Supplementary Methods

Data processing

We sequenced more than 120 mice, aiming to get 10x coverage for the 2 F0 mice and 36 F4 mice, and 20x coverage for the 11 F2 mice and 72 F5 mice. Sequencing was carried out on the Illumina Hiseq2500 platform for the 2 F0 mice and for 4 of the F2 mice, and on the Illumina Hiseq4000 platform for the remaining 7 F2 mice and all of the F4 and F5 mice. Genomic DNA was fragmented to an average size of 500 bp and subjected to DNA library creation using established Illumina paired-end protocols (Nextera DNA Library Prep). Sequencing reads were aligned to mm10 using BWA[2] (v. 0.7.0) followed by Stampy[3] (v. 1.0.23, option bamkeepgoodreads). We then used Picard tools (v. 1.115) (http://broadinstitute.github.io/picard) to merge bam files from different lanes for the same sample and mark the duplicated reads. Then we used GenomeAnalysisTK-3.3-0 (GATK) to do local Indel realignment using known Indel targets between B6 and CAST from the 4th version of the Mouse Genome Project (MGPv4) data[4], followed by base quality score recalibration using known sites from SNPs between B6 and CAST in MGPv4, and then we called the variants using UnitedGenotyper in GATK. Next we used the Variant Quality Score Recalibrator (VQSR) from the GATK for variant filtration, where we used the set of variants present on the Affymetrix Mouse Diversity Genotyping Array as a set of true positive variation[5]. We used the annotations "HRun", "HaplotypeScore", "DP", "QD", "FS", "MQ", "MQRankSum", and "ReadPosRankSum" to train the VQSR, and we used a sensitivity threshold of 90% for the true positive set to define the set of newly genotyped sites that passed VQSR filtration. After filtration, about 16 million variants remained. To remove potential hidden heterozygous sites from the F0 individuals and to get a more stringent set of SNPs to start with, we intersected our SNPs with variants that have the homologous reference

allele genotype for B6 and homologous alternative allele genotype for CAST from MGPv4 [4]. Only SNPs with a PASS quality score were used. After filtering, we obtained 13,946,562 and 13,940,079 reliable autosomal SNPs from F2 samples and F5 samples, respectively, as informative markers to detect recombination events, or roughly one SNP for about every 170 base pairs.

HMM algorithm to identify events

Using the information from the filtered strain-informative SNPs, we developed a Hidden Markov Model (HMM) to infer the strain origin of each broad segment of the genome. In our HMM, the three possible emitted genotype states B6/B6, B6/CAST and CAST/CAST are represented by 0, 1 and 2, respectively (i.e. the number of CAST allele copies at each strain-informative SNP site). Similarly, the hidden states representing background strain origin are encoded as 0, 1 and 2 copies of a CAST haplotype. Emitted states may be different from hidden states due to sequencing errors or real converted events (e.g. observing a homozygous CAST genotype on an otherwise heterozygous CAST/B6 background). A natural initial stationary distribution is (0.25, 0.5, 0.25) corresponding to state triple (0, 1, 2). The state transition between two sites is driven by recombination events, with the distance between two different states following an exponential distribution with a rate parameter equal to twice the recombination rate. Here we adopted a genome-wide average constant recombination rate of $r=0.625*10^{-8}$ per base pair per generation[6,7]. Thus, the probability of recombination from site *i* to site *j* can be written as follows:

$$P_{ij}=1-\exp(-2rD_{ij}),$$

where P_{ij} and D_{ij} stand for the recombination probability and distance between site *i* and *j*, respectively. The transition probability matrix from site *i* to site *j* is as follows:

$$\mathbf{P}_{ij}=(1-P_{ij})\mathbf{I_3}+P_{ij}\mathbf{Q}, \quad (2)$$

where I_3 is the 3×3 identity matrix and Q stands for the conditional transition matrix with the entry q_{mn} (m=0,1,2; n=0,1,2) describing the transition probability from state *m* to state *n*:

$$\mathbf{Q}=\begin{bmatrix} 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \end{bmatrix}. \quad (3)$$

There is no transition from state 0 to state 2, or vice versa, because it's unlikely that two independent recombination events would happen at exactly the same position with a small sample size. Conditional on there being a recombination event, state 0 or state 2 transitions to state 1 with probability 1, and state 1 transitions to either state 0 or state 2 with equal probability.

Here we defined the emission probabilities from each hidden state by using the quality metrics from GATK for states 0, 1 and 2. Given state *g* in each site *t*, GATK provides a quality score S for three states as follows:

$$s_g^t = -10\log_{10}\frac{p(D|G_t=g)}{\max_{k=0,1,2} p(D|G_t=k)}, \quad (4)$$

where $p(D|G_t = g)$ is the probability that we observe the data D, conditional on the hidden state G_t being *g*. Since for each site *t*, the maximum score is constant, we can inversely infer the probability of observing different states with a constant scale factor:

$$p(D|G_t = g) \propto 10^{-\frac{s_g}{10}}. \quad (5)$$

In our analysis, the scaling parameter was arbitrarily set to 1.

We applied the forward-backward algorithm to infer the posterior distribution of hidden states. Starting with prior state probabilities (0.25, 0.5, 0.25) at the first site, the forward probability of state *j* after seeing the first t sites is

$$A_t(j) = \sum_{i=0}^{2} \alpha_{t-1}(i) p_{ij}\ (t-1) e_j(t), \quad (6)$$

where p_{ij}(t-1) is the (*i*,*j*)th element of transition matrix **P** at site t-1, and $e_j(t) = p(D|G_t = j)$ is the emission probability conditioned on state *j* at site t given by equation 5. At the same time, we define a backward chain with an initialised probability (1, 1, 1) at the end of the site using the following:

$$\beta_t(j) = \sum_{k=0}^{2} \beta_{t+1}(k) p_{jk}\ (t) e_j(t+1), \quad (7)$$

and the probability of hidden state *j*, given the observed data (*j*=0,1,2) at site t is

$$p_t(j) = \frac{\alpha_t(j)\beta_t(j)}{\sum_{i=0}^{2} \alpha_t(i)\beta_t(i)}. \quad (8)$$

Finally, we can calculate the stationary distribution of states 0, 1 and 2 for each strain-informative SNP site given the sequencing data, and for each site we choose the hidden state with maximum probability as the real strain background state at that site.

Filters

Using the HMM method described above, we identified sites on the autosomes that have conflicting genotype calls with their background as potential converted sites. Most of these represent sequencing errors. We identified a region on Chromosome 6 from 37000000-56000000 (mm10) that is not fully inbred and is enriched with potential converted sites for most of our samples. We removed sites in this region from further analysis. We produced a series of filters to remove false positive sites, including hidden heterozygous sites in the F0 mice, false heterozygous calls due to unequal numbers of reads supporting the two alleles, false homozygous calls due to low read depth, and others (Table S1). After all these steps, the number of potential converted sites dropped sharply, e.g. from 863,082 SNPs potentially converted, to 183 distinct identified NCOs for the F2 mice.

By counting the inferred strain background transitions on the autosomes, we were able to identify potential COs. For COs in F2, we removed double COs (COs that are within 50 SNPs of each other and share the same flanking background) considering that pairs of nearby COs happening in one meiosis are rare. (We considered such events as potential NCO events, instead.). For COs in F5 animals, we used an HMM algorithm as described in[8] to identify *de novo* COs (those occurring in the germ cells of the F4 parents). For COs in F4s, if COs happen exactly at the same positions and share the same direction, they are treated as duplicates from the same ancestral CO. We kept one of these duplicates and also removed double COs, yielding a set of unique inherited COs for downstream analyses.

Of 1575 observed NCO events, only 9 of them are complex; of 1116 observed *de novo* CO events from F2 and F5, 7 were complex.

Power to identify NCOs

To estimate the power of our method to detect NCO events of varying tract lengths, we simulated NCOs with different mean tract lengths and ran our pipeline for discovering NCO events, including our filters. Because F2 events are controlled by both $Prdm9^{Hum}$ and $Prdm9^{Cast}$ and F5 *de novo* events are controlled by $Prdm9^{Hum}$ alone, we performed two sets of simulations by using data from 11 F2 samples and 72 F5 samples. Because most recombination events overlap hotspots, we simulated NCOs in hotspot regions. For each mean tract length, we sampled 2000 hotspots with probabilities proportional to their H3K4me3 enrichment. Within each sampled hotspot, we sampled the centre of the NCO tract according to the distribution of NCOs around PRDM9 motifs after correcting for SNP density, and we sampled its tract length from an

exponential distribution with a pre-defined mean tract length (which we varied from 10 to 100 bp with step size 10 and from 150 to 300 bp with step size 50). Sampled NCO tracts containing 0 SNPs were not counted as potentially detectable. Across these 2000 tracts, different animals possess different ancestral backgrounds. For each tract in each animal, we checked if any of the other animals has a different ancestral background consistent with a gene conversion event in the first animal. If so, we sampled such a "donor" mouse (other events were ignored). We copied the sequencing information corresponding to the converted sites from the donor mouse, such as the allele depth, and we copied the sequencing information for the background from the recipient, such as mate-pair information. Then, we applied the same filters to this simulated sequencing data at each sampled tract. We calculated our power by dividing the total number of simulated tracts left after filtering by the total number of simulated tracts overlapping at least one SNP (Extended Figure 1a, b).

Assignment of parental origin to CO and NCO events

For COs from F5 samples, we used an HMM algorithm[8] to identify *de-novo* COs and assign their parental origins by using genotypes from parents and children. All COs present in F4 parents are inherited COs. For NCOs, we are only able to confidently assign parental origin when one of the parents is heterozygous at the converted sites, but the other one is homozygous. When this is the case, the NCO must be inherited from the heterozygous parent.

DMC1 ChIP-seq

DMC1 ChIP-seq data were generated elsewhere[9] in joint submission with this manuscript. Briefly, single-stranded DNA sequencing (SSDS) DMC1 ChIP-seq was performed as previously described in Khil *et al*. 2012[10], using testes from one of the male F1 mice (B6xCAST)F1$^{Prdm9\ Hum/Cast}$ (C57BL/6J-*Prdm9*$^{Hum/Hum}$ mother, CAST/Eij father). ChIP and total chromatin DNA samples were sequenced in multiplexed paired-end Illumina HiSeq2500 libraries (rapid run), yielding 252 million 51-bp read pairs. We the processed the data for this study by following the algorithm provided by Khil *et al*. 2012 to map the reads to mm10 and obtain type I reads. We then used the same pipeline to call DMC1 peaks as described in Davies *et al.* 2016[1].

H3K4me3 ChIP-seq

We performed ChIP-seq against H3K4me3 as previously described[1] with several important modifications that increased ChIP stringency (noted here). An 8-week-old male (B6xCAST)F1-*Prdm9*$^{Hum/Cast}$ mouse C57BL/6J-*Prdm9*$^{Hum/Hum}$ mother, CAST/Eij father) was culled and the testes were immediately removed. Each testis was processed separately as a replicate. The testis tunica was removed, and the tubules were dissociated with tweezers and fixed in 1% formaldehyde in PBS for 5 minutes followed by glycine quenching (125 mM final concentration) for 5 minutes at room temperature. Following washing steps, pellets were resuspended in 900 µl cold 1% SDS lysis buffer (1% SDS, 10 mM EDTA, 50 mM Tris pH 8.0, 2x protease inhibitors), dounced 20 times and sonicated in 300 µl aliquots in a Bioruptor Twin sonication bath at 4°C for three 5-minute periods of 30s on, 30s off at high power, then cell debris were pelleted and removed and aliquots were pooled. Samples were then diluted 1:10 in IP wash buffer (100 mM Tris pH 7.5, 500 mM LiCl, 1% NP-40, 1% Na deoxycholate, 2x Protease Inhibitor, filtered). For

each sample, 100 µl of magnetic beads (with no primary antibody; Invitrogen Sheep Anti-Rabbit Dynabeads) were washed and resuspended three times in PBS/BSA (1X PBS, 5 mg/ml BSA, 1 tablet Roche Complete protease inhibitor per 50 ml, filtered with 0.45-micron filter) to a final volume of 100 µl and added to the chromatin samples for pre-clearing for 2 hours at 4°C with rotation. Beads were removed, and 100 µl of pre-cleared chromatin was set aside for the input control. 5 µl rabbit polyclonal anti-H3K4me3 antibody (Abcam ab8580) was added to the remaining pre-cleared chromatin and incubated overnight at 4°C with rotation. 50 µl fresh beads were washed and resuspended in PBS/BSA as before, then incubated with the chromatin samples for 2h at 4°C with rotation. Beads were then washed 5 times for 3 minutes each with cold IP wash buffer (with rotation), then washed once with cold 1X TE (10 mM Tris-HCl pH 7.5, 0.1 mM Na_2-EDTA). Bead pellets were resuspended in 200µl room-temperature IP elution buffer (1\% SDS, 0.1 M $NaHCO_3$, filtered with a 0.45-micron filter unit) and vortexed to mix. For input controls, 50 µl of pre-cleared chromatin was added to 150 µl IP elution buffer. Samples were incubated in a 65°C water bath for 1 hour with mixing at 15-minute intervals to uncouple beads from protein-DNA complexes. Samples were centrifuged (14,000 rpm, 3 mins) and placed on a magnet to pellet beads, and supernatants were isolated and then incubated at 65°C overnight to reverse crosslink. After de-crosslinking, samples were further incubated with 80 µg RNAse A at 37°C for 60 minutes and then with 80 µg Proteinase K at 55°C for 90 minutes. DNA was purified using a Qiagen MinElute reaction cleanup kit, yielding roughly 1 ng per testis (owing to higher stringency antibody incubation in wash buffer).

ChIP and input chromatin DNA samples were sequenced in multiplexed 51-bp paired-end Illumina HiSeq2500 libraries (rapid run). Raw read pair numbers were 90466668, 72740395, and

85980827 for the Input, ChIP replicate 1, and ChIP replicate 2 samples, respectively. Filtered read pair numbers were 69956995, 63065759, and 70685675, respectively. Sequencing reads were aligned to mm10 using BWA aln[2] (v. 0.7.0) followed by Stampy[3] (v. 1.0.23, option bamkeepgoodreads), and reads not mapped in a proper pair with insert size smaller than 10 kb were removed. Read pairs representing likely PCR duplicates were also removed by samtools rmdup[11]. Pairs for which neither read had a mapping quality score greater than 0 were removed. Fragment coverage was computed at each position in the genome and in 100-bp non-overlapping bins using in-house code and the samtools and bedtools[12] packages. Peaks were called as described in our previous work[1] and Altemose *et al.* 2017[13]. The percentage of ChIP-seq read pairs originating from signal (as opposed to background) was estimated to be 87.4%, a significant improvement over our prior, less stringent, experimental method (which yielded 62-71% of read pairs from signal)[1].

H3K4me3 enrichments were computed at DSB hotspots identified by DMC1 ChIP-seq, using our previously published force-calling method (*12,17*). H3K4me3 hotspots have also been called *de novo*, without using DMC1 peaks, using the same approach (*12,17*). The *de novo* calls were used to generate a list of regions likely to be trimethylated independently of PRDM9, by intersecting calls in mice with different *Prdm9* alleles (as described in *12*). In comparisons involving both DMC1 and H3K4me3 data, we used H3K4me3 enrichments force-called at DMC1 peaks. We only used *de novo* H3K4me3 calls for checking overlap between recombination events and PRDM9 binding sites.

Attribution of *Prdm9* control in hybrids

The *Prdm9* alleles in the hybrid mouse in this work are $Prdm9^{Hum}$ and $Prdm9^{Cast}$. Using DSB maps from other samples ($B6^{14}$, $B6^{B6/Hum}$, $B6^{Hum/Hum}$, $B6^{-/-14}$, B6xCAST, PWD, B6xCAST$^{Hum/Cast}$, (B6xCAST)F2$^{B6/Hum}$, PWDxB6, B6xPWD, PWDxB6$^{Hum/PWD}$, B6xPWD$^{Hum/PWD}$), we were able to classify DSB hotspots as being under the control of either the humanized or CAST *Prdm9* allele in the hybrids. The procedure we used is as follows, and we thank Anjali Hinch for suggesting it:

We create a superset of hotspots, which are combined across mice such that hotspots that have their centres within 600 bp of each other are merged to form a single hotspot. For each hotspot in this superset, we then create a maximal set of *Prdm9* variants that could potentially be responsible for activating it. For example, if a hotspot overlaps hotspots in B6xCAST and B6xCAST$^{Hum/Cast}$, the potential set of alleles that could activate the hotspot is $Prdm9^{B6}$, $Prdm9^{Cast}$, and $Prdm9^{Hum}$. From the maximal set, we reduce to a minimal set of alleles that can explain all of the mice in the set. In the example above it is $Prdm9^{Cast}$. Hotspots for which the minimal set consists of a single *Prdm9* variant are inferred to be activated by it. There are two special cases: Hotspots in the $B6^{-/-}$ are said to arise from a dummy allele $Prdm9^{KO}$. Hotspots that overlap with these hotspots are assigned the allele "KO". $Prdm9^{Cast}$ and $Prdm9^{PWD}$ have similarities in their zinc finger arrays and a large number of overlapping hotspots. If the maximal set contains more than one of these variants, we treat them as equivalent. It is not always possible to reduce the minimal set to a single *Prdm9* variant. For example, if a hotspot is found in B6xCAST, B6xCAST$^{Hum/Cast}$ and $B6^{Hum}$, then no single *Prdm9* variant can explain all the hotspots. The maximal set cannot be reduced from $Prdm9^{B6}$, $Prdm9^{Cast}$ and $Prdm9^{Hum}$. In this case, we take the following approach to assign alleles in the B6xCAST$^{Hum/Cast}$ mouse that is of interest in this work:

1. For hotspots where the minimal set contains both *Prdm9*^{Cast} and *Prdm9*^{Hum}, we say the allelic type is "*unknown*" or "*MULT*".

2. For hotspots where the minimal set contains *Prdm9*^{Cast}, but not *Prdm9*^{Hum}, the allelic type is "*CAST*".

3. For hotspots where the minimal set contains *Prdm9*^{Hum}, but not *Prdm9*^{Cast}, the allelic type is "*HUM*".

Testing for correlation of different recombination-related phenotypes at chosen scales

We assume that given an underlying vector of (binned) mean values W_k along the genome, the kth recombination-related quantity (number of observed recombination or non-crossover events in various classes), N_{ik}, follows a Poisson distribution with mean W_{ik}, in interval i. The W_k means vary along the genome and represent the underlying recombination rate parameters; this model is accurate provided (as is likely to be case) for a single meiosis, the number of expected events in each bin is small. Then the variance

$$Var(N_{ik}) = E\left(E\left({N_{ik}}^2 \middle| W_{ik}\right)\right) - E(W_{ik})^2 = E(W_{ik}) + Var(W_{ik})$$

This enables estimation of the variation in recombination rate along the genome, using the usual standard estimates of the mean and variance of the number of events, across bins genome-wide:

$$Var(W_{ik}) = Var(N_{ik}) - E(N_{ik})$$

Further the covariance

$$Covar\left(N_{ij}, N_{ik}\right) = E\left(E\left(N_{ij}N_{ik} \middle| W_{ij}, W_{ik}\right)\right) - E(W_{ik})E\left(W_{ij}\right) = Covar\left(W_{ij}, W_{ik}\right)$$

Combining these results enables estimation of the underlying correlation between W_j and W_k along the genome based on properties only of the observed Poisson counts N_j and N_k:

$$Cor(W_{ij}, W_{ik}) = \frac{Covar(N_{ij}, N_{ik})}{\sqrt{Var(N_{ik}) - E(N_{ik})}\sqrt{Var(N_{ik}) - E(N_{ik})}}$$

The quantities in the above equation are all estimated in the usual way using standard estimates of mean and variance from the observed vectors of counts.

At any interval size scale, we bootstrap resampled (10,000 times) the resulting disjoint intervals of the genome, to compute confidence intervals for the estimator.

Motif Analysis

We used a Bayesian, *de novo* motif finding algorithm to identify motifs within DSB hotspots (described in *14*, *26*). For each DSB hotspot that is controlled by *Prdm9*^{Cast}, a 1000-bp sequence (centred on the hotspot centre) was extracted from the reference sequence (mm10). *Ab initio* motif identification was performed on the centre 600-bp sequences from the top 1000 hotspots (ranked by DMC1 heat) that contained no bases overlapping annotated repeats. Motif calling proceeded in two stages: seeding motif identification, and motif refinement. Each seeding motif was obtained by first counting all 10-mers present in all input sequences, and from the top 50 most frequently occurring 10-mers, the one with the greatest over-representation in the central 300 bp of each peak sequence was chosen. This seeding 10-mer was then refined for 50 iterations as described in Davies *et al.* 2016. This refined motif was then force-called on the full set of the hotspots (without filtering) by rerunning the refinement algorithm, providing a probability of motif occurrence within each hotspot, and also identifying the most likely motif

location in each case. This motif was reported for each peak, along with position and strand. We did the same for DSB hotspots controlled by *Prdm9*^{Hum} and a 48-bp human motif was identified.

We used the SNPs generated as described above to determine if a SNP is in the motif or not. If a motif tract covers any SNP, then there is a SNP (SNPs) in the motif. The distance from a motif to an event is defined as the centre of the motif to the event tract (lower bound for NCOs). If the distance from the event to the nearest motif is less than 1000, we label the event as overlapping a motif. If any of the converted sites in a NCO overlaps a motif, we say that the NCO contains a SNP in the motif.

Estimation of NCO tract length for human-controlled and CAST-controlled events

To estimate NCO tract length, we assume the converted tract follows an exponential distribution with rate parameter λ, where 1/λ is the mean tract length. (If exponential tract lengths are not a fully accurate model, we can view this as a summary of tract properties, estimating the probability of co-conversion of pairs of markers as the distance between them increases.) We computed a composite likelihood function for our NCOs and estimated λ via maximal likelihood. Specifically, for each converted site, viewing this site as a "focal" site, we examine the SNPs nearby and record for each SNP its distance from the focal SNP, and whether that SNP is also converted. If the SNP is also converted, then it is still in the gene conversion tract, otherwise it is not. Using this approach allows our approach to be independent of SNP density, because we are conditioning on SNP positions in our analysis. The probability that a SNP nearby a converted site is also converted is

$$\Pr(\text{SNP nearby converted}) = \Pr(\text{in}) = \mathrm{e}^{-\lambda d}$$

d is the distance from the nearby SNP to the converted site. The probability that a SNP nearby a converted site is not in the tract is 1-Pr(in). All the NCOs are independent so we can multiple these probabilities for each SNP in the windows to get the (composite) likelihood of the data:

$$\Pr(\mathrm{D}) = \prod\prod_{\mathrm{all\_pairs}} \Pr(\mathrm{in})^{x} \left(1 - \Pr\left(\mathrm{in}\right)\right)^{1-x}$$

Here x=1 if the SNP nearby is also converted and x=0 otherwise. By maximise the likelihood using grid search for $1/\lambda$ from 1 to 1000 with step 0.1, we gained an estimate of tract length. Because pairs of SNPs are not in fact independent, this is not a true likelihood (though the resulting estimator is statistically consistent as the number of independent conversion events increases), and so to estimate uncertainty in the resulting estimates, we utilised bootstrapping of NCO events.

To perform bootstraps, we separate autosomal genomes into 258 non-overlapping 10 Mb blocks (the last block in each chromosome is shorter than 10 Mb). We resample 258 blocks with replacement, where the probability of sampling each block is proportional to the length of that block, and from the resulting bootstrapped set of NCO's, re-estimate tract length via the same procedure. Confidence intervals are calculated from a total of 10,000 bootstraps. We implemented this procedure for two sets of NCO events; those overlapping human-controlled, and those overlapping CAST-controlled, hotspots respectively.

Calculation of number of recombination events in one meiosis

We assume that the average number of DSBs per meiosis resolving as NCO events is K. Because each NCO affects only one of four chromatids, only one quarter of them will be seen in a single offspring.

We take F2 animals as an example. 22 meiosis occur, and generate 11 F2 animals. If D is hotspot SNP density, L is average NCO tract length, and "Power" represents the power to detect a SNP within a NCO event, then if N is the number of converted sites observed, we have:

$$E(N) = \frac{\mathrm{K}}{4} * 22 * \mathrm{Power} * \mathrm{L} * \mathrm{D}$$

Values for N, L, "Power" and D together allow estimation of K. We observe 0.0072 SNPs per bp within hotspots, and N=240 distinct converted sites in total; moreover, we estimate tract length L=30, and a power of 74.3% for these animals. This yields an estimate of $\widehat{K} = 274$ DSBs resolving as NCO events, per meiosis.

For crossover events, we have near 100% power to observe these, and half of all recombination CO events are transmitted to a particular offspring. Therefore, based on 295 observed CO events in these mice, the (sex-averaged) estimated number of CO events is 295x2/22=26.8 per meiosis.

The sum of these numbers is the total number of autosomal events repairing using the homologous chromosome, per meiosis (we neglect the X-chromosome in this calculation). To obtain confidence intervals for the number of NCOs, COs and the total number of recombination events per meiosis and for the NCO to CO ratio, we performed bootstrapping as to estimate the tract length of NCOs. For each bootstrapped sample (of 10,000), we obtained the number of NCOs and number of COs, and used these to re-estimate the total number of recombination events and the NCO/CO ratio.

Testing and characterizing the bias towards GC in NCO events

To test for the presence of GC bias in NCO events, we first combined inherited and *de novo* NCO events inferred to be under $Prdm9^{Hum}$ control, and occurring within (<1 kb from) DSB hotspots identified using our DMC1 data. This identified a total of 1,011 SNPs within such NCO tracts. We focus only on human-controlled events because the resulting hotspots are newly introduced, and so unlike $Prdm9^{Cast}$-controlled DSB hotspots cannot have been influenced by historical recombination, e.g. generating an excess of mutations towards G/C carried on the CAST genome.

Initially, we simply tested for an overall "GC bias" by testing for an excess of SNPs converted from A/T towards G/C vs. the converse direction, yielding strong evidence (59.9% towards GC, p=3.7x10^{-9} by 2-sided binomial test). This bias somehow occurs despite the great majority of our detected NCO events (encompassing 99% of all converted SNPs) representing simple, contiguous converted tracts. This raises the question of whether multiple co-converted SNPs still show a GC bias. More generally, a mechanistic understanding of the process of GC-biased gene conversion in mammals remains elusive.

We therefore separately tested for directional GC bias of converted SNPs where the adjacent SNPs in our set (among the 14,334,181 genome-wide) were either both non-converted, or where at least one of the adjacent SNPs was converted. These represent conversion of a single isolated SNP, vs. SNPs falling within a multiple-SNP tract, respectively. We observed a strong conversion bias for isolated SNPs (68.1% towards GC, p=1.9x10^{-15}). In complete contrast there is no evidence of bias whatsoever for SNPs within multiple-SNP tracts (50.4% towards GC,

p=0.921). This difference is highly significant ($p= 1.1x10^{-7}$, by Fisher's Exact Test). Thus, GC-biased gene conversion appears to effectively exclusively influence the shortest conversion tracts, of single SNPs. Our findings resolve the apparent paradox of why in our study, GC-biased gene conversion does not result in a high rate of observed complex NCO events within DSB hotspots. Below, we analyse occasional complex events that are observed.

We split NCO events according to sex, underlying hotspot heat and hotspot symmetry, and NCO position relative to the PRDM9 binding motif (Table S3; Figure S4): none of these had any effect on this pattern, or the quantitative GC-bias of single-SNP NCO events. To study symmetry, we split single-SNP NCO events into four quartiles based on their DMC1 symmetry (proportion of DMC1 signal coming from the B6 chromosome), ignoring SNPs within 20 bp of the centre of an identified (humanized) PRDM9 binding motif because such SNPs might be involved in driving asymmetry itself. No impact of symmetry on the bias is seen, so GC-bias cannot be driven by DSB initiation bias.

To distinguish whether GC-bias is prevented by, or else prevents, multiple-SNP conversion tracts, we tested for a relationship between the strength of bias and the distance of a SNP to other SNPs in hybrid mice. SNPs nearby other heterozygous SNPs tend to be co-converted with those SNPs, meaning they will normally lie within multiple-SNP tracts, if converted. Therefore, if GC-bias is prevented by multiple-SNP conversion SNPs, such SNPs will show little or no bias. We therefore plotted distance to the nearest flanking SNP versus the strength of GC-bias (Figure 4B). SNPs very near to other SNPs show no GC-bias, whilst those >100 bp from the nearest marker show approximately the 68% bias among all single-SNP conversion tracts. The

extremely high observed rate of co-conversion of nearby marker pairs (95.1% of adjacent markers within 10bp of a converted marker are also converted) implies that the number of SNPs in a conversion tract influences biased repair processing of DSBs towards or away from GC, rather than some process operating the other way around.

Thus, the strength of GC-bias depends on local SNP density, implying that the same SNP where one allele is A/T and the other is G/C will have different conversion rates and biases in different individuals, depending on alleles at surrounding SNPs. In humans, SNP densities are low (roughly 1 SNP per kilobase in Europeans) and so multiple-SNP conversion tracts (other than very long, typically complex tracts) are unusual; therefore, the similar GC-biases observed for single-SNP conversion tracts, of 68%, imply a common process might act in both species.

We reasoned that we could leverage the "non-biased" NCO events within longer tracts to understand whether the bias towards GC might depend on the invading (i.e. donor) haplotype, or the recipient haplotype (in which the DSB occurred). There are 12 possible combinations of donor and recipient alleles: we estimated underlying (i.e. non-biased) proportions of each from the multi-SNP NCO events (we averaged e.g. A<-G and T<-C mutations via obvious strand symmetry to yield 6 pooled types, in generating confidence intervals). In single-SNP observed NCO events, we then plot the fraction of each of the 12 possible types, normalised by these non-biased proportions (Figure 4C). We calculated 95% confidence intervals for the odds of each of the 12 possibilities relative to the multi-SNP tracts (binomial test; 95% CI were calculated from 6 pooled event types, after strand flipping, to aid power, though we show results for all 12 single-SNP NCO types for completeness).

As an alternative, we used markers not involved in gene conversion events but immediately adjacent to converted markers to estimate the background probabilities of different SNP types within hotspots in the same way. This yielded essentially identical results (Figure S4B), with no significant difference in SNP composition between non-converted markers and markers in multi-SNP conversion tracts ($p=0.59$ by Chi-squared test with 5 d.f). We observed odds ratios >1 for all event types involving a recipient allele which is an A or T, and odds ratios <1 for all event types involving a recipient allele which is G or C. This suggests a bias driven by the potential recipient allele. At DSB sites, it is possible in principle for nearby SNPs to be successfully converted from the homologous chromosome, or a potential conversion rejected. Given we do not observe mutations towards/away from GC bases altering the DSB rate, the observed GC-bias of NCO events is most simply explained by a tendency for the rejection of conversion of single SNPs, if the *recipient* haplotype has a G or a C at the same position, e.g. through the action of MMR, BER or NER proteins on heteroduplex DNA. The relative odds of conversion occurring is approximately half of that if the recipient haplotype (that on which the DSB occurred) carries an A or T allele at the same position, so it is a strong effect. (We cannot determine whether the bias impacts only G or C bases, or both, because we do not observe strand for our NCO events.) Importantly, this rejection does not seem to obviously depend on the type of the *donor* allele (other than it mismatching). For example, G->C conversions appear to occur at the lower rate while A->T conversions occur at the higher rate. As a caveat, there are relatively few such events, because these transversion mutations are relatively rare, so the different rates observed for these events do not reach statistical significance ($p>0.05$).

In conclusion, our data imply a mechanism of GC-biased gene conversion arising through identification of mismatching bases in the donor haplotype relative to the recipient, sometimes leading to rejection of the donor allele if the *recipient* allele is a G/C base pair (Fig. 6). However, this "checking" process normally only occurs if there is just a *single* mismatching base within the potential conversion tract, so that potential multi-SNP conversion tracts show no bias.

Finally, we reasoned that rare observed complex recombination events (i.e. non-contiguous non-crossover and crossover tracts) might be explained by every occasional ability of the same process to "escape", and act even in multi-SNP conversion tracts. We observed 12 SNPs not undergoing conversion but flanked by converted SNPs, i.e. within (8) complex NCO events. We reasoned that the non-converted SNPs might be "rejected" by the above biased process, if in rare cases it is able to operate even in the context of a multi-SNP potential conversion tract. If so, we would predict the *non-converted* markers should tend to show bias, with G/C bases on the recipient chromosome. Strikingly, we find all 12 SNPs show G/C bases on the potential recipient chromosome (and varying bases on the potential donor chromosome). The probability of observing this pattern by chance is approximately $1/2^{12}= 0.0002$. Similarly, we observe 7 complex *de novo* CO events. All involve a single "missing" SNP (not uniquely identifiable in 2 cases), which may therefore be explained as a rejected conversion of one SNP. In 6/7 cases this missing SNP has a G/C base on the potential recipient chromosome. Taken together yields a p-value among all complex events of p= 0.00004 (Binomial test, 1-sided). A previous human study[15] found a similar bias of up to 100% in apparently converted sites within complex crossovers, which also involved single SNPs, so it appears this phenomenon may extend to humans. Therefore, complex recombination events can be reinterpreted as otherwise normal,

continuous-tract events, but where a SNP is "rejected" for conversion, by a near 100% GC-biased process. Moreover, this process involves rejection of bases where the potential recipient chromosome carries a G or a C, exactly as in NCO GC-bias.

Therefore, the apparently distinct phenomena of GC-biased NCO events, and the occurrence of occasional complex NCO and CO events, may all be explained by a common underlying model of biased repair, involving rejection of single "incoming" SNPs where the existing allele is a G or C base. Given it does not correlate with DSB initiation, this phenomenon most plausibly arises via biased heteroduplex repair machinery. The bias appears to be close to 100%, but it does not impact all NCO events, yielding a maximal observed NCO bias (identical in humans, and in mice) of around 68%.

We inferred the rate at which a strongly GC-biased repair process would have to occur to yield the observed GC-biased gene conversion rate at single-SNP sites (68% of converted sites being A/T to G/C). Given the results in complex NCO events, we assume this process has a GC bias close to 100%, which prevents conversion where the recipient chromosome is a G/C. However, that process only acts some of the time, say with probability p, while the normal strand-biased process occurs with probability 1-p. Then, given there's a single-SNP mismatch in a tract, and assuming the mismatch is GC->AT as often as AT->GC, each with probability q=0.5:

$$\mathrm{P}(\text{conversion GC to AT}) = q(1-p) \text{ and}$$

$$\mathrm{P}(\text{no conversion GC to AT}) = qp, \text{ and}$$

$$\mathrm{P}(\text{conversion GC to AT}) = q,$$

with the probability of no conversion from AT to GC being small, given estimates of the number of DSBs versus the number of CO/NCO events per meiosis.

Given our data,

$$0.68 = \mathrm{P}(\text{conversion AT to GC} \mid \text{conversion observed}) = \frac{q}{q+q(1-p)} = \frac{1}{2-p}.$$

Solving this yields p=0.53. Thus, the data can be explained by simple model in which a distinct mismatch repair process acts to prevent gene conversion at 53% of sites where the recipient chromosome contains a G/C at a single mismatch site (or, for example, at 100% of sites where the recipient chromosome contains a G at a single mismatch site; Fig. 6).

Hotspot symmetry calculation

Sequence differences between the CAST and B6 genomes allowed us to quantify the fraction of ChIP-seq signal (either DMC1 or H3K4me3), coming from the B6 and CAST chromosomes. This also allows us to determine whether individual hotspots in these hybrids were 'symmetric', with DSBs occurring equally on both chromosomes, or 'asymmetric', with a preference towards either the CAST or B6 chromosome.

Using SNPs distinguishing the B6 and CAST genomes, each type I read pair from a hybrid DSB library (DMC1 ChIP-seq) is assigned to one of the categories 'B6', 'CAST', 'unclassified' or 'uninformative' as in[1], except we replace PWD with CAST. For each DSB hotspot, the B6 cutting ratio was then computed as the fraction of 'B6' reads mapped within 1 kb of the hotspot centre, over the sum of 'B6' and 'CAST' reads in that region.

We followed a similar approach for H3K4me3 ChIP-seq, but we further corrected for background signal, as described in[1]. For both DMC1 and H3K4me3, we only defined the B6 cutting ratio provided we had at least 10 informative reads.

To order hotspots based on their symmetry, if the fraction of cuts estimated on B6 and CAST chromosome respectively were x, and $1-x$, we defined the overall hotspot "symmetry" as $4x(1-x)$. We obtained additional results for events initiating on a known homologue by using "homologous heat", defined as xh, where h is the estimated total heat of the hotspot, for events initiating on the CAST chromosome, and $(1-x)h$ for events initiating on the CAST chromosome. Note that separate estimates of hotspot symmetry and homologous heat may be obtained from both H3K4me3 and DMC1 ChIP-Seq data, for the same collection of hotspots. Because e.g. the H3K4me3 homologous heat captures how well the homologous chromosome is bound by PRDM9, it may be of stronger direct interest; however, this is only directly available for NCO events, whose initiating homologue is known. For CO events, to be conservative (even though we could attempt to make assumptions regarding conversion tracts to estimate homologous heat), we mainly used hotspot symmetry, which is strand-symmetric and ranges from 0 to 1 for hotspots with events completely on one chromosome, versus equally on both chromosomes[1]. For one set of plots (Extended Data Fig. 6), we used average homologous heat, defined as $2hx(1-x)$ (this averages homologous heat over the strand an event occurs on).

Calculating the fraction of asymmetric/symmetric hotspots containing a disrupting variant in the motif

To estimate the proportion of hotspots of different levels of initiation on B6/CAST chromosomes containing SNPs within their PRDM9 binding motifs, we first filtered to include only hotspots containing a clear motif (posterior probability >0.99). Secondly, we required at least 20 informative reads in our DMC1 data in order to accurately estimate the proportion of reads from B6, and 5 reads from each homologue covering the motif region, to ensure there are enough reads to identify variants if present. In Figure S5F, we then plot the fraction of hotspots in each binned level of initiation on the B6 chromosome containing a SNP or Indel (as called by GATK prior to VQSR, or Platypus). We found 96% of identified highly asymmetric hotspots where this fraction was <5% or >95% contained such a SNP, after additionally requiring the P-value (binomial test) of asymmetry is $<10^{-10}$, to examine those hotspots most highly asymmetric.

Rejection sampling for COs and NCOs, construction of Figures 5 and S6, and testing for impacts of asymmetry on event resolution

In testing for impacts of asymmetry on the number of NCO and CO events observed, relative to expectations from ChIP-seq data, we allowed for the following factors.

(i) Asymmetric hotspots have higher SNP density around binding motifs so have higher power to identify NCOs (CO events are not affected). We corrected this in our analyses of how symmetry impacts the number of observed NCO events in different hotspot types, by incorporating SNP density information to estimate power to detect events in each hotspot.

(ii) We tested for differences in the impacts of symmetry in COs vs, NCOs, males vs. females, *de-novo* versus inherited events, and for $Prdm9^{Hum}$-controlled versus

Prdm9^{Cast}-controlled events. Because the CAST allele has co-evolved with the *castaneus* genome, some impacts of symmetry on recombination event resolution might be impacted (see below).

(iii) Strand can be identified for NCO events but not for CO events, allowing us to analyse events at asymmetric hotspots initiating on each strand separately. In particular, we can identify whether the homologue is strongly versus weakly marked by H3K4me3, for a given event.

In all analyses, we focused on NCO and CO events occurring in the subset of hotspots containing an identified motif, and with well-defined estimated heat on each strand for both H3K4me3 and DMC1.

To correct for SNP density in our analyses of how symmetry impacts the number of observed NCO events in different hotspot types, we directly leveraged SNP density information to estimate (relative) power to detect events in each hotspot. For the hotspots that contain an identified motif, we give each SNP near the motif (<1 kb) a weight according to its location relative to the motif. The weight is defined using the distribution of NCOs around motifs (Fig. 3d), and so estimates the probability an NCO event initiating within the hotspot will incorporate this SNP (up to a constant of proportionality). Therefore, this quantifies the over-representation of this hotspot compared to the true number of DSBs resolving as NCO events occurring within it. Summing this weight over all SNPs then yields the relative power to detect bases falling within NCO events in each hotspot (so if there are no SNPs in a hotspot, the power to detect NCO events is zero, while hotspots with many SNPs near the motif itself have highest power).

This yields a weight w_i for hotspot i. Multiplying the original hotspot heat (from DMC1 or H3K4me3) gives a power-corrected heat for the hotspot, used to define expectations for observable NCO events, and compare to actual observed NCOs.

Construction of Fig. 5A-B and Figure S6 panels A-D:

For these figures, we compare CO and NCO events, for events of different types (e.g. male vs. female) and in different categories (e.g. PRDM9^{Cast} versus PRDM9^{Hum} controlled). To compare NCO and CO events, we used the weights w_i and rejection sampling. Within a hotspot and event category, we started with all observed events, and associated weight w_i for event i. Because NCO events are over-represented on average w_i-fold, relative to the ChIP-seq observations and to CO events, we retained NCO events with probability $\min\{\alpha/w_i,1\}$, and CO events/hotspots with probability $\min\{1, w_i/\alpha\}$, where α is any constant. For any hotspot, the probability of retaining a NCO event is then immediately w_i/α times lower than that of a CO event, and so this perfectly reverses the over-representation of this hotspot in observed, versus initiated, NCO events (the constant α only impacts the overall number of NCO vs. CO events retained, not their spread, so does not impact the validity of this point). In practice, we used α=0.7 to retain similar numbers of CO and NCO events.

We then obtain comparable lists of hotspots with various H3K4me3 heats and symmetries, and DMC1 heats and symmetries, as well as observed NCO and CO events, in any given category. For Figure 5A and S6A,C, we next ordered hotspots by their H3K4me3 symmetry, defining 3 bins with equal expected number of events, according to DMC1-predicted overall heat of each hotspot. We order by H3K4me3 in order that our symmetry estimates are independent of the

estimated heats; in practice, ordering by DMC1 symmetry made almost no difference to results (not shown). We compared the binned predictions to the actual number of events of different types observed – both NCO, and COs. We also obtained 95% confidence intervals of the fraction of observed events in each category by bootstrapping events 1000 times. To obtain p-values for asymmetric hotspots, we obtained exact binomial p-values, to test the null hypothesis that the true proportion of events occurring in the asymmetric hotspots bin is 1/3.

For Fig. 5B and similar Figure S6 panels, we performed the same analysis, but now ordered hotspots by their DMC1 symmetry, defining 3 bins with equal expected number of events, according to H3K4me3-predicted overall heat of each hotspot. This tests whether H3K4me3-defined heats, which measure the extent of PRDM9 binding in each bin, accurately predict where CO and NCO events occur. As before, though less strongly because DMC1 shows inflation in asymmetric hotspots (Extended Data Fig. 5g), we observe fewer events of all types in asymmetric hotspots, relative to expectations.

Construction of Extended Data Fig. 6 panels e-h:

This group of panels is constructed as Fig. S6A-D, but studies only crossover events, so no rejection sampling was required. Rather than symmetry, we order hotspots based on their H3K4me3 or DMC1-estimated average homologous heat, and predict events within bins using their overall signal of DMC1 or H3K4me3. We separate hotspots depending on whether they are human-controlled or CAST- controlled. Homologous heat provides slightly stronger signals than symmetry itself, implying CO events avoid weak hotspots as well as asymmetric hotspots, i.e. all hotspots where the homologous chromosome is bound weakly.

Construction of Extended Data Fig. 6 panels I-l:

This group of panels is constructed as Extended Data Fig. 6a-d, but studies only non-crossover events. To account for power, we therefore resampled hotspots with weights proportional to w_i, and compare observed NCO events to expectations under this resampling. For NCO events, we can determine which homologue they occurred on. This allows us to test whether "homologous heat" (see "**Hotspot symmetry calculation"**), i.e. the strength of DMC1/H3K4me3 signal on the homologous chromosome, might more strongly determine whether NCO events occur than our overall single symmetry measure for a hotspot. We therefore now separated the two homologues for each hotspot, resulting in a predicted (DMC1 or H3K4me3) heat for each homologue, as well as two complementary homologous heats. We ordered hotspots by this homologous heat (as in Fig. 5, we defined the fraction of events occurring on each homologue using the independent ChIP-seq data) and then used predicted heat to bin hotspots so that (as in Fig. 5) 1/3 of events are predicted to occur in each bin. We again separate hotspots depending on whether they are PRDM9^{Hum}-controlled or PRDM9^{Cast}-controlled, and otherwise proceed as in Fig. 5.

This revealed a strengthened signal relative to previous tests – few events are seen in hotspots with low homologous heat, i.e. where PRDM9 does not bind the homologous chromosome, implying NCO events strongly avoid occurring on both the hot allele of highly asymmetric hotspots, and at very weak hotspots. Conversely, NCO events occur preferentially on the cold allele of asymmetric hotspots, or strong hotspots more generally. Again, this occurs for both *Prdm9*^{Hum} and *Prdm9*^{Cast}.

Notably, although both differ significantly from expectations from both DMC1 and H3K4me3 data, we see somewhat stronger signals for *Prdm9*^{Hum} than *Prdm9*^{Cast}. Asymmetry in *Prdm9*^{Cast}-controlled hotspots is largely the result of evolutionary erosion, which can only occur when NCO or CO events themselves occur. Therefore, hotspots which evolve by chance to become asymmetric are preferentially sampled from those more active for these marks. This predicts that asymmetric hotspots for this allele may tend to have higher NCO and CO rates relative to PRDM9 binding strength as measured by H3K4me3, compared to these rates at random hotspots. In contrast, asymmetry at *Prdm9*^{Hum}-controlled hotspots is mainly due to SNPs occurring at random within the PRDM9 binding motifs inside these hotspots, so is immune to biases in NCO and CO rate, and this seems likely to explain our observation.

Asymmetry rather than SNP density affects the generation of recombination events

We fitted a generalised linear regression model to discern whether hotspot asymmetry or local SNP density better predicts low CO and NCO rates. For each hotspot containing an identified PRDM9 binding motif, we indicate if there is a CO event overlapping this hotspot. We use this to produce a binary response vector, and fit a binomial generalised linear model. As predictors, we used:

(i) The symmetry of the hotspot

(ii) The log-transformed 'heat' of the hotspot measured by H3K4me3 (the H3K4me3 heat is incremented by a small value 0.0001 as there are a few hotspots with zero heat)

(iii) SNP densities around the PRDM9 binding motif at different scales (±100 bp, ± 500bp, ±800 bp)

We then tested various coefficients for significance, conditional on the others. We did the analysis for *Prdm9*^{Cast}–controlled COs (all of them were generated in the meiosis from F1 where there are two different *Prdm9* alleles) and *de novo Prdm9*^{Hum}–controlled COs (all of them were generated in the meiosis from F4 where there is only one type of *Prdm9* allele) separately to avoid the effect of competition between the two alleles. Results show that conditioned on the heat of H3K4me3 and symmetry of hotspots, SNP density has no significant effect on where COs happen (p-values from all three scales >0.08) while both heat and symmetry of hotspots have significant positive effects on CO events ($p<0.05$).

For NCO events, we performed a similar analysis, except that we resampled the above hotspots according to the weight generated as described in the section "Rejection sampling for COs and NCOs, construction of Fig. 5 and Extended Data Fig. 6, and testing for impacts of asymmetry on event resolution" to account for higher power to detect NCOs when there is greater local SNP density**.** Some hotspots appeared several times after rejection sampling. The number of these hotspots that are indicated as overlapping a NCO depending on how many NCOs overlap this hotspot. Then we applied the same GLM analysis used for COs. All results show that SNP density has no significant effect on where NCOs happen conditional on the heat of H3K4me3 and symmetry of hotspots ($p>0.2$). For all the *Prdm9*^{Hum}–controlled NCOs, results show that the heat of H3K4me3 and symmetry of hotspots have significant positive effects on NCOs ($p<0.003$). Results from *Prdm9*^{Cast}–controlled NCOs also suggest positive effects on prediction of NCOs, but p-values are not significant (<0.2). We explained the weaker effect of symmetry for *Prdm9*^{Cast}–controlled NCOs in the last section of the supplementary material.

Reference and Notes

1. Davies, A. B. *et al.* Re-engineering the zinc fingers of PRDM9 reverses hybrid sterility in mice. *Nature* **530,** 171–176 (2016).
2. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25,** 1754–1760 (2009).
3. Lunter, G. & Goodson, M. Stampy: A statistical algorithm for sensitive and fast mapping of Illumina sequence reads. *Genome Res.* **21,** 936–939 (2011).
4. Keane, T. M. *et al.* Mouse genomic variation and its effect on phenotypes and gene regulation. *Nature* **477,** 289–94 (2011).
5. Yang, H. *et al.* A customized and versatile high-density genotyping array for the mouse. *Nat. Methods* **6,** 663–666 (2009).
6. Jensen-Seaman, M. & Furey, T. Comparative recombination rates in the rat, mouse, and human genomes. *Genome Res.* 528–538 (2004). doi:10.1101/gr.1970304.1
7. Brunschwig, H. *et al.* Fine-scale maps of recombination rates and hotspots in the mouse genome. *Genetics* **191,** 757–764 (2012).
8. Hinch, A. G. *et al.* The landscape of recombination in African Americans. *Nature* **476,** 170–5 (2011).
9. Hinch, A. G. *et al.* Novel factors influencing meiotic recombination revealed by whole genome sequencing of single sperm cells (in press).
10. Khil, P. P., Smagulova, F., Brick, K. M., Camerini-Otero, R. D. & Petukhova, G. V. Sensitive mapping of recombination hotspots using sequencing-based detection of ssDNA. *Genome Res.* **22,** 957–965 (2012).
11. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25,**

2078–2079 (2009).

12. Quinlan, A. R. & Hall, I. M. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26,** 841–842 (2010).

13. Altemose, N. *et al.* A map of human PRDM9 binding provides evidence for novel behaviors of PRDM9 and other zinc-finger proteins in meiosis. *Elife* 1–46 (2017).

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

15. Arbeithuber, B., Betancourt, A. J., Ebner, T. & Tiemann-Boege, I. Crossovers are associated with mutation and biased gene conversion at recombination hotspots. *Proc. Natl. Acad. Sci.* **112,** 201416622 (2015).