Local and sex-specific biases in crossover vs. noncrossover outcomes at meiotic recombination hotspots in mouse

Esther de Boer, Maria Jasin, and Scott Keeney

List of Supplemental Material

Supplemental text describing the use of recombinant inbred mouse strains for recombination hotspot identification.
Figure S1. Breeding scheme for generating recombinant-inbred strains.
** Table S1: Polymorphisms, oligonucleotides, and recombinants at the central hotspot.
** Table S2: Polymorphisms, oligonucleotides, and recombinants at the distal hotspot.
Table S3: Additional primers for amplification and sequencing of the central hotspot region.
Table S4: Additional primers for amplification and sequencing of the distal hotspot region.
Table S5: Compilation of recombination and DSB (SSDS read count) data for the hotspots shown in Figure 3B.

** Provided as separate Excel files.

Identification of hotspots using recombinant inbred (RI) mouse strains

To compare recombination in male and female meiosis, we set out to identify two hotspots, preferentially on the same chromosome: one hotspot in a chromosomal region where males showed a clear preference for crossover formation over females, and one hotspot in a chromosomal region where males and females showed similar levels of crossover formation. At the time these studies were initiated, direct maps of DSBs via SSDS mapping were not yet available, so we initially made use of a method previously described by Bois (Bois, 2007) in which RI mouse strains are used as a crossover library (Figure S1) for prospective hotspot identification.

Based on cytological analysis using MLH1 foci as a marker for crossover formation, male mice form crossovers preferentially in the centromere-distal subtelomeric regions of many chromosomes (Baker *et al.*, 1996; Froenicke *et al.*, 2002). In contrast, females show much less of a preference for crossing over in this region (de Boer *et al.*, 2006). Using these cytological data, we selected the following centromere-distal subtelomeric regions as having a male-specific crossover preference: 90–100% of chromosomes 1 and 2, 75–90% of chromosome 18 and 70–90% of chromosome 19 (with the centromere set at 0%). We further defined the following central regions as having no clear sex-specific preference for crossing over: 40–60% for chromosomes 1, 2 and 19; and 45–65% for chromosome 18.

Using SNP data from various RI panels (Shifman *et al.*, 2006), sites where a crossover had occurred during RI inbreeding can be easily identified. We used strains from five RI lines: BxD (C57BL/6J x DBA; 34 strains), BxA and AxB (C57BL/6J x A/J and vice versa; 18 and 22 strains, respectively), BxH (C57BL/6J x C3H/HeJ; 16 strains) and CxB (BALB/cByJ x C57BL/6J; 13 strains) to identify crossover breakpoints located in the chromosomal regions of interest. In all RI lines, one of the parents was C57BL/6J because the cytological data were from mice of this background (de Boer *et al.*, 2006).

To select hotspot candidates for further analysis, we took into account the number of occurrences of crossovers in different RI lines, presence of reciprocal recombinant configurations in different lines, and the number of known SNPs in the region. For each target chromosomal region, we selected two RI crossover breakpoints (16 in total; see table below). The selected breakpoints mapped to very large regions, ranging from ~70 kb to ~850 kb, so we used the mouse phenome database from Jackson labs (http://phenome.jax.org) to locate SNPs at various positions within the CO breakpoint region. Short stretches surrounding the SNPs (typically ~1 kb) were amplified from genomic DNA (purchased from the Jackson Laboratories) from both parental strains as well as from the RI strains involved. SNPs were subsequently

genotyped by sequencing. This was repeated until a ~10 kb region containing the crossover breakpoint remained, which we then sequenced.

Despite an initial group of 16 crossover breakpoints as possible hotspots, we were only able to use one in this study because most hotspot candidates did not meet the requirements for analysis. Since the conversion tracts that are associated with noncrossovers are short, a high SNP density within the hotspot is required for the detection of noncrossovers. In addition, sufficient SNPs located outside but nearby (within ~5 kb) are required for the allele-specific PCR needed in both the crossover and noncrossover/crossover assays. Finally, because both assays are PCR-based, it is important to have few repeat elements surrounding the hotspot. As many as 13 candidate breakpoints had to be discarded because they failed to meet one or more of these criteria (see table below). Thus, while the crossover library in RI strains may be useful for prospective hotspot identification in some circumstances (Bois, 2007), its overall success rate in identifying hotspots suitable for molecular assays is too low to make it a viable alternative to more direct DSB mapping via SSDS or SPO11-oligo sequencing.

References

Baker, S. M.*, et al.* (1996). "Involvement of mouse Mlh1 in DNA mismatch repair and meiotic crossing over." Nat Genet **13**(3): 336-342.

Bois, P. R. J. (2007). "A highly polymorphic meiotic recombination mouse hot spot exhibits incomplete repair." Mol.Cell.Biol. **27**(20): 7053-7062.

Cole, F.*, et al.* (2010). "Comprehensive, fine-scale dissection of homologous recombination outcomes at a hot spot in mouse meiosis." Molecular Cell **39**(5): 700-710.

de Boer, E.*, et al.* (2006). "Two levels of interference in mouse meiotic recombination." Proc Natl Acad Sci USA **103**(25): 9607-9612.

Froenicke, L.*, et al.* (2002). "Male mouse recombination maps for each autosome identified by chromosome painting." Am.J.Hum.Genet. **71**(6): 1353-1368.

Kauppi, L.*, et al.* (2004). "Where the crossovers are: recombination distributions in mammals." Nat.Rev.Genet. **5**(6): 413-424.

Shifman, S.*, et al.* (2006). "A high-resolution single nucleotide polymorphism genetic map of the mouse genome." PLoS Biol **4**(12): e395.

Analyzed crossover breakpoints from RI strains.

			Initial breakpoint from RI strains				Results of refined breakpoint mapping				
					Located between				Located between		
Breakpoint	Chrom.	Region	Initial size breakpoint	RI strains	SNP (position)	… and SNP (position)	Narrowed down to	RI strains	SNP (position)	… and SNP (position)	Comment
BP1_2	1	central	~372 kb	AxB8 BxA13	rs6182343 (78,315,951)	rs13475939 (78,687,715)	~0.2 kb	BxA13	rs6292825 (78,590,768)	Indel (78,590,953)	used in this study.
							~11 kb	AxB8	rs30859973 (78,605,548)	rs30663826 (78,616,144)	discarded; many repeats, low SNP density.
BP1_6	1	central	~433 kb	AxB21 AxB23 BxD18 BxD31	rs6234856 (87,916,394)	rs13475970 (88,349,770)	~217 kb	AxB21 AxB23 BxD18	rs30301268 (88,133,257)	rs13475970 (88,349,770)	
							~17 kb	BxD31	rs30305484 (88,080,647)	rs31323177 (88,097,782)	
BP1_21	1	distal	~272 kb	CxB6 AxB15	rs13476283 (184,905,583)	rs6154379 (185,177,962)	~0.5 kb	AxB15	rs33827280 (185,015,170)	rs33827277 (185,015,554)	discarded; low SNP density, repeats in region surrounding the breakpoint.
							~19 kb	CxB6	rs33830201 (185,145,767)	rs33827771 (185,164,928)	discarded, no known SNPS in ~19 kb region between parental strains involved
BP1_30	1	distal	~158 kb	BxD13	rs3708442 (192,343,890)	rs3689947 (192,501,995)	~13 kb	BxD13	rs33062268 (192,421,794)	rs32103256 (192,435,216)	
BP2_2	2	central	~847 kb	BxA26	rs13476594 (79,768,919)	rs3722345 (80,615,688)	~34 kb	BxA26	rs28315693 (80,242,258)	rs28302339 (80,276,369)	discarded; low SNP density in ~34 kb region.
BP2_8	2	central	~513 kb	BxA26 AxB2	rs13476687 (106,727,593)	rs6242654 (107,240,990)	~0.5 kb	BxA26	rs28330238 (106,751,765)	rs28330235 (106,752,168)	discarded; CO breakpoint in 403 bp SNP-free region, surrounding region also low SNP density
							~36 kb	AxB2	rs27492383 (107,049,776)	rs27492209 (107,085,355)	

BP2_12	2	distal	~111 kb	BxA7 BxD22 BxD9	rs3726974 (169,841,980)	rs3143843 (169,952,882)	~5 kb	BxA7 BxD22 BxD9	rs27628756 (169,886,635)	rs27628737 (169,891,409)	discarded; CO breakpoints in 4774 bp SNP-free region
BP2_16	2	distal	~71 kb	BxD22 BxD5	rs13476936 (179,208,083)	rs6335805 (179,278,998)	~6 kb	BxD22 BxD5	rs52534685 (179,250,964)	rs27683534 (179,257,053)	discarded; CO breakpoint in ~6.1 kb region without SNPs between parental strains
BP18_2	18	central	~231 kb	AxB8	rs3676196 (42,923,810)	rs13483328 (43,154,669)	~17 kb	AxB8	rs29578559 (43,097,768)	rs31900117 (43,115,130)	discarded; low SNP density in ~17.4 kb region.
BP18_11	18	central	~132 kb	BxD31	rs13483381 (58,263,501)	rs13483382 (58,395,122)	~13 kb	BxD31	rs30260734 (58,372,106)	rs30205618 (58,384,987)	discarded; low SNP density and multiple repeats in ~13 kb region.
BP18_13	18	distal	~374 kb	AxB2 BxA26	rs13483437 (73,844,689)	rs13483438 (74,218,827)	~77 kb	BxA26	rs13483437 (73,844,689)	rs38325583 (73,921,549)	
							~86 kb	AxB2	rs37541434 (73,923,173)	rs29556648 (74,008,689)	
BP18_16	18	distal	~412 kb	BxD13 BxD27	rs13483436 (73,432,362)	rs13483437 (73,844,689)	~204 kb	BxD13 BxD27	rs13483436 (73,432,362)	rs6377427 (73,637,771)	
BP19_2	19	central	~483 kb	AxB21	rs6291559 (27,573,224)	rs6238322 (28,056,254)	~36 kb	AxB21	rs36294339 (27,597,466)	rs38490227 (27,633,343)	discarded; low SNP density and multiple repeats in ~36 kb region.
BP19_10	19	central	~137 kb	AxB5 AxB13	rs3710829 (38,168,496)	rs13483617 (38,305,197)	~137 kb	AxB5 AxB13	rs3710829 (38,168,496)	rs13483617 (38,305,197)	trouble with SNP typing because of repeats
BP19_13	19	distal	~139 kb	AxB2 CxB5	rs8257607 (47,864,318)	rs13483653 (48,003,548)	~5 kb	AxB2 CxB5	rs30668843 (47,958,306)	rs31272779 (47,963,411)	discarded; low SNP density at and surrounding the CO breakpoint
BP19_15	19	distal	~236 kb	BxA26 BxD36 BxD21	rs3023496 (48,137,866)	rs3660360 (48,373,875)	~13 kb	BxA26 BxD21	rs31222821 (48,314,544)	rs49222411 (48,327,789)	discarded; CO breakpoint in ~13 kb region without SNPs
							~21 kb	BxD36	rs30851180 (48,211,635)	rs39432548 (48,232,486)	discarded; low SNP density and multiple repeats in ~21 kb region.

Table S3A. Hotspot in the central region of chromosome 1: primers used for amplification and sequencing

Primer	Sequence
SNP1_BP1_2_for1	TTTGCTGGTTTATTGCTGGGTC
SNP1_BP1_2_rev1	AGCACACAAAGAACTCCAAATCC
SNP2_BP1_2_for1	ATGCTTAGAAGCCAGCTGTATGTG
SNP2_BP1_2_rev1	TGCAGCGATCATCCACTCAG
BP1_2c_for_1	CCCAGAACAATGTCAATCAGC
BP1_2c_rev_1	CTTTTGAAGGTAGCTTTGCGG
BP1_2c_for2	TCTGCCCTGAAGAGTTTTGAGC
BP1_2c_rev2	AGGGTTTCTGGCCTGTCATTG
BP1_2e_for1	GTACGCAATACAGTGGAGACAGGAG
BP1_2e_rev1	TGTGACAATAATGGCAGCAGGAG
BP1_2f_for1	GATGCCAGCTTCTCCGACCTT
BP1_2f_rev1	TATGAAGGGAAGAGAAAGGCAACC
BP1_2i_for1	TCTTCCCTGGCTACAACCTTCAA
BP1_2i_rev1	TGCAAGGATGAGAGTGACAGAGG
BP1_2j_for1	GATGCCAACAACCTATGACTGCTC
BP1_2j_rev1	AGTGTGCTCTCAATGCCAAATCTC
BP1_2jc_for1	ATGGCCTTCCTTCAAATCTTCCT
BP1_2jc_for2	AAGAAATTAAAGGGCTGGCAAGAT
BP1_2jc_for3	TAGAAAGGCAGCAGGGTCACAA
BP1_2jc_for4	CCCAGAACGTATTTCCACCCAA
BP1_2jc_for5	TCCTCCCAGAACAATGTCAATCAG
BP1_2jc_for6	GTGATGCTAAGGGGTGAAGAAGAC
BP1_2jc_rev1	AGATGGCACAGAACACAAATAATGG
BP1_2jc_rev2	TCAGCACTGGGTACTTTCCTCTTT
BP1_2jc_rev3	TGGAATGCTGGCTTAAATGAAATC
BP1_2jc_rev4	CACCCCTTAGCATCACATCTTCAT
BP1_2jc_rev5	CCGTGTGATGTGATGTCCAGTATT
BP1_2jc_rev6	CTAGAGGCAATAAATCTTTTCCACATC
BP1_2jc_rev7	GTATGTACTTTGAGGCAGGATGTCTAC
BP1_2jc_rev8	CTGTATTGTTAGGCTTGATTTATGTGTCT
BP1_2jc_rev9	GGGGCAGCAACTCCTTTCAG
BP1_2jc_for7	CTCACTCCTCACATGAACAATGCT
BP1_2jc_for8	AGGCCTTATTAATGCTGCTACTGC
BP1_2jc_for9	CCTTCCCCCTTCATCAATGC
BP1_2jc_for10	GACAGCTCCCTTCTCCTCCTCT
BP1_2jc_for11	GCCAGTAAGATGAGTCAGCAGGTA
BP1_2jc_for12	GAGAAACTCTTCTAGCCAGAGCATTG
BP1_2jc_for13	GATGGGGTTTCTCAGTTGGGTT
BP1_2jc_for14	GACAAACCAGCCACCCTCACTT
BP1_2jc_for15	GCTGCCATGCCTGAAGACTC

Table S3B. Hotspot in the central region of chromosome 1: Universal PCR primers.

Primer	Sequence
BP1_2jc_3467749F1_U	AAGCCTGATCATGGTGACACAGGC
BP1_2jc_for4	CCCAGAACGTATTTCCACCCAA
BP1_2jc_rev14	CTAATGGCAGATGCTTCCGTTTTG
BP1_2jc_rev2	TCAGCACTGGGTACTTTCCTCTTT

Table S4A. Hotspot in the distal region of chromosome 1: primers used for amplification and sequencing.

Primer	Sequence
seqG1_1761814F	ATCCCTAATTCTTGCCCCTACTT
seqG1_1762630F	GTGTCCGGTGTGGTCTATCTTGAA
seqG1_1763061F	AAATCCAGACGAACCCAGAGAAGG
seqG1_1763869F	CGATAATTCATTCAGACTTCAGGA
seqG1_1764109F	GGATTCAAGGGGAGCCATTTACA
seqG1_1764427F	AGTAGGGCTGTTTCTGGGTTATTG
seqG1_1764850F	GGCAGAACACCCTAACACCCTACA
seqG1_1764960R	CCAGCAGAGGCCAAGGACT
seqG1_1764390R	AGGCCTCACTTCCCACACTATT

Table S4B. Hotspot in the distal region of chromosome 1: Universal PCR primers.

Primer	Sequence
1761560F1_U	GCAAAAAGTGTAGTCGGGCATAAC
1760576F1_U	TTTGTGCATCATTAGCTCAGGTCT
1765562R1_U	CAGAATCATCGAGACGTTTGCTAA
1766602R1_U	CTCAAGCTCCCAAGGTCTATCTAC

Table S5. Hotspot compilation used in Figure 3B

Hotspot	Chrom.	Start	End	Interval size	SSDS hotspot start*	SSDS hotspot end*	DSB activity (SSDS count)*	CO frequency (cM)	F1 hybrid assayed for CO	CO freq. source**	Notes
HS14.9	19	14954681	14961578	6897	14957522	14959522	1066	0.0052	BxD	1,2	Similar initiation frequency inferred on both alleles
HS18.2	19	18219446	18224030	4584	18220397	18222397	123	0.0135	BxD	1,2	Reciprocal CO asymmetry, predicts DBA allele hotter than B6 allele
HS23.9	19	23913421	23918537	5116	23916042	23918042	180	0.0027	BxD	1,2	Reciprocal CO asymmetry, predicts DBA allele hotter than B6 allele
HS44.2	19	44222000	44228101	6101	44222189	44224189	3610	0.024	BxD	1,2	Insufficient data to evaluate symmetry
HS48.3	19	48285649	48290677	5028	48288465	48290465	311	0.0015	BxD	1,2	Similar initiation frequency inferred on both alleles
HS61.1	19	61085324	61090636	5312	61087502	61089502	738	0.0506	BxD	1,2	Similar initiation frequency inferred on both alleles
HS61.2	19	61226601	61232402	5801	61228218	61230218	2178	0.0578	BxD	1,2	Similar initiation frequency inferred on both alleles
A3	1	161954686	161956666	1981	161954749	161956749	1984	0.26	BxD	3	Similar initiation frequency inferred on both alleles
M1	1	145721107	145726510	5403	145723315	145725315	79	0.005	BxC	4	Similar initiation frequency inferred on both alleles
Ebeta	17	34446742	34452746	6004	34449044	34451044	167	0.02	B10xA?	5	Similar initiation frequency inferred on both alleles
central	1	78585902	78588950	3048	78586497	78588497	2814	0.92	AxB	This study	Similar initiation frequency inferred on both alleles
subtelomeric	1	187089370	187093372	4002	187090129	187092832	9247	0.42	AxB	This study	Similar initiation frequency inferred on both alleles

Scnm1	3	95521800	95525145	3345	95521294	95523294	300	0.1105	BxC	6	Pedigree data: 85 cM/Mb in 1.3 kb hotspot
HS9	19	10741609	10745069	3461	10743299	10745299	2117	0.0477	BxD	2	Insufficient data to evaluate symmetry
HS22	19	23033612	23038358	4747	23035147	23037147	329	0.0217	BxD	1,2	Reciprocal CO asymmetry, predicts B6 allele hotter than DBA allele
HS37	19	38747388	38755916	8529	38751793	38753793	1084	0.0345	BxD	2	Insufficient data to evaluate symmetry

All coordinates are NCBI Build 37

*SSDS hotspot coordinates and strength are from Brick et al. (2012).

**References:
1. Wu et al. (2010).
2. Bois (2007).
3. Cole et al. (2010).
4. Kauppi et al. (2007).
5. Yauk et al. (2003).
6. Buchner et al. (2002).

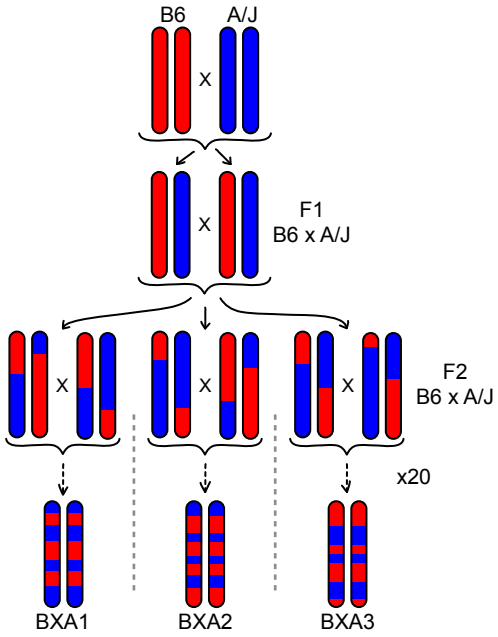

Figure S1. Breeding scheme for generating recombinant-inbred strains.
RI strains are named using abbreviations for the maternal and paternal progenitor strains. Each RI strain can be viewed as a library of crossovers that occurred during the generations of inbreeding prior to fixation of homozygosity. Adapted from Bois (2007).