ARTICLES

High-resolution mapping of meiotic crossovers and non-crossovers in yeast

Eugenio Mancera[1]*, Richard Bourgon[2]*, Alessandro Brozzi[2], Wolfgang Huber[2] & Lars M. Steinmetz[1]

Meiotic recombination has a central role in the evolution of sexually reproducing organisms. The two recombination outcomes, crossover and non-crossover, increase genetic diversity, but have the potential to homogenize alleles by gene conversion. Whereas crossover rates vary considerably across the genome, non-crossovers and gene conversions have only been identified in a handful of loci. To examine recombination genome wide and at high spatial resolution, we generated maps of crossovers, crossover-associated gene conversion and non-crossover gene conversion using dense genetic marker data collected from all four products of fifty-six yeast (*Saccharomyces cerevisiae*) meioses. Our maps reveal differences in the distributions of crossovers and non-crossovers, showing more regions where either crossovers or non-crossovers are favoured than expected by chance. Furthermore, we detect evidence for interference between crossovers and non-crossovers, a phenomenon previously only known to occur between crossovers. Up to 1% of the genome of each meiotic product is subject to gene conversion in a single meiosis, with detectable bias towards GC nucleotides. To our knowledge the maps represent the first high-resolution, genome-wide characterization of the multiple outcomes of recombination in any organism. In addition, because non-crossover hotspots create holes of reduced linkage within haplotype blocks, our results stress the need to incorporate non-crossovers into genetic linkage analysis.

In most eukaryotes, homologous chromosomes exchange genetic information through recombination during meiosis. This process increases genetic diversity by breaking haplotypes, but it may also homogenize alleles through gene conversion[1,2]. Furthermore, recombination is fundamental to sexual reproduction because it provides physical connections between homologues during the first meiotic division, contributing to correct chromosome segregation[3]. In the current model, meiotic recombination starts with the formation of a double-strand break (DSB)[4,5]. The break is then repaired through a series of steps, involving resection, synthesis and ligation, using the homologous chromosome as a template. Repair results in either a crossover—reciprocal exchange accompanied by a tract subject to gene conversion—or a non-crossover—a tract subject to conversion but not associated with reciprocal exchange[4,6]. At least two pathways form crossovers: the Msh4/Msh5-dependent pathway, which proceeds through a double Holliday junction, and the Mus81/Mms4-dependent pathway[7,8]. In contrast, non-crossovers are thought to be the result of synthesis-dependent strand annealing[9]. It is known that DSB[10–14] and crossover rates[15] vary along chromosomes. Non-crossovers and crossover-associated gene conversions have not been characterized genome wide, however, because this requires monitoring recombination between closely spaced markers along the genomes of all four meiotic products[2].

High-resolution mapping of recombination

In *Saccharomyces cerevisiae* we achieved a detailed characterization of recombination outcomes by genotyping ~52,000 markers in all four viable spores derived from 51 meioses of an S288c/YJM789 hybrid strain[16,17] (Fig. 1). Genomic DNA from parental strains and each of the 204 spores was hybridized to high-density microarrays that tile the genomes of both S288c and YJM789 with a median probe offset of 4 base pairs (bp). To infer genotypes from the hybridization intensities of the probes covering each marker (eight probes per marker on average), we developed a new algorithm, ssGenotyping, based on semi-supervised clustering (see Methods). The high density of polymorphism and probes resulted in spore genotypes with a median distance of 78 bp between consecutive markers (Supplementary Fig. 1). This resolution is over 20 times higher than in the current yeast genetic map[15] and more than 360 times higher than in the most recent human crossover map[18].

Owing to their high resolution, our maps invert the traditional relationship between markers and recombination events: there are multiple markers within most recombination events rather than vice versa. This allows characterization of both crossover-associated and non-crossover gene conversion tracts, which are typically thought to be only 1–2-kilobases (kb) long[2]. Genotype calls from all four spores in each wild-type tetrad were used to infer a total of 4,163 crossovers and 2,126 non-crossovers (see Methods). We expect to have detected nearly all crossovers but, because non-crossovers have no effect on flanking markers, to have missed non-crossovers that completely fell between two markers, or non-crossovers in which mismatch repair restored the original genotype. We observed an average of 90.5 crossovers and 46.2 non-crossovers per meiosis. A total of 30.1% of observed crossovers occurred between two consecutive markers, and therefore had no detectable conversion tract. Taking this percentage as an estimate of the fraction of unobserved non-crossovers, we obtained a corrected total, 90.5 crossovers plus 66.1 non-crossovers, which is remarkably similar to a recent estimate of 140–170 DSBs per meiosis[13].

All chromosomes but one had at least one crossover, in agreement with the essential role that crossovers have in chromosome segregation[3]. The average number of crossovers was linearly related to chromosome length, with an intercept of 1.0, corresponding to one obligate crossover, plus an additional 6.1 crossovers per megabase

[1]European Molecular Biology Laboratory, Meyerhofstrasse 1, 69117 Heidelberg, Germany. [2]European Molecular Biology Laboratory, European Bioinformatics Institute, Cambridge CB10 1SD, UK.
*These authors contributed equally to this work.

(Mb) (Supplementary Fig. 5). Notably, non-crossovers behaved similarly (3.4 non-crossovers per Mb), but with a lower intercept (0.3).

The median size of conversion tracts was 2.0 kb for those associated with crossovers, and 1.8 kb for non-crossover conversion tracts (see Methods). The difference in medians is statistically significant (Wilcoxon rank-sum test, $P < 0.0001$). These sizes are consistent with previous estimates made at a single yeast hotspot[19], but are considerably larger than single-locus estimates in human[20]. Our finding that crossover tracts tend to be larger than non-crossover tracts also corroborates previous, single-locus observations in yeast and human[20,21].

We observed 57 non-crossover conversion tracts larger than 5 kb in size, the largest being 40.8 kb (minimal length). Three of these were

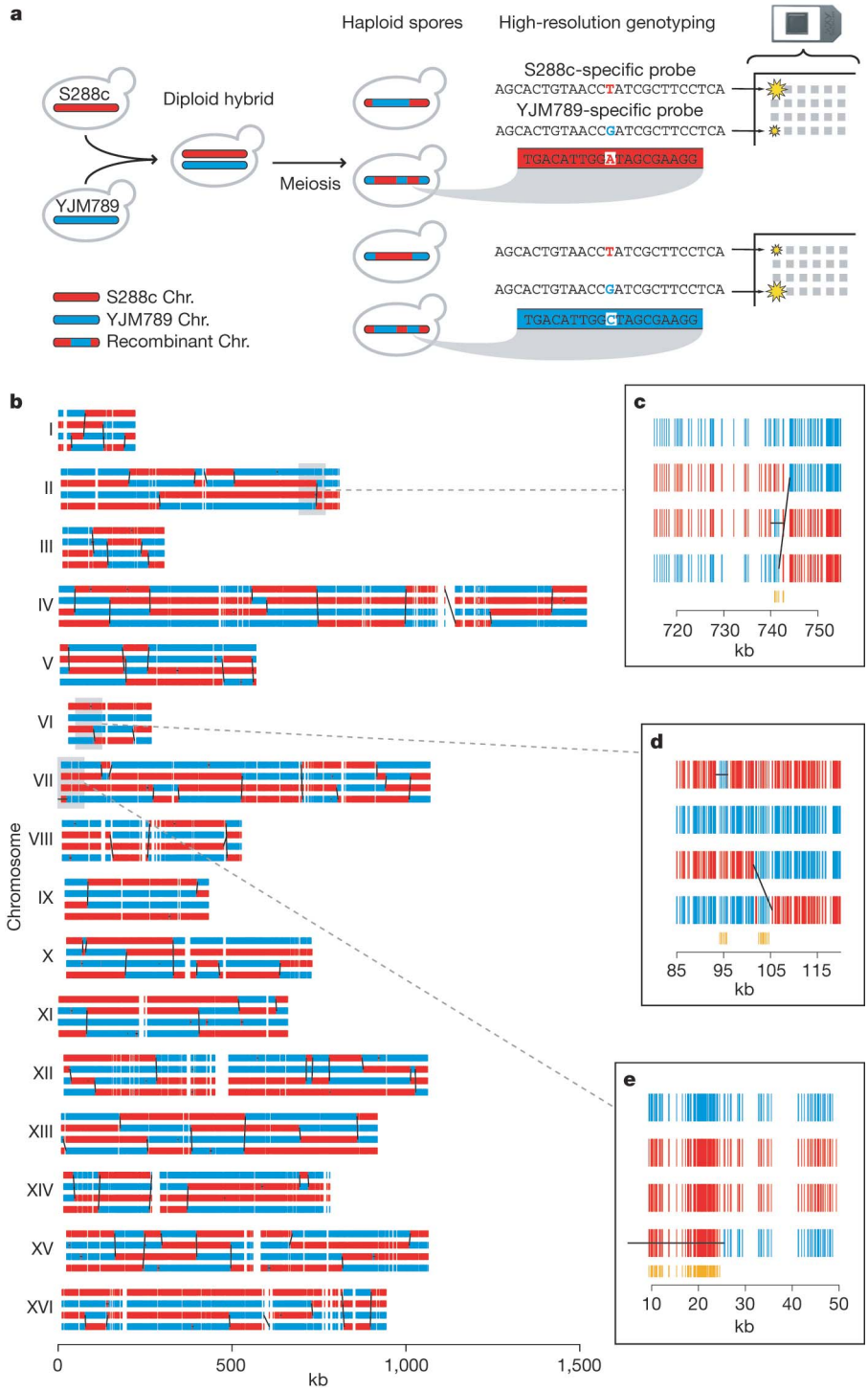

Figure 1 | High-resolution mapping of meiotic recombination along the yeast genome. **a**, Schematic description of the recombination mapping approach. Meiotic products from a hybrid derived from highly polymorphic strains were individually genotyped using microarrays. **b**, Genotype calls at ~52,000 markers for all four spores resulting from a single meiosis. Diagonal/vertical black lines are inferred crossovers; horizontal lines are non-crossovers. **c**–**e**, Close-ups of a crossover overlapped by an independent non-crossover in a third spore (**c**); a crossover with a complex gene conversion tract, and a nearby, independent non-crossover (**d**); and a long non-crossover at the end of the chromosome (**e**). (See Methods for full annotation procedure.) In close-ups, orange vertical segments denote markers with non-mendelian ratios (1:3 or 0:4).

found at the end of chromosomes, suggesting that they could be the result of meiotic break-induced replication, as has been proposed for long non-crossover tracts at the *HIS4* locus[22] (Fig. 1e). Three also showed complete loss of allelic variation across all four meiotic products (4:0 segregation), consistent with either mitotic or complex meiotic events.

We also observed that 11.5% of the conversion tracts accompanying crossovers exhibited complex patterns of genotype change (Fig. 1d). A total of 11.1% had more than one genotype change on just one of the involved chromatids, and 0.4%, on both chromatids. Such tracts are predicted to result from the resolution of a double Holliday junction owing to multiple distinct patches of heteroduplex in a single crossover event[6], but they could also possibly result from mismatch repair alternating between conversion and restoration. 3.4% of single-chromatid non-crossover events were also detected to have complex conversion tracts.

Recombination hotspots

To estimate the local recombination rate along the genome, we counted the events overlapping each intermarker interval and adjusted for the size of the interval (see Methods, Fig. 2b and Supplementary Figs 8 and 9). This novel approach was necessary because recombination events typically overlapped multiple markers, making existing rate estimation methods designed for low-resolution data inappropriate. Recombination hotspots were defined as runs of contiguous intermarker intervals involved in more recombination events than expected under a homogeneous genomic rate ($P < 0.001$, see Methods and Supplementary Information). We identified hotspots for crossover, non-crossover and overall recombination activity separately. At the hottest of the 179 resulting overall recombination hotspots, 27.7% of spores showed observable evidence of involvement in a crossover or non-crossover event (58.7% of meioses). At the hottest crossover and non-crossover hotspots, 21.7% and 8.7% of spores showed observable evidence of a crossover or a non-crossover, respectively. This corresponds to 21.7% and 17.4% of spores being involved in a crossover or non-crossover event, because a single crossover produces two spores with observable evidence whereas a non-crossover produces only one. Given that some non-crossovers may have been missed, we therefore observed similar rates for both outcomes at their hottest locations in the genome.

It is known that most DSBs occur in promoter regions[10,11], and indeed, 84% of hotspots overlap a promoter. Nonetheless, hotspot intervals primarily overlap coding sequence: only 25% of the bases in hotspot intervals overlap promoters, whereas 68% overlap coding sequences.

Centromere-proximal regions showed low recombination rates, and no recombination event overlapped a centromere on any chromosome (Supplementary Figs 8 and 9 and Supplementary Table 1). However, many chromosomes did have at least one event less than 4 kb away, including a crossover only 341 bp from *CEN5* (Supplementary Table 1). Telomeres could not be directly interrogated owing to repetitive sequence. We did, however, observe some chromosomes with a complete lack of recombination activity well before the telomeres; others showed strong activity near a telomere (Supplementary Figs 8 and 9).

Validating our approach, all previously known yeast recombination hotspots except for *HIS2* are within or adjacent to one of our hotspots (*HIS4*, *ARG4*, *CYS3*, *DED81*, *ARE1-IMG1*, *CDC19*, *THR4*, *LEU2-CEN3*)[23]. Furthermore, despite differences in strain background and the numerous heterozygosities in our hybrid strain, our recombination rates are in close agreement with a recently generated genome-wide DSB rate map in a homozygous SK1 strain[13] (Fig. 3). In addition to showing correspondence between the initiation of recombination and its resolution, this agreement suggests that the distribution of meiotic recombination is largely persistent within a species. Some fine-scale differences, however, do exist, possibly reflecting within-species variation in recombination rate[18].

Distinct crossover and non-crossover distributions

It is expected that the distribution of meiotic recombination is determined by the location of initiating DSBs as well as by how the DSBs are repaired[4]. It has not been clear, however, whether crossovers and

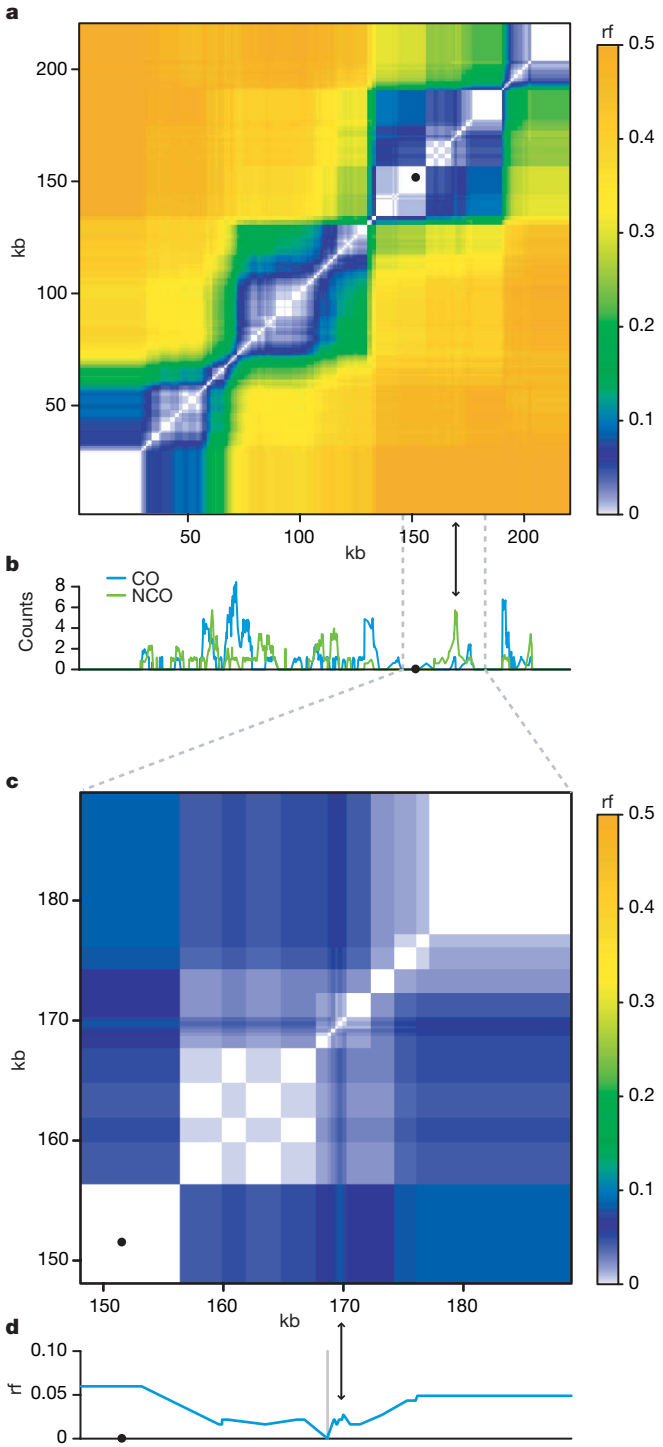

Figure 2 | Crossover and non-crossover rates along chromosome I and their effect on recombination fraction (rf). **a**, The recombination fraction was calculated for every pair of markers as the portion of segregants in which the markers have opposite genotype. Black dots denote the centromere. **b**, Crossover (CO, blue) and non-crossover (NCO, green) counts, adjusted for varying intermarker interval size (see Methods). **c**, Close-up of a non-crossover-biased region (indicated by arrows) on chromosome I. **d**, Recombination fraction relative to a single reference marker (grey vertical line) upstream of the non-crossover-biased region, showing the non-monotonic relationship between genetic and physical maps caused by such a region.

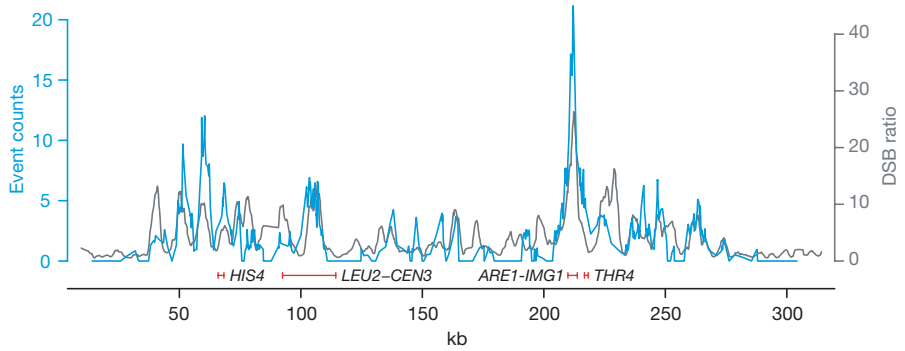

Figure 3 | Comparison of DSB and recombination rates along chromosome III. DSB smoothed fluorescence ratios in a SK1 strain (*dmc1*Δ, grey)[13] are compared with our recombination event counts (blue), after adjusting the latter for varying intermarker interval size (see Methods). Peak locations largely agree despite distinct strain backgrounds, although some fine-scale differences exist. Previously known hotspots are indicated by red segments.

non-crossovers always occur in similar proportions or whether there are crossover- or non-crossover-specific hotspots. Whereas a recent study reported mild crossover/non-crossover differences for two hotspots[24], our maps allow investigation of such differences genome wide (Fig. 2b). Using an approach that accounts for unobserved non-crossovers which fall completely between two markers, we identified regions with biased crossover/non-crossover ratios, and found more intermarker intervals with extreme ratios than expected by chance ($P < 0.0005$, see Methods). We observed an average excess of ~60 intervals favouring crossovers and ~170 intervals favouring non-crossovers, spanning ~100 kb of genomic sequence in total (see Methods). Notably, we estimated that such differences affect at least 1.4% of the genomic regions exhibiting one or more recombination event. The crossover/non-crossover event ratios at the regions showing the strongest evidence of bias, after accounting for the effect of marker spacing, were 14:0 and 0:7. Our findings therefore suggest that a significant fraction of the genome exhibits differences in crossover/non-crossover ratio.

The observed dissimilarity in crossover/non-crossover distribution has implications for linkage analysis. In contrast to crossover hotspots, regions with high non-crossover frequency can be expected to have reduced linkage to their surroundings, but to maintain linkage between loci to either side. By estimating the recombination fraction between all pairs of markers on each chromosome, we show that crossover hotspots are associated with linkage block boundaries, whereas non-crossover-biased regions correspond to regions with reduced linkage within blocks (Fig. 2). Non-crossover-biased regions result in a non-monotonic relationship between the genetic and physical distance, and create holes within linkage blocks. Over generations, non-crossover-biased hotspots would form genomic regions with low linkage disequilibrium relative to their surroundings, and thus be difficult to track with markers outside the non-crossover-biased region[25,26].

The existence of regions with a crossover/non-crossover bias suggests that the bifurcation between the two outcomes might, in fact, be a controlled process, influenced by local chromosomal properties. Recombination hotspots were found to contain short poly(A) stretches (20–41 bp) more frequently than expected, and to be significantly associated with several gene ontology (GO) terms (see Supplementary Information). Nonetheless, we found no sequence motifs to be specifically associated with crossover- or non-crossover-biased regions, and only one GO term ('cell ageing') to exhibit a significant association with such regions. A comparison of our results with measurement of transcriptional activity during meiosis in W303 and SK1 strains[27] showed that hotspot-proximal genes were significantly enriched in two specific expression profiles: a transcription peak around 2 h after meiotic induction ($P < 0.0001$, see Supplementary Information, Fig. 4a), and a transcription decrease between 8 and 10 h ($P = 0.0046$, Fig. 4b). In addition, a cluster with genes upregulated 4 h after meiotic induction contained genes from non-crossover-biased regions, but no genes from crossover-biased regions (Fisher exact test $P = 0.015$, Fig. 4c). This relationship between specific transcriptional behaviour and proximity to recombination hotspots supports a role for chromatin accessibility and transcription factor binding in meiotic recombination[28].

Crossovers and non-crossovers in recombination mutants

To assess the differences between the generation of crossovers and non-crossovers further, we mapped recombination events in *msh4* and *mms4* null mutants, in which either the Msh4/Msh5-dependent or the Mus81/Mms4-dependent crossover pathway is disturbed[7]. Five full tetrads of the *msh4* mutant were genotyped. Given the role of *MSH4* in crossover generation, its deletion is expected to reduce the number of crossovers but maintain the number of non-crossovers[29]. Consistent with this expectation, we observed that the non-crossover frequency showed no statistically significant change (*t*-test, $P = 0.12$), whereas the average number of crossovers per meiosis was markedly reduced from 90.5 in the wild type to 46.6 in *msh4* (*t*-test, $P < 0.0001$, Fig. 5a). Furthermore, in contrast to the wild type, all *msh4* tetrads except one had one or more chromosomes with no crossovers at all (6.3% of all chromosomes). Unexpectedly, the median size of *msh4* crossover conversion tracts was 479 bp larger than for wild type (Wilcoxon rank-sum, $P = 0.0003$). The median size of *msh4* non-crossover recombination tracts, however, was 338 bp shorter than for wild type (Wilcoxon rank-sum, $P = 0.0008$). Therefore, deletion of *MSH4* reduced genome-wide frequency of crossovers, as expected given its role in the Msh4/Msh5-dependent pathway, but affected tract size of both crossovers and non-crossovers (Supplementary Fig. 6).

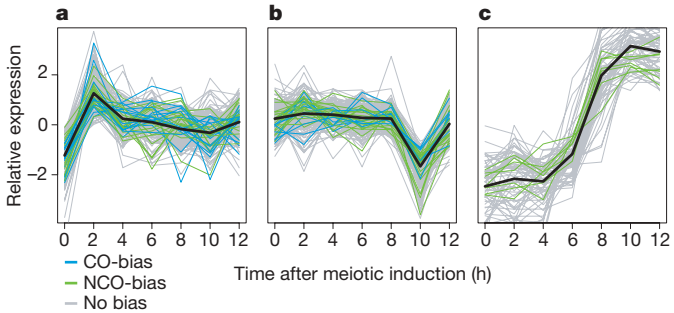

Figure 4 | Association between gene expression and recombination activity. After centring by their mean over time points, genes were clustered (see Supplementary Information) by their W303 expression profile during meiosis[27]. **a**, **b**, Two clusters enriched for genes overlapped by overall recombination hotspots. Genes overlapped by regions of crossover (CO) or non-crossover (NCO) bias, however, are present in similar proportion in these clusters. **c**, Gene cluster containing genes overlapped by regions with non-crossover bias, but no genes overlapped by regions with crossover bias.

The observation that, in the *msh4* mutant, the frequency of one event type was altered with respect to wild type whereas the other was not has two important implications. First, because Msh4 is thought to function downstream of DSB formation[30], we expect the *msh4* null mutant to have the same number of DSBs as the wild type. (This is known to be the case for *MSH5*, the functional partner of *MSH4* (ref. 31).) Our data therefore suggest that a fraction of DSBs are not resolved towards crossovers or non-crossovers, but may instead be repaired by alternative mechanisms such as sister chromatid exchange[32] or non-homologous end joining[4]. Second, we have perturbed a DSB-resolution pathway and seen strong but distinct effects on the global crossover/non-crossover balance. If this pathway has regional preferences, this may contribute to observed crossover/non-crossover bias.

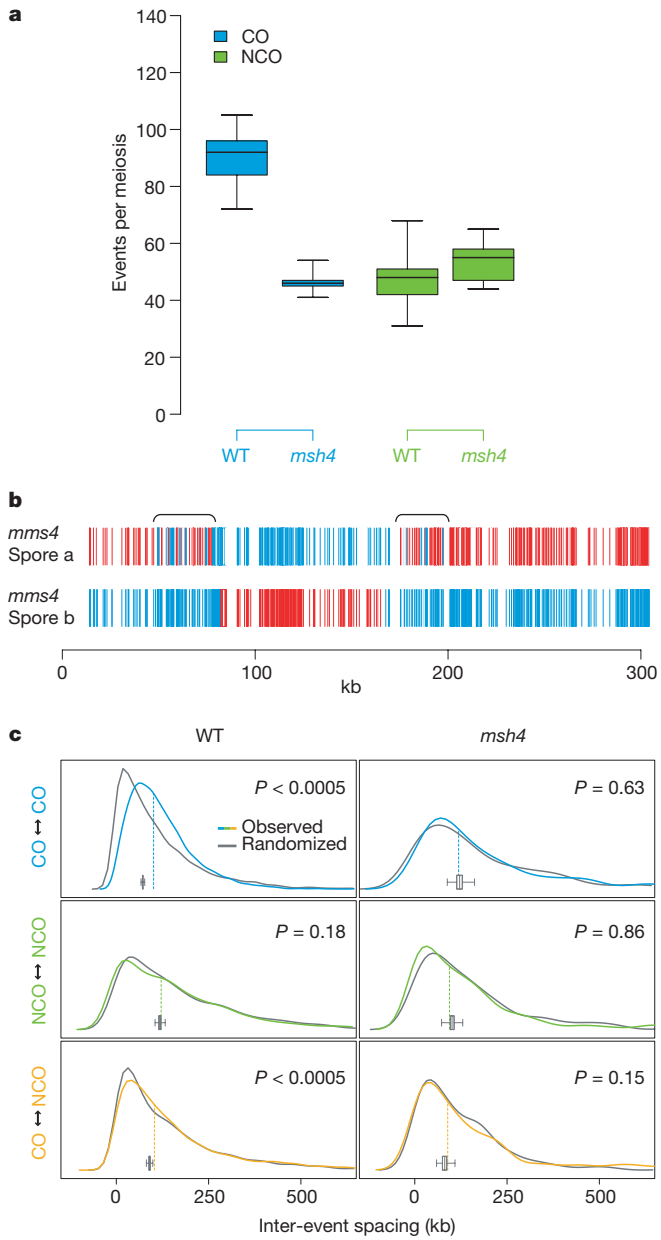

Figure 5 | Meiotic recombination in *msh4* and *mms4* strains. a, Genomic frequencies of crossovers (CO) and non-crossovers (NCO), per meiosis (wild type $n = 46$, *msh4* $n = 5$; box-plots show minimum, first quantile, median, third quantile and maximum). **b**, Genotype calls (S288c/YJM789, red/blue) along chromosome III in an *mms4* dyad showing large regions of frequent genotype change (calls in brackets validated by sequencing). **c**, Density estimates for the distance between adjacent recombination events, to measure interference. Coloured lines show the real data distribution, whereas grey lines denote one corresponding tetrad-randomized distribution. Dashed vertical lines show the real data median, and the box-plot shows the distribution of medians for 2,000 randomizations. *P*-values for difference in median distance are given within each panel (see Supplementary Information).

The *mms4* mutant exhibited low sporulation efficiency and spore viability, which impeded recovery of complete tetrads, so we only genotyped 6 dyads (12 spores) and 8 single *mms4* spores. Surprisingly, the *mms4* spores showed several regions (~7 per spore) exhibiting unusually frequent genotype changes (Fig. 5b)—up to ~70 kb in size and typically associated with apparent crossovers. For example, one such 63-kb region contained a total of 31 genotype changes. The mechanism responsible for these genotype changes is not known, but their presence may help elucidate the way in which the Mus81–Mms4 nuclease complex generates crossovers[8]. We chose not to pursue recombination event inference for the *mms4* spores owing to both the presence of such regions and the inherent difficulty in distinguishing between single non-crossovers and pairs of nearby crossovers in a single-spore context.

Crossover and non-crossover interference

Interference, where a recombination event reduces the probability that an additional recombination event occurs nearby[33], is an important determinant of the distribution of meiotic recombination, and could also contribute to differences in crossover/non-crossover rates. So far, interference has been reported only between crossovers[34]. To assess interference, we considered the distances between adjacent, same-tetrad recombination events. These distances were compared with those in tetrad-randomized data sets (see Supplementary Information). Tetrad randomization preserves hotspot and cold-spot structure along the genome, but removes interference effects. The distance between consecutive crossovers was larger in wild-type meioses than expected by chance: a median inter-crossover distance of 101.1 kb in observed data versus 71.8 kb under tetrad randomization ($P < 0.0005$, see Supplementary Information and Fig. 5c). No such effect was seen for non-crossovers. Notably, and in contrast to previous reports[34], crossovers and non-crossovers also exhibited interference: the median observed distance from a crossover to the nearest non-crossover was 13.1 kb larger in real data than under tetrad randomization ($P < 0.0005$). In the *msh4* null mutant, crossovers did not show interference ($P = 0.63$). This is consistent with the hypothesis that only crossovers generated by the Msh4/Msh5-dependent pathway exhibit interference[7]. Furthermore, in the *msh4* mutant, evidence of interference between crossovers and non-crossovers disappears as well ($P = 0.15$). These results support the existence of at least two types of crossovers with differences in interference, and yield genome-wide evidence for interference between crossovers, and among crossovers and non-crossovers.

We also observed an over-representation of overlapping events within the same meiosis in the wild-type strain, which is surprising given the observed patterns of interference. For example, 2.6% of crossover conversion tracts had an overlapping non-crossover partner on a third spore, and an additional 0.6% had an overlapping crossover partner involving the other two spores (Figs 1c and Supplementary Fig. 12). Such overlapping events could result from paired DSBs in two different chromatids; but, they could also be the consequence of a single DSB, the resolution of which involves multiple rounds of strand invasion and extension from different templates[35]. We also observed 110 pairs of partially or exactly overlapping non-crossovers with reciprocal genotypes. The existence of such pairs is relevant to current models for non-crossover formation (see Section 8 of Supplementary Information for discussion).

Genomic effect of gene conversion

Having observed differences in crossover and non-crossover distributions as well as interference between events, we next considered the effects of gene conversion tracts. We determined the portion of the

yeast genome that is involved in crossover-associated and non-crossover gene conversion. A total of 2.1% of the polymorphic positions was converted to the opposite genotype per meiosis. Furthermore, across the genomes of all four wild-type meiotic products, crossover tracts covered between 92 kb and 320 kb per meiosis (minimal and maximal), and the non-crossover tracts, between 62 kb and 148 kb. Therefore, as much as 1% of a meiotic product's genome may be subject to conversion in a single meiosis.

Genomic regions active in gene conversion are susceptible to the effect of gene conversion on allelic frequency, and also to mutation-prone processes[36]. We therefore analysed GC content and single-nucleotide polymorphism (SNP) density in converted regions and hotspots. For both crossover-associated and non-crossover gene conversions, we detected mismatch repair bias favouring GC nucleotides (Supplementary Information). Relative to the base content at SNP positions in the parental genomes, we observed a 1.4% GC increase in the converted sequences of the spores (χ^2 $P = 0.0001$, Supplementary Table 2). This bias could contribute to the association between recombination hotspots and GC-richness that we observed (χ^2 $P < 0.0001$)—an association that has also been found for DSBs[11]. Although on an evolutionary timescale, GC bias could potentially homogenize alleles, comparison to low-depth genome sequences of 37 *S. cerevisiae* strains showed that our hotspots were actually associated with greater genetic diversity (see Supplementary Information). Therefore, GC conversion bias may be counteracted by other processes, such as those that increase AT content[37,38]. We find no evidence of allelic homogenization at recombination hotspots, despite the presence of GC bias during mismatch repair.

Conclusion

The recombination maps presented here constitute the first survey of non-crossovers and both crossover-associated and non-crossover gene conversion across an entire genome in any organism. In addition to permitting detection and characterization of gene conversion, the high resolution of our approach reveals phenomena which would otherwise be difficult to observe, such as complex conversion tracts and large regions of frequent genotype changes (Figs 1d and 5b). The data uncover regions of interest for further investigation, and the approach is applicable to other mutants and conditions. It could thus contribute to answering questions about the mechanisms of interference and crossover homeostasis[24,39], or possible alternative DSB-resolution pathways[4–6].

Although the degree of polymorphism between the parental strains results in unprecedented marker resolution, polymorphisms may also affect recombination propensity[40,41]. Nonetheless, several observations suggest that recombination is not markedly perturbed in our hybrid: the agreement between our maps and the DSB map from a homozygous SK1 strain[13]; consistency between our overall number of crossovers and the number generated from genetic-map estimates[42]; and the detection of previously known recombination hotspots[23]. Furthermore, outside laboratory conditions, most sexually reproducing organisms are heterozygous. Individuals in natural populations may, therefore, resemble our hybrid more than they do a homozygous strain.

Our maps show the existence of locations with distinct preferences for either crossovers or non-crossovers, suggesting a role for genomic position in determining DSB resolution outcome. Given that chromatin conformation is known to be important for recombination generally[28], it is plausible that local chromosomal properties could influence the crossover/non-crossover bifurcation. Such properties may not, however, be the sole determinants of crossover/non-crossover bias. Through interference, both crossover–crossover and crossover–non-crossover, the decision could also depend on recombination activity in nearby regions.

Our maps also stress the relevance of non-crossovers, and gene conversion generally, in genetic analysis. Crossover is the major determinant of linkage disequilibrium, but both crossover-associated and non-crossover gene conversion weaken linkage disequilibrium between nearby loci. Models that incorporate gene conversion will therefore be able to relate linkage disequilibrium and physical distance more accurately. Furthermore, crossover-associated and non-crossover conversion tracts have different effects on the fine structure of haplotypes[26]. As shown in Fig. 2, gene conversion at crossover hotspots softens the boundaries of linkage blocks, whereas non-crossover-biased regions create holes within blocks. Both phenomena have implications for genetic association analyses. Although these regions are highly localized and have an impact on only a fraction of meioses, their effect can accumulate over generations, hiding genetic variants with phenotypic relevance (for example, disease genes). Having a higher density of markers in regions with frequent gene conversion may thus help to uncover genetic factors contributing to phenotypic variation.

METHODS SUMMARY

A S96/YJM789 hybrid strain was sporulated[43], and genomic DNA—from 51 wild type and 5 *msh4* tetrads as well as from 20 *mms4*, 13 S96 parental, and 12 YJM789 parental spores—was extracted from single-colony cultures and hybridized to a custom-designed tiling microarray[44]. (S96 is isogenic to S288c (refs 16, 17).) Normalized[45] fluorescence intensities corresponding to the set of probes covering each polymorphism were analysed by applying multivariate semi-supervised clustering to the combined parental and segregant data. Segregant genotypes were assigned using posterior probability of class membership. To reduce genotyping errors, we applied filters to whole arrays, to probe sets and to individual genotype calls. DNA sequencing of ~60 kb confirmed 100% of filtered genotype calls. After grouping data by tetrad, pairs of adjacent genotype change points isolated from all other changes were called non-crossovers if they involved one spore, or crossovers if they involved two. Complex groups of genotype changes were annotated as described in Supplementary Fig. 3. To calculate event rate along the genome, it was necessary to adjust for varying intermarker interval size. Because individual recombination events typically overlapped multiple intermarker intervals, a novel adjustment procedure was used (Supplementary Information). We defined three types of hotspots—crossover, non-crossover and overall recombination events—by identifying runs of contiguous intermarker intervals involved in more recombination events than expected under a homogeneous genomic rate. To assess crossover/non-crossover bias, we compared the number and size of intermarker intervals exhibiting more/fewer crossovers than expected to the corresponding null distribution, generated via simulation. We tested for interference—between consecutive events of the same type and also between crossovers and non-crossovers—by comparing the median distance between adjacent, same-tetrad events to medians computed after tetrad label randomization. This randomization strategy preserved hot- and cold-spot structure but removed interference.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 10 March; accepted 30 May 2008.
Published online 9 July 2008.

1. Gordo, I. & Charlesworth, B. Genetic linkage and molecular evolution. *Curr. Biol.* **11,** R684–R686 (2001).
2. Chen, J. M. *et al.* Gene conversion: mechanisms, evolution and human disease. *Nature Rev. Genet.* **8,** 762–775 (2007).
3. Page, S. L. & Hawley, R. S. Chromosome choreography: the meiotic ballet. *Science* **301,** 785–789 (2003).
4. Baudat, F. & de Massy, B. Regulating double-stranded DNA break repair towards crossover or non-crossover during mammalian meiosis. *Chromosome Res.* **15,** 565–577 (2007).
5. Bishop, D. K. & Zickler, D. Early decision; meiotic crossover interference prior to stable strand exchange and synapsis. *Cell* **117,** 9–15 (2004).
6. Whitby, M. C. Making crossovers during meiosis. *Biochem. Soc. Trans.* **33,** 1451–1455 (2005).
7. Argueso, J. L., Wanat, J., Gemici, Z. & Alani, E. Competing crossover pathways act during meiosis in *Saccharomyces cerevisiae*. *Genetics* **168,** 1805–1816 (2004).
8. Hollingsworth, N. M. & Brill, S. J. The Mus81 solution to resolution: generating meiotic crossovers without Holliday junctions. *Genes Dev.* **18,** 117–125 (2004).
9. Allers, T. & Lichten, M. Differential timing and control of noncrossover and crossover recombination during meiosis. *Cell* **106,** 47–57 (2001).
10. Baudat, F. & Nicolas, A. Clustering of meiotic double-strand breaks on yeast chromosome III. *Proc. Natl Acad. Sci. USA* **94,** 5213–5218 (1997).
11. Gerton, J. L. *et al.* Inaugural article: global mapping of meiotic recombination hotspots and coldspots in the yeast *Saccharomyces cerevisiae*. *Proc. Natl Acad. Sci. USA* **97,** 11383–11390 (2000).

12. Borde, V. *et al.* Association of Mre11p with double-strand break sites during yeast meiosis. *Mol. Cell* **13,** 389–401 (2004).
13. Buhler, C., Borde, V. & Lichten, M. Mapping meiotic single-strand DNA reveals a new landscape of DNA double-strand breaks in *Saccharomyces cerevisiae. PLoS Biol.* **5,** 2797–2808 (2007).
14. Blitzblau, H. G. *et al.* Mapping of meiotic single-stranded DNA reveals double-stranded-break hotspots near centromeres and telomeres. *Curr. Biol.* **17,** 2003–2012 (2007).
15. Cherry, J. M. *et al.* Genetic and physical maps of *Saccharomyces cerevisiae. Nature* **387** (suppl.), 67–73 (1997).
16. McCusker, J. H., Clemons, K. V., Stevens, D. A. & Davis, R. W. Genetic characterization of pathogenic *Saccharomyces cerevisiae* isolates. *Genetics* **136,** 1261–1269 (1994).
17. Mortimer, R. K. & Johnston, J. R. Genealogy of principal strains of the yeast genetic stock center. *Genetics* **113,** 35–43 (1986).
18. Coop, G. *et al.* High-resolution mapping of crossovers reveals extensive variation in fine-scale recombination patterns among humans. *Science* **319,** 1395–1398 (2008).
19. Borts, R. H. & Haber, J. E. Length and distribution of meiotic gene conversion tracts and crossovers in *Saccharomyces cerevisiae. Genetics* **123,** 69–80 (1989).
20. Jeffreys, A. J. & May, C. A. Intense and highly localized gene conversion activity in human meiotic crossover hot spots. *Nature Genet.* **36,** 151–156 (2004).
21. Terasawa, M. *et al.* Meiotic recombination-related DNA synthesis and its implications for cross-over and non-cross-over recombinant formation. *Proc. Natl Acad. Sci. USA* **104,** 5965–5970 (2007).
22. Merker, J. D., Dominska, M. & Petes, T. D. Patterns of heteroduplex formation associated with the initiation of meiotic recombination in the yeast *Saccharomyces cerevisiae. Genetics* **165,** 47–63 (2003).
23. Lichten, M. & Goldman, A. S. Meiotic recombination hotspots. *Annu. Rev. Genet.* **29,** 423–444 (1995).
24. Martini, E., Diaz, R. L., Hunter, N. & Keeney, S. Crossover homeostasis in yeast meiosis. *Cell* **126,** 285–295 (2006).
25. Ardlie, K. *et al.* Lower-than-expected linkage disequilibrium between tightly linked markers in humans suggests a role for gene conversion. *Am. J. Hum. Genet.* **69,** 582–589 (2001).
26. Wall, J. D. Close look at gene conversion hot spots. *Nature Genet.* **36,** 114–115 (2004).
27. Primig, M. *et al.* The core meiotic transcriptome in budding yeasts. *Nature Genet.* **26,** 415–423 (2000).
28. Petes, T. D. Meiotic recombination hot spots and cold spots. *Nature Rev. Genet.* **2,** 360–369 (2001).
29. Ross-Macdonald, P. & Roeder, G. S. Mutation of a meiosis-specific MutS homolog decreases crossing over but not mismatch correction. *Cell* **79,** 1069–1080 (1994).
30. Kunz, C. & Schar, P. Meiotic recombination: sealing the partnership at the junction. *Curr. Biol.* **14,** R962–R964 (2004).
31. Borner, G. V., Kleckner, N. & Hunter, N. Crossover/noncrossover differentiation, synaptonemal complex formation, and regulatory surveillance at the leptotene/zygotene transition of meiosis. *Cell* **117,** 29–45 (2004).
32. Schwacha, A. & Kleckner, N. Interhomolog bias during meiotic recombination: meiotic functions promote a highly differentiated interhomolog-only pathway. *Cell* **90,** 1123–1135 (1997).
33. Hillers, K. J. Crossover interference. *Curr. Biol.* **14,** R1036–R1037 (2004).
34. Malkova, A. *et al.* Gene conversion and crossing over along the 405-kb left arm of *Saccharomyces cerevisiae* chromosome VII. *Genetics* **168,** 49–63 (2004).
35. Oh, S. D. *et al.* BLM ortholog, Sgs1, prevents aberrant crossing-over by suppressing formation of multichromatid joint molecules. *Cell* **130,** 259–272 (2007).
36. Hurles, M. How homologous recombination generates a mutable genome. *Hum. Genomics* **2,** 179–186 (2005).
37. Birdsell, J. A. Integrating genomics, bioinformatics, and classical genetics to study the effects of recombination on genome evolution. *Mol. Biol. Evol.* **19,** 1181–1197 (2002).
38. Lindahl, T. Instability and decay of the primary structure of DNA. *Nature* **362,** 709–715 (1993).
39. Kleckner, N. *et al.* A mechanical basis for chromosome function. *Proc. Natl Acad. Sci. USA* **101,** 12592–12597 (2004).
40. Borts, R. H. & Haber, J. E. Meiotic recombination in yeast: alteration by multiple heterozygosities. *Science* **237,** 1459–1465 (1987).
41. Chen, W. & Jinks-Robertson, S. The role of the mismatch repair machinery in regulating mitotic and meiotic recombination between diverged sequences in yeast. *Genetics* **151,** 1299–1313 (1999).
42. Weiner, B. M. & Kleckner, N. Chromosome pairing via multiple interstitial interactions before and during meiosis in yeast. *Cell* **77,** 977–991 (1994).
43. Rockmill, B., Sym, M., Scherthan, H. & Roeder, G. S. Roles for two RecA homologs in promoting meiotic chromosome synapsis. *Genes Dev.* **9,** 2684–2695 (1995).
44. David, L. *et al.* A high-resolution map of transcription in the yeast genome. *Proc. Natl Acad. Sci. USA* **103,** 5320–5325 (2006).
45. Huber, W. *et al.* Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics* **18** (suppl. 1), S96–S104 (2002).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank S. Clauder-Münster, M. Granovskaia, M. Sieber, T. Bähr-Ivacevic, M. Nguyen, V. Benes, Z. Xu, L. Ettwiller, P. McGettigan and the EMBL Genomics Core Facility for technical help; M. Knop for discussions; A. Akhtar, A. Ladurner, A. De Luna and M. Knop for critical comments on the manuscript; E. Louis, R. Durbin and D. Carter for making data from the Saccharomyces Genome Resequencing Project available; and the contributors to the Bioconductor (http://www.bioconductor.org) and R (http://www.R-project.org) projects for making their software available. This work was supported by grants to L.M.S. from the National Institutes of Health and the Deutsche Forschungsgemeinschaft, and to W.H. from the Human Frontier Science Program; and by a Darwin Trust's Jeff Shell Scholarship awarded to E.M.

Author Information Raw data are available from ArrayExpress (http://www.ebi.ac.uk/arrayexpress) under accession number E-TABM-470. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to L.M.S. (larsms@embl.de).

METHODS

Strains and media. The hybrid strain S288c/YJM789 was generated by crossing S96, isogenic to S288c (ref. 17), with YJM789[16]. To generate the homozygous *msh4*Δ and *mms4*Δ hybrid strains, the corresponding gene was replaced by a *nat*MX4 or *kan*MX4 drug-resistance marker[46] in each of the haploid parental strains, which were then crossed. Sporulation was induced by transferring overnight cultures from liquid YEPD to 2% potassium acetate[43].

DNA extraction and hybridization. Fifty-one complete wild-type and five complete *msh4* tetrads were dissected for genotyping. Twenty *mms4* viable spores were also selected, as were thirteen S96 and twelve YJM789 parentals. Spores were allowed to grow in YEPD solid medium and then streaked out to obtain single colonies, only one of which was used for genotyping. Note that starting from a single colony prevented analysis of heterozygosities within a single spore arising from post-meiotic segregation. Genomic DNA was extracted from an overnight, 100 ml, YEPD, saturated culture of each spore using a QIAGEN Genomic-tip according to manufacturer's protocol. Ten micrograms of genomic DNA were fragmented, biotin-labelled and hybridized to a custom Affymetrix microarray, as described previously[44]. All probes were remapped (exonerate[47]) to the S288c genome and the aligned portion of the YJM789 genome[48]. Only probes with one exact match (25 matching bases) and no near matches (22 to 24) were retained, yielding 287,000 S288c-specific probes, 112,000 YJM789-specific probes, and 2.37 million probes interrogating non-polymorphic sequence.

Genotyping. Fluorescence intensities were normalized with *vsn*[45,49]. SNPs, insertions and deletions were identified using the S288c/YJM789 alignment[48], and for each polymorphism, a probe set was formed from probes interrogating the position(s) involved. Nearby polymorphisms producing identical probe sets were treated as a single marker. Genotype labels were available for parental data, so to genotype segregants, semi-supervised clustering was applied to the combined parental and segregant data. For each probe set, a two-component gaussian mixture model—with fixed mixture proportions (0.5) but distinct covariance matrices—was fit using the EM algorithm. For the small fraction of probe sets with >10 probes (probe sets interrogating large indels), principal components dimension reduction ($d = 10$) was applied first. Segregant genotypes were assigned using posterior probability of class membership. For *mms4*, genotypes were assigned in a supervised fashion, using the distributions previously estimated from the wild-type and *msh4* data.

Filtering of genotype calls. We deliberately opted for a high no-call rate with fewer errors, to reduce the chance of spurious short non-crossovers. Five wild-type and two *mms4* arrays exhibiting excessive genotype switching and large Mahalanobis residuals were set aside. A small fraction (0.7%) of probe sets exhibiting >2 classes—probably due to cross-hybridization with unlinked loci—were discarded (Supplementary Fig. 2c). Misclassification rates were estimated using inferred mixture distributions, and probe sets (4.6%) for which this estimate exceeded 1% were also discarded (Supplementary Fig. 2b). For retained probe sets, individual calls (4.9%) were discarded if the posterior probability of assigned class membership was too far from 1, or if the Mahalanobis residual was large (Supplementary Fig. 2a). In two sequencing validation data sets covering ~60 kb—one focused on calls from 16 different wild-type spores, and another, on two regions of an *mms4* segregant exhibiting frequent genotype switching—100% of filtered genotype calls were confirmed.

Recombination event annotation. After collecting genotype data into tetrads, genotype change points were grouped by proximity. Most cases were simple: pairs of changes isolated from all other changes were called non-crossovers if they involved one spore, or crossovers if they involved two (Supplementary Fig. 3a). A fraction of cases, however, were more complex, admitting several distinct interpretations. To treat such cases systematically, cutoff-based rules reflecting basic assumptions about the recombination process were used. See Supplementary Information Section 1 for details. Importantly, we explored a variety of plausible alternative annotation sets, and found no qualitative change in our main results.

Conversion tract length. Tract size estimates obtained using midpoints of flanking intermarker intervals were used for most calculations (see Supplementary Information Section 3). Where indicated, we also computed lower and upper bounds, using the regions spanned by converted markers (minimal), and delimited by the two nearest unconverted markers (maximal)[2]. For summary statistics, we combined simple (Supplementary Fig. 3a) and complex (Supplementary Fig. 3b, c) conversion tracts.

Event rate adjustment. Intermarker interval size affects the probability of involvement in recombination events. To adjust for this, we used a semi-parametric statistical model (Supplementary Information) to relate size to the probabilities of (1) involvement in and (2) detection of recombination events. The model's extension length distribution was estimated empirically. Given this estimate, we then counted recombination events overlapping each intermarker interval, and estimated remaining parameters by Poisson regression.

Defining hotspots. Using model parameter estimates, expected crossover and non-crossover counts were computed for each intermarker interval under a null hypothesis of rate homogeneity. We identified three types of hotspots: crossover, non-crossover and overall recombination events. To identify crossover hotspots, we performed a one-tailed test ($\alpha = 0.001$) using the Poisson distribution and the expected crossover counts. Hot intermarker intervals separated by <500 bp were merged. Non-crossover and overall recombination hotspots were identified similarly. Note that the three types of hotspots are statistically related, but crossover and non-crossover hotspot counts need not sum to the overall count.

Crossover/non-crossover bias testing. To assess crossover/non-crossover bias, we used expected crossover and non-crossover counts to compute expected crossover fractions. Conditioning on the observed number of events overlapping each interval, we then compared observed and expected counts using two one-tailed binomial distribution tests. The resulting *P*-values correspond to either an excess or deficiency of crossovers. Despite the large sample size, individual intermarker intervals were rarely involved in >10 events, so we chose to treat crossover/non-crossover bias *P*-values collectively rather than individually. We simulated data ($B = 2000$) under the same binomial distributions used for *P*-value calculations—conditioning on observed counts so that rate inhomogeneity across the genome was preserved—and examined (1) the average number of simulated *P*-values falling below 0.10, and (2) the average total size of intermarker intervals associated with such *P*-values. The former permitted estimation of false discovery rate, and the latter, estimation of the total size of intermarker intervals associated with true crossover/non-crossover bias.

46. Goldstein, A. L. & McCusker, J. H. Three new dominant drug resistance cassettes for gene disruption in *Saccharomyces cerevisiae*. *Yeast* **15**, 1541–1553 (1999).
47. Slater, G. S. & Birney, E. Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics* **6**, 31 (2005).
48. Wei, W. *et al.* Genome sequencing and comparative analysis of *Saccharomyces cerevisiae* strain YJM789. *Proc. Natl Acad. Sci. USA* **104**, 12825–12830 (2007).
49. Gentleman, R. C. *et al.* Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* **5**, R80 (2004).