosopsida	Sapindales	Rutaceae	5,24
<i>Solanum lycopersicum</i> (L.)	Rosopsida	Solanales	Solanaceae	1,55
<i>Vitis vinifera</i> (L.)	Rosopsida	Vitales	Vitaceae	0,82
<i>Cryptomeria japonica</i> (Thunb. ex L. f.) D. Don)	Pinopsida	Pinales	Cupressaceae	1,09
<i>Larix kaempferi</i> ((Lamb.) Carrière)	Pinopsida	Pinales	Pinaceae	1,92
<i>Picea abies</i> (L.) H. Karst.)	Pinopsida	Pinales	Pinaceae	0,07
<i>Picea glauca</i> ((Moench) Voss)	Pinopsida	Pinales	Pinaceae	0,09
<i>Pinus pinaster</i> (Aiton)	Pinopsida	Pinales	Pinaceae	0,32
<i>Trypanosoma brucei</i> (Plimmer & Bradford)	Kinetoplastea	Trypanosomatida	Trypanosomidae	52,2
<i>Plasmodium falciparum</i> (Welch)	Aconoidasida	Haemosporida	Plasmodiidae	67,7
<i>Cochliobolus sativus</i> (S. Ito & Kurbitt.) Drechsler ex Dastur)	Dothideomycetes	Pleosporales	Pleosporaceae	25,7
<i>Zymoseptoria tritici</i> (Desm.) Quaedvl. & Crous)	Dothideomycetes	Capnodiales	Mycosphaerellaceae	71,16
<i>Saccharomyces cerevisiae</i> (Meyen ex E. C. Hansen)	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	0,4 cM/kb
<i>Schizosaccharomyces pombe</i> (Lindner)	Schizosaccharomycetes	Schizosaccharomycetales	Schizosaccharomycetaceae	3 COs/Mb
<i>Coprinospin cinerea</i> ((Schaeff.) Redhead, Vilgalys & Moncalvo)	Agaricomycetes	Agaricales	Psathyrellaceae	37,4
<i>Cryptococcus neoformans</i> ((San Felice) Vuill.)	Tremellomycetes	Tremellales	Tremellaceae	75
<i>Caenorhabditis elegans</i> (Maupas)	Secernentea	Rhabditida	Rhabditidae	12,59

Druh	Třída	Řád	Čeleď	Mira rekombinace (cM/Mb)
<i>Crassostrea gigas</i> (Thunberg)	Bivalvia	Ostreida	Flemingostreidae	1,2
<i>Daphnia pulex</i> (Leydig)	Branchiopoda	Diplostroca	Daphniidae	5,1
<i>Rhyzopertha dominica</i> (Fabricius)	Insecta	Coleoptera	Bostrichidae	0,8
<i>Tribolium confusum</i> (Jacquelin du Val)	Insecta	Coleoptera	Tenebrionidae	4
<i>Aedes aegypti</i> (L.)	Insecta	Diptera	Culicidae	0,3
<i>Anopheles gambiae</i> (Giles)	Insecta	Diptera	Culicidae	0,8
<i>Drosophila melanogaster</i> (Meigen)	Insecta	Diptera	Drosophilidae	1,6
<i>Apis mellifera</i> (L.)	Insecta	Hymenoptera	Apidae	16
<i>Bombus terrestris</i> (L.)	Insecta	Hymenoptera	Apidae	4,4
<i>Nasonia vitripennis</i> (Walker)	Insecta	Hymenoptera	Pteromalidae	2,5
<i>Trichogramma brassicae</i> (Bezdenko)	Insecta	Hymenoptera	Trichogrammatidae	5,4
<i>Acromyrmex echinator</i> (Forel)	Insecta	Hymenoptera	Formicidae	6,7
<i>Bombyx mori</i> (L.)	Insecta	Lepidoptera	Bombycidae	2,6
<i>Takifugu rubripes</i> (Temminck & Schlegel)	Actinopterygii	Tetraodontiformes	Tetraodontidae	3,1
<i>Ambystoma tigrinum</i> (Green)	Amphibia	Ambystomatoidea	Ambystomatidae	0,2
<i>Gallus gallus</i> (L.)	Aves	Galliformes	Phasianidae	3,1
<i>Monodelphis domestica</i> (Wagner)	Mammalia	Didelphimorphia	Didelphidae	0,18
<i>Macropus eugenii</i> (Demsarest)	Mammalia	Diprotodontia	Macropodidae	0,22
<i>Mus musculus</i> (L.)	Mammalia	Rodentia	Muridae	0,4
<i>Equus caballus</i> (L.)	Mammalia	Perissodactyla	Equidae	0,9
<i>Canis familiaris</i> (L.)	Mammalia	Carnivora	Canidae	1,55
<i>Felis catus</i> (L.)	Mammalia	Carnivora	Felidae	1,1
<i>Macaca mulatta</i> (Zimmermann)	Mammalia	Primates	Cercopitheidae	0,73
<i>Papio hamadryas</i> (L.)	Mammalia	Primates	Cercopitheidae	0,65
<i>Homo sapiens</i> (L.)	Mammalia	Primates	Hominidae	1,13

4. Závěr

Tato práce shrnuje základní poznatky o rekombinačních hot-spotech napříč různými organismy. Obsahuje také výčet některých faktorů, které mají na umístění rekombinace, a tedy i na umístění rekombinačních hot-spotů, významnější vliv. Rekombinační hot-spoty jsou u různých organismů určeny odlišnými mechanismy. Mezi tyto mechanismy patří především histonové modifikace, sekvenční motivy, transkripční počátky, konce transkripce a protein PRDM9 (Obr. 7). Hot-spoty jsou zde rozdělené na ancestrální a odvozené. K ancestrálním hot-spotům patří hot-spoty hub (především kvasinek), ptáků a rostlin. K odvozeným hot-spotům jsou přiřazeni savci (s proteinem PRDM9).



Obrázek 7: Porovnání mechanismu určování hot-spotů mezi jednotlivými druhy. Na obrázku jsou uvedeny nejčastější mechanismy určení rekombinačních hot-spotů z hlediska jejich lokalizace. Je vždy uveden primární mechanismus určení u daného druhu (tedy u člověka a myši je uveden pouze protein PRDM9, který vytváří histonové modifikace a NDRs, ačkoliv ani jedno v tabulce zaškrtnuté není). Jsou zde druhy ze skupiny kvasinek (*S. cerevisiae* a *S. pombe*), ze třídy savců (*H. sapiens*, *M. musculus* a *C. familiaris*), z říše rostlin (*A. thaliana*), dále 2 druhy zastupující skupinu organismů bez hot-spotů (*D. melanogaster* a *C. elegans*) a druh z třídy ptáků (*T. guttata*). Jednotlivé barvy přiřazené ke druhům označují v tabulce určující mechanismy hot-spotů daného druhu (Cooper et al., 2016; Tiemann-Boege et al., 2017).

Proteinu PRDM9, který se nachází především u savců, se práce věnuje detailněji. PRDM9 se skládá ze 3 hlavních funkčních domén a každá z nich má jinou funkci. KRAB doména zajišťuje připojení hot-spotu (který je na chromozomové smyčce) k chromozomové ose. PR/SET doména zajišťuje tvorbu epigenetických značek (H3K4me3 a H3K36me3) a poslední C2H2 doména obsahující ZnFs určuje místo rekombinace. C2H2 doména prochází rychlou evolucí, aby vyřešila tzv. „hot-spot paradox“.

PRDM9 se ale nevyskytuje ve funkční formě u všech savců. U psovitých kvůli mutacím došlo k vytvoření pseudogenu, který vytváří nefunkční formu PRDM9. Psovití tak vykazují podobné určení hot-spotů jako druhy s ancestrálními hot-spoty.

Pro člověka existuje několik typických sekvenčních motivů, na které se PRDM9 váže a ten nejčastější z nich (CCNCCNTNNCCNC) se vyskytuje ve 40 % hot-spotů. Problémy a chyby při procesu rekombinace mohou způsobit několik vážnějších nemocí jako je Charcot-Marie-Tooth syndrom, hemofilie A nebo azoospermie.

Některé druhy vůbec hot-spoty nemají. Blíže prozkoumanými druhy jsou *C. elegans* a některé druhy z rodu *Drosophila*. U obou jsou hot-spoty určeny nejspíše sekvenčními motivy.

Organismy se od sebe výrazně liší mírou rekombinace. Nejvyšší míru rekombinace vykazují houby a parazitické mikroorganismy. Svojí hodnotou míry rekombinace několikanásobně převyšují druhy z jiných skupin organismů.

Rekombinační hot-spoty jsou stále aktuálním tématem a bylo by potřeba je prozkoumat detailněji. Ačkoliv je v současné době dostupné velké množství informací o rekombinačních hot-spotech, jsou tyto informace často neúplné. U mnoha druhů (včetně zde zmíněného *C. elegans*), řádů i tříd (třídy z podkmene obratlovců mimo savců) je buď velmi málo studií nebo dokonce žádné. K detailnějšímu poznání hot-spotů by tak bylo vhodné prozkoumat více druhů, aby byly informace dostatečně reprezentativní a abychom získali lepší představu o určení hot-spotů napříč různými druhy. Odlišný způsob určení a umístění hot-spotů je důležité více zkoumat hlavně s ohledem na to, že i údaje, které jsou k dispozici dnes, se hodně liší mezi různými skupinami druhů, a to dokonce i mezi relativně blízce příbuznými druhy. Možný budoucí směr výzkumu v této oblasti, který by mohl mít i praktické uplatnění, je například rekombinace a hot-spoty u člověka. V několika studiích je zmíněno, že chyby v rekombinaci by mohly mít vliv na rakovinu a plodnost, což jsou velmi aktuální problémy.

Seznam použité literatury

- Arnheim, N., Calabrese, P., and Tiemann-Boege, I. (2007). Mammalian Meiotic Recombination Hot Spots. *Annu. Rev. Genet.* *41*, 369–399.
- Auton, A., Fledel-Alon, A., Pfeifer, S., Venn, O., Séguirel, L., Street, T., Leffler, E.M., Bowden, R., Aneas, I., Broxholme, J., et al. (2012). A Fine-Scale Chimpanzee Genetic Map from Population Sequencing. *Science* *336*, 193–198.
- Auton, A., Li, Y.R., Kidd, J., Oliveira, K., Nadel, J., Holloway, J.K., Hayward, J.J., Cohen, P.E., Greally, J.M., Wang, J., et al. (2013). Genetic Recombination Is Targeted Towards Gene Promoter Regions in Dogs. *PLoS Genet.* *9*, e1003984.
- Axelsson, E., Webster, M.T., Ratnakumar, A., The LUPA Consortium, Ponting, C.P., and Lindblad-Toh, K. (2012). Death of PRDM9 Coincides with Stabilization of the Recombination Landscape in the Dog Genome. *Genome Res.* *22*, 51–63.
- Baker, C.L., Walker, M., Kajita, S., Petkov, P.M., and Paigen, K. (2014). PRDM9 Binding Organizes Hotspot Nucleosomes and Limits Holliday Junction Migration. *Genome Res.* *24*, 724–732.
- Baker, C.L., Kajita, S., Walker, M., Saxl, R.L., Raghupathy, N., Choi, K., Petkov, P.M., and Paigen, K. (2015). PRDM9 Drives Evolutionary Erosion of Hotspots in *Mus musculus* through Haplotype-Specific Initiation of Meiotic Recombination. *PLoS Genet.* *11*, e1004916.
- Baker, Z., Schumer, M., Haba, Y., Bashkirova, L., Holland, C., Rosenthal, G.G., and Przeworski, M. (2017). Repeated Losses of PRDM9-Directed Recombination Despite the Conservation of PRDM9 across Vertebrates. *ELife* *6*, e24133.
- Barton, A.B., Pkosz, M.R., Kurvathi, R.S., and Kaback, D.B. (2008). Meiotic Recombination at the Ends of Chromosomes in *Saccharomyces cerevisiae*. *Genetics* *179*, 1221–1235.
- Baudat, F., Buard, J., Grey, C., Fledel-Alon, A., Ober, C., Przeworski, M., Coop, G., and de Massy, B. (2010). PRDM9 Is a Major Determinant of Meiotic Recombination Hotspots in Humans and Mice. *Science* *327*, 836–840.
- Baudat, F., Imai, Y., and de Massy, B. (2013). Meiotic Recombination in Mammals: Localization and Regulation. *Nat. Rev. Genet.* *14*, 794–806.
- Berg, I.L., Neumann, R., Lam, K.-W.G., Sarbajna, S., Odenthal-Hesse, L., May, C.A., and Jeffreys, A.J. (2010). PRDM9 Variation Strongly Influences Recombination Hot-Spot Activity and Meiotic Instability in Humans. *Nat. Genet.* *42*, 859–863.
- Berg, I.L., Neumann, R., Sarbajna, S., Odenthal-Hesse, L., Butler, N.J., and Jeffreys, A.J. (2011). Variants of the Protein PRDM9 Differentially Regulate a Set of Human Meiotic Recombination Hotspots Highly Active in African Populations. *Proc. Natl. Acad. Sci.* *108*, 12378–12383.
- Beye, M., Gattermeier, I., Hasselmann, M., Gempe, T., Schioett, M., Baines, J.F., Schlipalius, D., Mougél, F., Emore, C., Rueppell, O., et al. (2006). Exceptionally High Levels of Recombination across the Honey Bee Genome. *Genome Res.* *16*, 1339–1344.
- Billings, T., Parvanov, E.D., Baker, C.L., Walker, M., Paigen, K., and Petkov, P.M. (2013). DNA Binding Specificities of the Long Zinc-Finger Recombination Protein PRDM9. *Genome Biol.* *14*, R35.
- Blitzblau, H.G., Bell, G.W., Rodriguez, J., Bell, S.P., and Hochwagen, A. (2007). Mapping of Meiotic Single-Stranded DNA Reveals Double-Strand-Break Hotspots Near Centromeres and Telomeres. *Curr. Biol.* *17*, 2003–2012.

- Borde, V., Robine, N., Lin, W., Bonfils, S., Géli, V., and Nicolas, A. (2009). Histone H3 Lysine 4 Trimethylation Marks Meiotic Recombination Initiation Sites. *EMBO J.* 28, 99–111.
- Börner, G.V., Kleckner, N., and Hunter, N. (2004). Crossover/Noncrossover Differentiation, Synaptonemal Complex Formation, and Regulatory Surveillance at the Leptotene/Zygotene Transition of Meiosis. *Cell* 117, 29–45.
- Boulton, A., Myers, R.S., and Redfield, R.J. (1997). The Hotspot Conversion Paradox and the Evolution of Meiotic Recombination. *Proc. Natl. Acad. Sci.* 94, 8058–8063.
- Brick, K., Smagulova, F., Khil, P., Camerini-Otero, R.D., and Petukhova, G.V. (2012). Genetic Recombination Is Directed away from Functional Genomic Elements in Mice. *Nature* 485, 642–645.
- Brion, C., Legrand, S., Peter, J., Caradec, C., Pflieger, D., Hou, J., Friedrich, A., Llorente, B., and Schacherer, J. (2017). Variation of the Meiotic Recombination Landscape and Properties over a Broad Evolutionary Distance in Yeasts. *PLoS Genet.* 13, e1006917.
- Brown, T.C., and Jiricny, J. (1988). Different Base/Base Mispairs Are Corrected with Different Efficiencies and Specificities in Monkey Kidney Cells. *Cell* 54, 705–711.
- Campbell, C.L., Furlotte, N.A., Eriksson, N., Hinds, D., and Auton, A. (2015). Escape From Crossover Interference Increases with Maternal Age. *Nat. Commun.* 6, 1–7.
- de Castro, E., Soriano, I., Marín, L., Serrano, R., Quintales, L., and Antequera, F. (2012). Nucleosomal Organization of Replication Origins and Meiotic Recombination Hotspots in Fission Yeast. *EMBO J.* 31, 124–137.
- Choi, K., and Henderson, I.R. (2015). Meiotic Recombination Hotspots - A Comparative View. *Plant J.* 83, 52–61.
- Choi, K., Zhao, X., Kelly, K.A., Venn, O., Higgins, J.D., Yelina, N.E., Hardcastle, T.J., Ziolkowski, P.A., Copenhaver, G.P., Franklin, F.C.H., et al. (2013). *Arabidopsis* Meiotic Crossover Hot Spots Overlap with H2A.Z Nucleosomes at Gene Promoters. *Nat. Genet.* 45, 1327–1336.
- Choo, Y., and Klug, A. (1994). Toward a Code for the Interactions of Zinc Fingers with DNA: Selection of Randomized Fingers Displayed on Phage. *Proc. Natl. Acad. Sci.* 91, 11163–11167.
- Choo, Y., and Klug, A. (1997). Physical Basis of a Protein-DNA Recognition Code. *Curr. Opin. Struct. Biol.* 7, 117–125.
- Cirulli, E.T., Kliman, R.M., and Noor, M.A.F. (2006). Fine-Scale Crossover Rate Heterogeneity in *Drosophila pseudoobscura*. *J. Mol. Evol.* 64, 129–135.
- Cokus, S.J., Feng, S., Zhang, X., Chen, Z., Merriman, B., Haudenschild, C.D., Pradhan, S., Nelson, S.F., Pellegrini, M., and Jacobsen, S.E. (2008). Shotgun Bisulphite Sequencing of the *Arabidopsis* Genome Reveals DNA Methylation Patterning. *Nature* 452, 215–219.
- Cole, F., Keeney, S., and Jasin, M. (2010). Comprehensive, Fine-Scale Dissection of Homologous Recombination Outcomes at a Hotspot in Mouse Meiosis. *Mol. Cell* 39, 700–710.
- Colegrave, N. (2002). Sex Releases the Speed Limit on Evolution. *Nature* 420, 664–666.
- Comeron, J.M., Ratnappan, R., and Bailin, S. (2012). The Many Landscapes of Recombination in *Drosophila melanogaster*. *PLoS Genet.* 8, e1002905.
- Coop, G., and Myers, S.R. (2007). Live Hot, Die Young: Transmission Distortion in Recombination Hotspots. *PLoS Genet.* 3, e35.

- Cooper, T.J., Garcia, V., and Neale, M.J. (2016). Meiotic DSB Patterning: A Multifaceted Process. *Cell Cycle* 15, 13–21.
- Croll, D., Lendenmann, M.H., Stewart, E., and McDonald, B.A. (2015). The Impact of Recombination Hotspots on Genome Evolution of a Fungal Plant Pathogen. *Genetics* 201, 1213–1228.
- Cromie, G.A., Hyppa, R.W., Cam, H.P., Farah, J.A., Grewal, S.I.S., and Smith, G.R. (2007). A Discrete Class of Intergenic DNA Dictates Meiotic DNA Break Hotspots in Fission Yeast. *PLoS Genet.* 3, e141.
- Dreissig, S., Mascher, M., and Heckmann, S. (2019). Variation in Recombination Rate Is Shaped by Domestication and Environmental Conditions in Barley. *Mol. Biol. Evol.* 36, 2029–2039.
- Dressler, D., and Potter, H. (1982). Molecular Mechanisms in Genetic Recombination. *Annu. Rev. Biochem.* 51, 727–761.
- Drouaud, J., Khademian, H., Giraut, L., Zanni, V., Bellalou, S., Henderson, I.R., Falque, M., and Mézard, C. (2013). Contrasted Patterns of Crossover and Non-crossover at *Arabidopsis thaliana* Meiotic Recombination Hotspots. *PLoS Genet.* 9, e1003922.
- Dumont, B.L., and Payseur, B.A. (2008). Evolution of the Genomic Rate of Recombination in Mammals. *Evolution* 62, 276–294.
- Dumont, B.L., and Payseur, B.A. (2011). Genetic Analysis of Genome-Scale Recombination Rate Evolution in House Mice. *PLoS Genet.* 7, e1002116.
- Duret, L., and Arndt, P.F. (2008). The Impact of Recombination on Nucleotide Substitutions in the Human Genome. *PLoS Genet.* 4, e1000071.
- Eizirik, E., Murphy, W.J., Koepfli, K.-P., Johnson, W.E., Dragoo, J.W., Wayne, R.K., and O'Brien, S.J. (2010). Pattern and Timing of Diversification of the Mammalian Order Carnivora Inferred from Multiple Nuclear Gene Sequences. *Mol. Phylogenet. Evol.* 56, 49–63.
- Field, Y., Kaplan, N., Fondufe-Mittendorf, Y., Moore, I.K., Sharon, E., Lubling, Y., Widom, J., and Segal, E. (2008). Distinct Modes of Regulation by Chromatin Encoded Through Nucleosome Positioning Signals. *PLoS Comput. Biol.* 4, e1000216.
- Fowler, K.R., Sasaki, M., Milman, N., Keeney, S., and Smith, G.R. (2014). Evolutionarily Diverse Determinants of Meiotic DNA Break and Recombination Landscapes Across the Genome. *Genome Res.* 24, 1650–1664.
- Fox, M.E., Yamada, T., Ohta, K., and Smith, G.R. (2000). A Family of cAMP-Response-Element-Related DNA Sequences With Meiotic Recombination Hotspot Activity in *Schizosaccharomyces pombe*. *Genetics* 156, 59–68.
- Garcia, V., Gray, S., Allison, R.M., Cooper, T.J., and Neale, M.J. (2015). Tel1/ATM-Mediated Interference Suppresses Clustered Meiotic Double-Strand-Break Formation. *Nature* 520, 114–118.
- Gerton, J.L., DeRisi, J., Shroff, R., Lichten, M., Brown, P.O., and Petes, T.D. (2000). Global Mapping of Meiotic Recombination Hotspots and Coldspots in the Yeast *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci.* 97, 11383–11390.
- Grey, C., Barthès, P., Chauveau-Le Fricq, G., Langa, F., Baudat, F., and de Massy, B. (2011). Mouse PRDM9 DNA-Binding Specificity Determines Sites of Histone H3 Lysine 4 Trimethylation for Initiation of Meiotic Recombination. *PLoS Biol.* 9, e1001176.
- Groenen, M.A.M., Wahlberg, P., Foglio, M., Cheng, H.H., Megens, H.-J., Crooijmans, R.P.M.A., Besnier, F., Lathrop, M., Muir, W.M., Wong, G.K.-S., et al. (2008). A High-Density

- SNP-Based Linkage Map of the Chicken Genome Reveals Sequence Features Correlated with Recombination Rate. *Genome Res.* *19*, 510–519.
- Han, L., and Zhao, Z. (2009). Contrast Features of CpG Islands in the Promoter and Other Regions in the Dog Genome. *Genomics* *94*, 117–124.
- Hayashi, K., Yoshida, K., and Matsui, Y. (2005). A Histone H3 Methyltransferase Controls Epigenetic Events Required for Meiotic Prophase. *Nature* *438*, 374–378.
- Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Van Calcar, S., Qu, C., Ching, K.A., et al. (2007). Distinct and Predictive Chromatin Signatures of Transcriptional Promoters and Enhancers in the Human Genome. *Nat. Genet.* *39*, 311–318.
- Higgins, J.D. (2004). The *Arabidopsis* MutS Homolog AtMSH4 Functions at an Early Step in Recombination: Evidence for Two Classes of Recombination in *Arabidopsis*. *Genes Dev.* *18*, 2557–2570.
- Hinch, A.G., Tandon, A., Patterson, N., Song, Y., Rohland, N., Palmer, C.D., Chen, G.K., Wang, K., Buxbaum, S.G., Akylbekova, E.L., et al. (2011). The Landscape of Recombination in African Americans. *Nature* *476*, 170–175.
- Hossain, M.A., Barrow, J.J., Shen, Y., Haq, M.I., and Bungert, J. (2015). Artificial Zinc Finger DNA Binding Domains: Versatile Tools for Genome Engineering and Modulation of Gene Expression. *J. Cell. Biochem.* *116*, 2435–2444.
- Howie, J.M., Mazzucco, R., Taus, T., and Nolte, V. (2019). DNA Motifs Are Not General Predictors of Recombination in Two *Drosophila* Sister Species. *Genome Biol Evol* *11*, 1345–1357.
- Hyppa, R.W., and Smith, G.R. (2010). Crossover Invariance Determined by Partner Choice for Meiotic DNA Break Repair. *Cell* *142*, 243–255.
- Imai, Y., Baudat, F., Taillepierre, M., Stanzione, M., Toth, A., and de Massy, B. (2017). The PRDM9 KRAB Domain Is Required for Meiosis and Involved in Protein Interactions. *Chromosoma* *126*, 681–695.
- Jaramillo-Correa, J.P., Verdu, M., and Gonzalez-Martinez, S.C. (2010). The Contribution of Recombination to Heterozygosity Differs Among Plant Evolutionary Lineages and Life-Forms. *BMC Evol. Biol.* *10*, 1–10.
- Jones, G.H., and Franklin, F.C.H. (2006). Meiotic Crossing-Over: Obligation and Interference. *Cell* *126*, 246–248.
- Kaplan, N., Moore, I.K., Fondufe-Mittendorf, Y., Gossett, A.J., Tillo, D., Field, Y., LeProust, E.M., Hughes, T.R., Lieb, J.D., Widom, J., et al. (2009). The DNA-Encoded Nucleosome Organization of a Eukaryotic Genome. *Nature* *458*, 362–366.
- Kauppi, L., Jeffreys, A.J., and Keeney, S. (2004). Where the Crossovers Are: Recombination Distributions in Mammals. *Nat. Rev. Genet.* *5*, 413–424.
- Kaur, T., and Rockman, M.V. (2014). Crossover Heterogeneity in the Absence of Hotspots in *Caenorhabditis elegans*. *Genetics* *196*, 137–148.
- Kerstes, N.A.G., Béréños, C., Schmid-Hempel, P., and Wegner, K.M. (2012). Antagonistic Experimental Coevolution with a Parasite Increases Host Recombination Frequency. *BMC Evol. Biol.* *12*, 1–9.
- Kirkpatrick, D.T., Fan, Q., and Petes, T.D. (1999). Maximal Stimulation of Meiotic Recombination by a Yeast Transcription Factor Requires the Transcription Activation Domain and a DNA-Binding Domain. *Genetics* *152*, 101–115.

- Kon, N., Krawchuk, M.D., Warren, B.G., Smith, G.R., and Wahls, W.P. (1997). Transcription Factor Mts1/Mts2 (Atf1/Pcr1, Gad7/Pcr1) Activates the *M26* Meiotic Recombination Hotspot in *Schizosaccharomyces pombe*. *Proc. Natl. Acad. Sci. U. S. A.* *94*, 13765–13770.
- Kong, A., Barnard, J., Gudbjartsson, D.F., Thorleifsson, G., Jonsdottir, G., Sigurdardottir, S., Richardsson, B., Jonsdottir, J., Thorgeirsson, T., Frigge, M.L., et al. (2004). Recombination Rate and Reproductive Success in Humans. *Nat. Genet.* *36*, 1203–1206.
- Kovalchuk, I., Kovalchuk, O., Kalck, V., Boyko, V., Filkowski, J., Heinlein, M., and Hohn, B. (2003). Pathogen-Induced Systemic Plant Signal Triggers DNA Rearrangements. *Nature* *423*, 760–762.
- Kulathinal, R.J., Bennett, S.M., Fitzpatrick, C.L., and Noor, M.A.F. (2008). Fine-Scale Mapping of Recombination Rate in *Drosophila* Refines Its Correlation to Diversity and Divergence. *Proc. Natl. Acad. Sci. U. S. A.* *105*, 10051–10056.
- Lam, I., and Keeney, S. (2015). Nonparadoxical Evolutionary Stability of the Recombination Initiation Landscape in Yeast. *Science* *350*, 932–937.
- Le Comber, S.C., Ainouche, M.L., Kovarik, A., and Leitch, A.R. (2010). Making a Functional Diploid: From Polysomic to Disomic Inheritance. *New Phytol.* *186*, 113–122.
- Leflon, M., Grandont, L., Eber, F., Huteau, V., Coriton, O., Chelysheva, L., Jenczewski, E., and Chèvre, A.-M. (2010). Crossovers Get a Boost in *Brassica* Allotriploid and Allotetraploid Hybrids. *Plant Cell* *22*, 2253–2264.
- Lenormand, T., Engelstädter, J., Johnston, S.E., Wijnker, E., and Haag, C.R. (2016). Evolutionary Mysteries in Meiosis. *Philos. Trans. R. Soc. B Biol. Sci.* *371*, 20160001.
- Li, B., Carey, M., and Workman, J.L. (2007). The Role of Chromatin during Transcription. *Cell* *128*, 707–719.
- Lichten, M., and Goldman, A.S.H. (1995). Meiotic Recombination Hotspots. *Annu. Rev. Genet.* *29*, 423–444.
- Liu, S., Yeh, C.-T., Ji, T., Ying, K., Wu, H., Tang, H.M., Fu, Y., Nettleton, D., and Schnable, P.S. (2009). Mu Transposon Insertion Sites and Meiotic Recombination Events Co-Localize with Epigenetic Marks for Open Chromatin Across the Maize Genome. *PLoS Genet.* *5*, e1000733.
- Marais, G., Charlesworth, B., and Wright, S.I. (2004). Recombination and Base Composition: The Case of the Highly Self-Fertilizing Plant *Arabidopsis thaliana*. *Genome Biol.* *5*, R45.
- Marand, A.P., Zhao, H., Zhang, W., Zeng, Z., Fang, C., and Jiang, J. (2019). Historical Meiotic Crossover Hotspots Fueled Patterns of Evolutionary Divergence in Rice. *Plant Cell* *31*, 645–662.
- de Massy, B. (2013). Initiation of Meiotic Recombination: How and Where? Conservation and Specificities Among Eukaryotes. *Annu. Rev. Genet.* *47*, 563–599.
- Mihola, O., Trachtulec, Z., Vlcek, C., Schimenti, J.C., and Forejt, J. (2009). A Mouse Speciation Gene Encodes a Meiotic Histone H3 Methyltransferase. *Science* *323*, 373–375.
- Miyamoto, T., Koh, E., Sakugawa, N., Sato, H., Hayashi, H., Namiki, M., and Sengoku, K. (2008). Two Single Nucleotide Polymorphisms in *PRDM9* (*MEISETZ*) Gene May Be a Genetic Risk Factor for Japanese Patients with Azoospermia by Meiotic Arrest. *J. Assist. Reprod. Genet.* *25*, 553–557.
- Morran, L.T., Parmenter, M.D., and Phillips, P.C. (2009). Mutation Load and Rapid Adaptation Favor Outcrossing Over Self-Fertilization. *Nature* *462*, 350–352.

- Muñoz-Fuentes, V., Di Rienzo, A., and Vilà, C. (2011). Prdm9, a Major Determinant of Meiotic Recombination Hotspots, Is Not Functional in Dogs and Their Wild Relatives, Wolves and Coyotes. *PLoS ONE* 6, e25498.
- Muñoz-Fuentes, V., Marcet-Ortega, M., Alkorta-Aranburu, G., Linde Forsberg, C., Morrell, J.M., Manzano-Piedras, E., Söderberg, A., Daniel, K., Villalba, A., Toth, A., et al. (2015). Strong Artificial Selection in Domestic Mammals Did Not Result in an Increased Recombination Rate. *Mol. Biol. Evol.* 32, 510–523.
- Myers, S., Bottolo, L., Freeman, C., McVean, G., and Donnelly, P. (2005). A Fine-Scale Map of Recombination Rates and Hotspots Across the Human Genome. *Science* 310, 321–324.
- Myers, S., Freeman, C., Auton, A., Donnelly, P., and McVean, G. (2008). A Common Sequence Motif Associated with Recombination Hot Spots and Genome Instability in Humans. *Nat. Genet.* 40, 1124–1129.
- Myers, S., Bowden, R., Tumian, A., Bontrop, R.E., Freeman, C., MacFie, T.S., McVean, G., and Donnelly, P. (2010). Drive Against Hotspot Motifs in Primates Implicates the PRDM9 Gene in Meiotic Recombination. *Science* 327, 876–879.
- Narasimhan, V.M., Hunt, K.A., Mason, D., Baker, C.L., Karczewski, K.J., Barnes, M.R., Barnett, A.H., Bates, C., Bellary, S., Bockett, N.A., et al. (2016). Health and Population Effects of Rare Gene Knockouts in Adult Humans with Related Parents. *Science* 352, 474–477.
- Oliver, P.L., Goodstadt, L., Bayes, J.J., Birtle, Z., Roach, K.C., Phadnis, N., Beatson, S.A., Lunter, G., Malik, H.S., and Ponting, C.P. (2009). Accelerated Evolution of the *Prdm9* Speciation Gene Across Diverse Metazoan Taxa. *PLoS Genet.* 5, e1000753.
- Oliver, T.R., Middlebrooks, C., Harden, A., Scott, N., Johnson, B., Jones, J., Walker, C., Wilkerson, C., Saffold, S.-H., Akinseye, A., et al. (2016). Variation in the Zinc Finger of PRDM9 is Associated with the Absence of Recombination Along Nondisjoined Chromosomes 21 of Maternal Origin. *J. Syndr. Chromosome Abnorm.* 2, 1–11.
- Otto, S.P., and Lenormand, T. (2002). Resolving the Paradox of Sex and Recombination. *Nat. Rev. Genet.* 3, 252–261.
- Paigen, K., and Petkov, P.M. (2018). PRDM9 and Its Role in Genetic Recombination. *Trends Genet.* 34, 291–300.
- Pan, J., Sasaki, M., Kniewel, R., Murakami, H., Blitzblau, H.G., Tischfield, S.E., Zhu, X., Neale, M.J., Jasin, M., Succi, N.D., et al. (2011). A Hierarchical Combination of Factors Shapes the Genome-Wide Topography of Yeast Meiotic Recombination Initiation. *Cell* 144, 719–731.
- Pao, W.K., and Li, H.W. (1948). Desynapsis and Other Abnormalities Induced by High Temperature. *J. Genet.* 48, 297–310.
- Parvanov, E.D., Petkov, P.M., and Paigen, K. (2010). *Prdm9* Controls Activation of Mammalian Recombination Hotspots. *Science* 327, 835.
- Parvanov, E.D., Tian, H., Billings, T., Saxl, R.L., Spruce, C., Aithal, R., Krejci, L., Paigen, K., and Petkov, P.M. (2016). PRDM9 Interactions with Other Proteins Provide a Link Between Recombination Hotspots and the Chromosomal Axis in Meiosis. *Mol. Biol. Cell* 28, 488–499.
- Patel, A., Zhang, X., Blumenthal, R.M., and Cheng, X. (2017). Structural Basis of Human PR/SET Domain 9 (PRDM9) Allele C-Specific Recognition of Its Cognate DNA Sequence. *J. Biol. Chem.* 292, 15994–16002.
- Pekowska, A., Benoukraf, T., Zacarias-Cabeza, J., Belhocine, M., Koch, F., Holota, H., Imbert, J., Andrau, J.-C., Ferrier, P., and Spicuglia, S. (2011). H3K4 Tri-Methylation Provides an Epigenetic Signature of Active Enhancers. *EMBO J.* 30, 4198–4210.

- Petes, T.D. (2001). Meiotic Recombination Hot Spots and Cold Spots. *Nat. Rev. Genet.* 2, 360–369.
- Phadnis, N., Hyppa, R.W., and Smith, G.R. (2011). New and Old Ways to Control Meiotic Recombination. *Trends Genet.* 27, 411–421.
- Ponticelli, A.S., Sena, E.P., and Smith, G.R. (1988). Genetic and Physical Analysis of the *M26* Recombination Hotspot of *Schizosaccharomyces pombe*. *Genetics* 119, 491–497.
- Ponting, C.P. (2011). What Are the Genomic Drivers of the Rapid Evolution of PRDM9? *Trends Genet.* 27, 165–171.
- Powers, N.R., Parvanov, E.D., Baker, C.L., Walker, M., Petkov, P.M., and Paigen, K. (2016). The Meiotic Recombination Activator PRDM9 Trimethylates Both H3K36 and H3K4 at Recombination Hotspots In Vivo. *PLoS Genet.* 12, e1006146.
- Pratto, F., Brick, K., Khil, P., Smagulova, F., Petukhova, G.V., and Camerini-Otero, R.D. (2014). Recombination Initiation Maps of Individual Human Genomes. *Science* 346, 1256442.
- Ptak, S.E., Hinds, D.A., Koehler, K., Nickel, B., Patil, N., Ballinger, D.G., Przeworski, M., Frazer, K.A., and Pääbo, S. (2005). Fine-Scale Recombination Patterns Differ Between Chimpanzees and Humans. *Nat. Genet.* 37, 429–434.
- Radman-Livaja, M., and Rando, O.J. (2010). Nucleosome Positioning: How Is It Established, and Why Does It Matter? *Dev. Biol.* 339, 258–266.
- Rice, W.R. (2002). Experimental Tests of the Adaptive Significance of Sexual Recombination. *Nat. Rev. Genet.* 3, 241–251.
- Ritz, K.R., Noor, M.A.F., and Singh, N.D. (2017). Variation in Recombination Rate: Adaptive or Not? *Trends Genet.* 33, 364–374.
- Rockman, M.V., and Kruglyak, L. (2009). Recombinational Landscape and Population Genomics of *Caenorhabditis elegans*. *PLoS Genet.* 5, e1000419.
- Rockmill, B., Voelkel-Meiman, K., and Roeder, G.S. (2006). Centromere-Proximal Crossovers Are Associated With Precocious Separation of Sister Chromatids During Meiosis in *Saccharomyces cerevisiae*. *Genetics* 174, 1745–1754.
- San Filippo, J., Sung, P., and Klein, H. (2008). Mechanism of Eukaryotic Homologous Recombination. *Annu. Rev. Biochem.* 77, 229–257.
- de los Santos, T., Hunter, N., Lee, C., Larkin, B., Loidl, J., and Hollingsworth, N.M. (2003). The Mus81/Mms4 Endonuclease Acts Independently of Double-Holliday Junction Resolution to Promote a Distinct Subset of Crossovers During Meiosis in Budding Yeast. *Genetics* 164, 81–94.
- Sasaki, M., Lange, J., and Keeney, S. (2010). Genome Destabilization by Homologous Recombination in the Germline. *Nat. Rev. Mol. Cell Biol.* 11, 182–195.
- Shilo, S., Melamed-Bessudo, C., Dorone, Y., Barkai, N., and Levy, A.A. (2015). DNA Crossover Motifs Associated with Epigenetic Modifications Delineate Open Chromatin Regions in *Arabidopsis*. *Plant Cell* 27, 2427–2436.
- Singhal, S., Leffler, E.M., Sannareddy, K., Turner, I., Venn, O., Hooper, D.M., Strand, A., Li, Q., Raney, B., Balakrishnan, C.N., et al. (2015). Stable Recombination Hotspots in Birds. *Science* 350, 928–932.
- Sirviö, A., Johnston, J.S., Wenseleers, T., and Pamilo, P. (2011). A High Recombination Rate in Eusocial Hymenoptera: Evidence from the Common Wasp *Vespula vulgaris*. *BMC Genet.* 12, 1–7.

- Smagulova, F., Gregoretto, I.V., Brick, K., Khil, P., Camerini-Otero, R.D., and Petukhova, G.V. (2011). Genome-Wide Analysis Reveals Novel Molecular Features of Mouse Recombination Hotspots. *Nature* 472, 375–378.
- Smukowski Heil, C.S., Ellison, C., Dubin, M., and Noor, M.A.F. (2015). Recombining without Hotspots: A Comprehensive Evolutionary Portrait of Recombination in Two Closely Related Species of *Drosophila*. *Genome Biol. Evol.* 7, 2829–2842.
- Spencer, C.C.A., Deloukas, P., Hunt, S., Mullikin, J., Myers, S., Silverman, B., Donnelly, P., Bentley, D., and McVean, G. (2006). The Influence of Recombination on Human Genetic Diversity. *PLoS Genet.* 2, 1375–1385.
- Stapley, J., Feulner, P.G.D., Johnston, S.E., Santure, A.W., and Smadja, C.M. (2017). Variation in Recombination Frequency and Distribution Across Eukaryotes: Patterns and Processes. *Philos. Trans. R. Soc. B Biol. Sci.* 372, 20160455.
- Stevison, L.S., and Noor, M.A.F. (2010). Genetic and Evolutionary Correlates of Fine-Scale Recombination Rate Variation in *Drosophila persimilis*. *J. Mol. Evol.* 71, 332–345.
- Stukenbrock, E.H., and Dutheil, J.Y. (2018). Fine-Scale Recombination Maps of Fungal Plant Pathogens Reveal Dynamic Recombination Landscapes and Intragenic Hotspots. *Genetics* 208, 1209–1229.
- Sun, F., Fujiwara, Y., Reinholdt, L.G., Hu, J., Saxl, R.L., Baker, C.L., Petkov, P.M., Paigen, K., and Handel, M.A. (2015). Nuclear Localization of PRDM9 and Its Role in Meiotic Chromatin Modifications and Homologous Synapsis. *Chromosoma* 124, 397–415.
- Termolino, P., Cremona, G., Consiglio, M.F., and Conicella, C. (2016). Insights into Epigenetic Landscape of Recombination-Free Regions. *Chromosoma* 125, 301–308.
- Thibault-Sennett, S., Yu, Q., Smagulova, F., Cloutier, J., Brick, K., Camerini-Otero, R.D., and Petukhova, G.V. (2018). Interrogating the Functions of PRDM9 Domains in Meiosis. *Genetics* 209, 475–487.
- Thomas, J.H., Emerson, R.O., and Shendure, J. (2009). Extraordinary Molecular Evolution in the *PRDM9* Fertility Gene. *PLoS ONE* 4, e8505.
- Tiemann-Boege, I., Schwarz, T., Striedner, Y., and Heissl, A. (2017). The Consequences of Sequence Erosion in the Evolution of Recombination Hotspots. *Philos. Trans. R. Soc. B Biol. Sci.* 372, 20160462.
- Tiley, G.P., and Burleigh, J.G. (2015). The Relationship of Recombination Rate, Genome Structure, and Patterns of Molecular Evolution Across Angiosperms. *BMC Evol. Biol.* 15, 1–14.
- Tsai, I.J., Burt, A., and Koufopanou, V. (2010). Conservation of Recombination Hotspots in Yeast. *Proc. Natl. Acad. Sci.* 107, 7847–7852.
- Úbeda, F., Russell, T.W., and Jansen, V.A.A. (2019). PRDM9 and the Evolution of Recombination Hotspots. *Theor. Popul. Biol.* 126, 19–32.
- Villeneuve, A.M., and Hillers, K.J. (2001). Whence Meiosis? *Cell* 106, 647–650.
- Wahls, W.P., Siegel, E.R., and Davidson, M.K. (2008). Meiotic Recombination Hotspots of Fission Yeast Are Directed to Loci that Express Non-Coding RNA. *PLoS ONE* 3, e2887.
- Wang, Y., and Copenhaver, G.P. (2018). Meiotic Recombination: Mixing It Up in Plants. *Annu. Rev. Plant Biol.* 69, 577–609.

- White, M.A., Dominska, M., and Petes, T.D. (1993). Transcription Factors are Required for the Meiotic Recombination Hotspot at the *HIS4* Locus in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci.* *90*, 6621–6625.
- Wijnker, E., Velikkakam James, G., Ding, J., Becker, F., Klasen, J.R., Rawat, V., Rowan, B.A., de Jong, D.F., de Snoo, C.B., Zapata, L., et al. (2013). The Genomic Landscape of Meiotic Crossovers and Gene Conversions in *Arabidopsis thaliana*. *ELife* *2*, e01426.
- Wilfert, L., Gadau, J., and Schmid-Hempel, P. (2007). Variation in Genomic Recombination Rates Among Animal Taxa and the Case of Social Insects. *Heredity* *98*, 189–197.
- Winckler, W., Myers, S.R., Richter, D.J., Onofrio, R.C., McDonald, G.J., Bontrop, R.E., McVean, G.A.T., Gabriel, S.B., Reich, D., Donnelly, P., et al. (2005). Comparison of Fine-Scale Recombination Rates in Humans and Chimpanzees. *Science* *308*, 107–111.
- Wu, H., Mathioudakis, N., Diagouraga, B., Dong, A., Dombrovski, L., Baudat, F., Cusack, S., de Massy, B., and Kadlec, J. (2013). Molecular Basis for the Regulation of the H3K4 Methyltransferase Activity of PRDM9. *Cell Rep.* *5*, 13–20.
- Yamada, S., Ohta, K., and Yamada, T. (2013). Acetylated Histone H3K9 Is Associated with Meiotic Recombination Hotspots, and Plays a Role in Recombination Redundantly with Other Factors Including the H3K4 Methylase Set1 in Fission Yeast. *Nucleic Acids Res.* *41*, 3504–3517.
- Yamada, S., Okamura, M., Oda, A., Murakami, H., Ohta, K., and Yamada, T. (2017). Correlation of Meiotic DSB Formation and Transcription Initiation Around Fission Yeast Recombination Hotspots. *Genetics* *206*, 801–809.
- Yamada, T., Mizuno, K., Hirota, K., Kon, N., Wahls, W.P., Hartsuiker, E., Murofushi, H., Shibata, T., and Ohta, K. (2004). Roles of Histone Acetylation and Chromatin Remodeling Factor in a Meiotic Recombination Hotspot. *EMBO J.* *23*, 1792–1803.
- Yant, L., Hollister, J.D., Wright, K.M., Arnold, B.J., Higgins, J.D., Franklin, F.C.H., and Bomblies, K. (2013). Meiotic Adaptation to Genome Duplication in *Arabidopsis arenosa*. *Curr. Biol.* *23*, 2151–2156.
- Zalevsky, J., MacQueen, A.J., Duffy, J.B., Kempfues, K.J., and Villeneuve, A.M. (1999). Crossing Over During *Caenorhabditis elegans* Meiosis Requires a Conserved MutS-Based Pathway That Is Partially Dispensable in Budding Yeast. *Genetics* *153*, 1271–1283.
- Zhang, L., Kim, K.P., Kleckner, N.E., and Storlazzi, A. (2011). Meiotic Double-Strand Breaks Occur Once per Pair of (Sister) Chromatids and, via Mec1/ATR and Tel1/ATM, Once per Quartet of Chromatids. *Proc. Natl. Acad. Sci.* *108*, 20036–20041.