WHERE THE CROSSOVERS ARE: RECOMBINATION DISTRIBUTIONS IN MAMMALS

Liisa Kauppi[*‡], *Alec J. Jeffreys*[*] *and Scott Keeney*[‡]

Until recently, recombination studies in humans and mice had identified only a few anecdotal examples of crossover hot spots. Recently, the pace of discovery has accelerated. In every genomic segment that has been examined at sufficiently high resolution, recombination events have a punctate recombination distribution: they are clustered within small (1–2-kb) regions that are surrounded by large stretches of recombinationally suppressed DNA. Here, we review progress in understanding the distribution of mammalian recombination events, tie mammalian results together with informative studies in budding yeast and discuss the consequences of these findings for genome diversity and evolution.

HAPLOTYPE
A combination of alleles at different loci that is transmitted together from one generation to the next.

In most sexually reproducing organisms, crossover recombination events provide physical connections between homologous chromosomes. These connections work in conjunction with sister-chromatid cohesion to ensure proper chromosome segregation during the first meiotic division[1,2]. The formation of crossovers also has the effect of breaking up HAPLOTYPES, therefore increasing genetic diversity from one generation to the next. Because of this effect, meiotic recombination exerts a profound influence on genome diversity and evolution. Studies in many organisms demonstrate that recombination events are more likely to occur in some parts of the genome than in others. Understanding this nonrandom distribution has important implications for explaining genome plasticity and for our ability to map disease-related loci.

In the past few years, our understanding of the distribution of recombination events in humans has drastically changed. Recent studies strongly support the view that most recombination events occur at highly localized hot spots, whereas the bulk of the DNA is 'cold'. In this review, we discuss low-resolution experiments that first documented the existence of recombination hot spots in humans and mice, and review more recent high-resolution studies that reveal the highly localized clustering of recombination events. Because the emerging pattern in mammals seems similar to that documented in budding yeast (*Saccharomyces cerevisiae*), we review molecular details in this model organism that are informative for understanding the mechanisms that dictate recombination distributions in mammals. We then discuss the implications that this highly punctuated recombination distribution has for understanding genome architecture.

Several experimental approaches have been used to characterize the frequency and spatial distribution of meiotic recombination events. All show that recombination in mammals is nonrandomly distributed, at various levels of spatial resolution (BOX 1). The classic tool in gene mapping is to examine parent-to-offspring transmission of alleles at several loci. This so-called pedigree approach has identified localized genomic regions, at the 1–10-kb scale, in which recombination events seem to cluster and that historically were referred to as recombination hot spots. However, this approach to measuring meiotic recombination rates is limited by difficulties in detecting low-frequency events and by the variable density of polymorphic markers that are used to map crossover positions. Although recombination events are clearly distributed nonrandomly across the genome, it has been difficult to pinpoint precisely where they occur. This issue is important for understanding the

[*]*Department of Genetics, University of Leicester, Leicester LE1 7RH, UK.*
[‡]*Molecular Biology Program, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, NewYork, New York 10021, USA. Correspondence to L.K., A.J.J. or S.K.*
e-mails: kauppil@mskcc.org; ajj@le.ac.uk; keeneys@mskcc.org
doi:10.1038/nrg1346

Box 1 | **Crossover distributions in mammals — an historical perspective**

Genetic versus physical distances

Crossover activity in mammals is usually expressed in CENTIMORGANS (cM) per megabase pair (Mb). Comparisons of genetic distances with physical distances reveal, at the megabase level, long DNA domains of elevated or suppressed recombination that alternate along the length of each chromosome (see, for example, REFS 61,91). Because estimates of genetic distance are based on the number of offspring observed with recombinant genotypes, the resolution of genetic maps is typically limited to 0.1–1 Mb, so little is known about how crossovers are distributed within 'hot' and 'cold' domains. Linkage maps also reveal that females have a higher recombination rate than males, with the female:male genetic-map-length ratio being approximately 1.6:1 (REFS 40,92). Female:male ratios are typically highest around centromeres and lowest near telomeres[40,60,92,93].

Cytological studies

Crossover distributions can also be examined cytologically by observing CHIASMATA or through immunofluorescent detection of chromosome-associated protein complexes that are involved in recombination. These cytological approaches reveal that most chromosomes in humans and mice undergo one or two crossovers per chromosome or chromosome arm in each meiosis, with species-specific and sex-specific variations in position along chromosomes[88,94–100]. However, these studies do not provide a high-resolution view of recombination distributions because of low sample sizes and microscopic resolution.

Pedigree and linkage-disequilibrium analysis

The average crossover frequency is 0.5 cM Mb^{-1} in mice[101] and 1.1 cM Mb^{-1} in humans[40]. This means that, within a 1-kb interval, crossovers in humans will occur on average in only 1 gamete in 90,000, and therefore only the most intense hot spots will be detected by pedigree analysis. An alternative approach is to use LINKAGE DISEQUILIBRIUM (LD) measures of the association of alleles at separate loci to estimate the amount of historical recombination that has occurred between the loci (BOX 2), and therefore to infer the presence and location of crossover hot spots[102]. Pedigree approaches, often guided by LD analysis, have identified several putative crossover hot spots in the human genome, although the number of crossovers that are detected in each region are low (for example, the β-globin gene region[103,104], the *PGM1* gene[105], the *ADA* gene[106], the *LRP5* gene[107] and three regions of the major histocompatibility complex (MHC)[87,108]). In the mouse MHC, 4 regions that are active in recombination have been defined at a resolution of <10 kb[109–116]. In fact, the first direct evidence for crossover hot spots in any mammal came from mouse MHC studies[113]. Also, the elegant work of Shiroishi and co-workers on the mouse *Lmp2* hot spot highlights the potential importance of *cis*-acting factors in mammalian recombination[114,117].

molecular mechanism behind nonrandom recombination distributions because our perspective is strongly influenced by the spatial resolution of the methods that are used to detect recombination events[3].

Characterizing mammalian hot spots

Progress in this field has recently accelerated owing to two technical advances: the development of efficient re-sequencing methods for SNP discovery, which have greatly increased marker density[4], and, more importantly, the advent of single-molecule methods that allow recombinant DNA molecules to be recovered directly from sperm DNA. There are two variations on the latter theme. One relies on the PCR amplification of haploid genomes of individual flow-sorted sperm[5] (reviewed in REF. 6). The other approach uses allele-specific PCR to amplify recombinant DNA molecules directly from large batches of sperm DNA[7] (FIG. 1); this approach allows many thousands of meiotic products from a single individual to be screened, and provides the means to detect and quantify even very low-frequency recombination events (10^{-6} or less per sperm). Sperm analysis can be coupled to linkage disequilibrium (LD) surveys (BOX 2; FIG. 2), using the latter as a guide to the location of putative hot spots. Crossover-breakpoint mapping and the resolution of hot spots is limited only by the density of markers within the target region; for humans, resolution to less than 0.5 kb is usually possible by choosing individuals with appropriate genotypes. Of course, the obvious drawback of sperm PCR techniques is that they can only be used to analyse meiotic crossovers in males. Family data therefore remain essential for analysing female recombination rates in humans.

Analysis of flow-sorted single sperm allowed the localization of a hot spot to an 11-kb region in the human β-globin locus[8] and the demonstration of regional heterogeneity in crossover rates along the human major histocompatibility complex (MHC)[9]. However, further narrowing of hot spots in both studies was limited by the number of crossovers that were recovered and the density of markers that were analysed. Other studies using the batch PCR approach have been able to identify and characterize several crossover hot spots in humans and mice at even higher resolution.

High-resolution analysis of human crossover hot spots. The first hot spot to be characterized by sperm DNA analysis lies next to the MINISATELLITE MS32 on chromosome 1 (REF. 7). This hot spot is highly localized (it is ~1.5-kb wide) and is centred 200 bp away from the minisatellite, with crossovers extending into the array of repeat DNA. The peak crossover rate is 50 times the male genome average.

High-resolution LD mapping across 210 kb of the MHC class II region revealed several localized regions of LD breakdown (see FIG. 2b for an example). Allele-specific PCR analyses of sperm DNA identified 6 localized crossover hot spots (some of them in clusters that span

CENTIMORGAN
(cM). A measure of genetic distance. 1 cM corresponds to a 1% frequency of recombinant progeny.

CHIASMA
(pl. chiasmata). A cytologically visible physical connection between homologous chromosomes that corresponds to the position of a meiotic crossover.

LINKAGE DISEQUILIBRIUM
(LD). A measure of whether alleles at two loci coexist in a population in a nonrandom fashion. Alleles that are in LD are found together on the same haplotype more often than would be expected by chance.

MINISATELLITES
Regions of DNA in which repeat units of 6–100 bp are arranged in tandem arrays that are 0.5–30 kb in length.

less than 10 kb) in these regions of LD breakdown[10,11] (FIG. 3). All of these hot spots show clustering of crossovers within a narrow (1–2-kb) region, and are surrounded by DNA sequences that are virtually inert for recombination (FIG. 3). Peak recombination activities at these MHC hot spots vary markedly, from 0.3 cM Mb^{-1} to 110 cM Mb^{-1}. The peak activity at the weakest hot spot is actually below the genome average frequency, but the hot spot is nevertheless evident given the extremely low recombination rate in surrounding sequences, which is estimated to be approximately 0.05 cM Mb^{-1}(REF. 11). For this reason, it is useful to define hot spots as localized sites of recombination rather than as sites that show recombination frequencies above the genome average. With this concept of a hot spot, it is worth noting that there is no need to invoke the existence of specific chromosomal sites with opposite properties — that is, 'cold spots'. Indeed, most of the DNA sequences in the genome are cold.

The 2.6-Mb pseudoautosomal pairing region (*PAR1*) at the subtelomeric region of Xp/Yp is a site of obligatory crossover in male meiosis. High-resolution LD mapping around the *SHOX* gene in *PAR1* showed no evidence of clear LD block structure[12], which is consistent with the high rate of crossing over in males (20-fold above the genome average)[13]. However, a sperm-crossover assay targeted to 1 of the regions of free association in the *SHOX* gene revealed a highly localized (2-kb-wide) hot spot, with peak crossover activities of 140–370 cM Mb^{-1} in the 3 men tested[12].

High-resolution analysis of mouse recombination hot spots. Recently, two mouse MHC hot spots have been subjected to similar crossover analyses in sperm[14,15]. The *Lmp2* (now re-named *Psmb9*) gene region contains a hot spot that is similar in width (1.8 kb) to all human hot spots and with an intense peak activity of 1,300 cM Mb^{-1} (REF. 14). At the *Eb* hot spot, which is again of similar width, the influence of haplotype on male meiotic recombination has been analysed in detail[15]. Analysis of sperm crossovers from mice that are heterozygous for different haplotypes showed large variation in recombination rates that correlated inversely with sequence divergence between haplotypes in and near the hot spot, although the position of the hot spot remained unchanged.

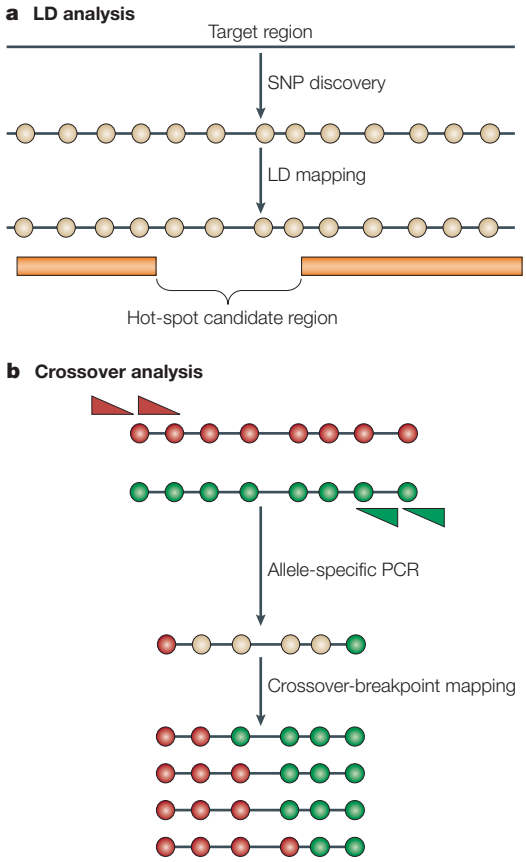

Figure 1 | **The PCR approach for recovering crossover molecules directly from sperm DNA. a** | The first step is SNP discovery (SNPs are shown as beige circles) and linkage-disequilibrium (LD) mapping, which allow the identification of LD BLOCKS (orange boxes) and regions where LD breaks down (gap between orange boxes). Such regions of LD breakdown are putative recombination hot spots. **b** | For putative recombination-hot-spot regions, two rounds of allele-specific PCR are carried out on batches of sperm DNA from an individual with multiple SNP heterozygosities (red and green circles). Allele-specific primers are shown as red and green arrowheads. Recombinant sperm DNA molecules are selectively amplified, and the location of crossover breakpoints can be subsequently mapped by typing internal SNPs.

Properties of crossover hot spots

The studies described above reveal several features of recombination hot spots. Crossovers at all mammalian hot spots identified so far cluster within small (1–2-kb) regions (FIG. 3b); these common widths imply that there are common mechanisms of recombination initiation and resolution at different hot spots. Crossover breakpoints are distributed symmetrically (that is, they have a 'normal distribution'; see FIG. 3c) within these narrow regions. Peak recombination rates at human hot spots vary from below the genome average at 1 hot spot in the MHC to 370-fold above the genome average at the *SHOX* hot spot. Hot spots are located in different genomic environments and share no obvious sequence similarities with one another. Interestingly, some hot spots are located in repeat elements (such as *Alu* elements and *LINEs*), which raises the issue of how ectopic recombination events are prevented in the human genome. Thuriaux pointed out that recombination might be associated with genes, because in various organisms the overall amount of recombination correlates with gene content rather than the physical size of the genome[16]. However, there does not seem to be a strict relationship of this sort for mammalian hot spots. The human and mouse MHC hot spots and the *SHOX* hot spot are all located at or near genes (within ~8 kb), but, as the MHC is a gene-rich region and *SHOX* was pre-selected for analysis, this association is not necessarily meaningful. Only one of these recombination hot

LD BLOCK
A DNA segment within which markers are in significant linkage disequilibrium with each other, which implies that there is low recombination activity within the block.

Box 2 | D′ as a measure of linkage disequilibrium

Linkage disequilibrium (LD) is a population-based parameter and reflects recombination in all of the generations since the rise of the individual polymorphic alleles that are under study. So, LD analyses do not suffer from the same limitations as do pedigree analyses with respect to detecting low-frequency events. However, patterns of LD can be influenced by the age of crossover events, GENE CONVERSION, selective advantages or disadvantages of single or linked alleles, and demographic history, such as founder effects and ADMIXTURE[84,118,119].

In principle, if strong LD is found between two loci, little if any historical recombination has occurred between them. Conversely, little or no LD implies that there has been historical recombination activity. There are various LD measures for characterizing the statistical association between alleles at different loci (see, for example, REF. 120). A simple measure of LD is $D_{ij} = x_{ij} - p_i q_j$, where x_{ij} is the observed frequency of haplotype A_iB_j, p_i and q_j are the frequencies of alleles A_i and B_j at loci A and B, respectively, and the expected frequency of gamete A_iB_j is p_iq_j assuming no statistical association between the alleles. The range of this measure is a function of the allele frequencies, and is therefore not ideal.

Consider 2 SNPs in which SNP1 carries alleles C and G (at a population frequency of 0.90 and 0.10, respectively), and SNP2 carries alleles A and T (at a population frequency of 0.55 and 0.45, respectively). There are then 4 possible haplotypes for these 2 SNPs (CA, CT, GA and GT); assuming no statistical association, we expect to see these haplotypes at frequencies of 0.495, 0.405, 0.055 and 0.045, respectively. If the observed haplotype frequencies in the population are 0.450, 0.450, 0.100 and 0, then $D_{CA} = 0.450 - 0.495 = -0.045$.

The normalized measure D′ (REF. 121) is perhaps the most widely used, as its range is the same for all allele frequencies. It does tend to be over-inflated when sample size is small, however[122]. D′ is defined as follows: $D'_{ij} = D_{ij}/D_{max}$, where $D_{max} = \min[p_iq_j, (1-p_i)(1-q_j)]$ when $D_{ij} < 0$, or $D_{max} = \min[p_i(1-q_j), (1-p_i)q_j]$ when $D_{ij} > 0$. For the above example of the CA haplotype, $D_{max} = 0.045$, which is the lesser value of $0.9 \times 0.55 = 0.495$ and $(1-0.9)(1-0.55) = 0.045$. Therefore, $D'_{CA} = D_{CA}/D_{CA\,max} = -0.045/0.045 = -1$. D′ can range from –1 to 1; whether the sign is positive or negative depends on the arbitrary labeling of alleles, and therefore, the absolute value |D′| is mostly used. If |D′| = 1, this is referred to as complete LD, and occurs if 3 haplotypes at most are present in the population. The presence of three haplotypes can be explained by mutation alone, and therefore, there is no evidence for historical recombination. If |D′| = 0, this is referred to as no LD, or free association, and implies that there has been extensive recombination between the markers, or (in theory) recurrent mutation. LD plots can be generated to visualize the extent of LD (see FIG. 2). The statistical significance of each LD value can also be incorporated into the plots.

spots (the weakest MHC hot spot) lies in a promoter region, in contrast to budding yeast in which promoter sequences are preferential sites for hot spots[17]. Furthermore, some human MHC genes probably do not have adjacent hot spots[11], and the MS32 hot spot lies in the middle of a 77-kb non-coding DNA segment, again arguing against a 1:1 correlation between genes and hot spots.

Crossover hot spots are also gene-conversion hot spots. There is evidence that recombination hot spots in mice and humans are sites of initiation and resolution of both crossovers and non-crossover gene conversions[7,14,18]. At *Lmp2* in the mouse MHC, conversions that involve a single chosen marker in the hot spot revealed short conversion tracts of <540 bp[14]. More recently, an ENRICHMENT-BASED ASSAY[19] has been developed to capture conversion events that involve even single SNP sites[18]. Analysis of three human crossover hot spots showed that they are also extremely localized conversion hot spots, with peak crossover and conversion activity mapping to the same location at the centre of the hot spot. The number of non-crossover conversion events dominated the number of crossovers by a 4:1 to 15:1 ratio, conversion tracts were short (<300 bp on average) and initiation events were inferred to cluster within a small zone of ~400 bp[18]. The existence of initiating lesions in hot spots has been confirmed by Qin *et al.*[20], who developed a PCR-based approach to reveal the presence of meiosis-specific strand breaks with 3′-OH ends in the mouse *Ea* recombination hot spot.

GENE CONVERSION
The non-reciprocal transfer of information between homologous DNA sequences as a consequence of heteroduplex formation during recombination, which is followed by repair of mismatches in the heteroduplex.

ADMIXTURE
The mixing of two genetically differentiated populations.

ENRICHMENT-BASED ASSAY
An allele-specific hybridization method for enriching DNA molecules that carry certain allelic combinations from bulk genomic DNA.

The how and why of nonrandom distributions

The existence of localized recombination hot spots in mammals raises the question of how this punctate crossover distribution arises. A large body of work, primarily in budding yeast, provides a direct insight into the molecular mechanisms that are behind hot-spot activity.

The DNA events of meiotic recombination. The meiotic recombination pathway in budding yeast has at its heart the formation and subsequent repair of double-strand breaks (DSBs; reviewed in REFS 21–23; see FIG. 4). The repair of any given meiotic DSB results in either the reciprocal exchange of the chromosome arms that flank the break (a crossover) or no exchange of flanking arms (a non-crossover or parental configuration). It is highly probable that DSBs initiate meiotic recombination in mammals and other organisms as well, not least because the central player in DSB formation, Spo11, is structurally and functionally conserved[24,25].

Crossover distributions reflect double-strand-break distributions. In budding yeast, the single most important determinant of the crossover distribution is the distribution of the DSBs that initiate meiotic recombination[17,26]. As for crossover distributions in humans (BOX 1), DSB distributions in budding yeast show many levels of spatial organization (FIG. 5). There are hot (DSB-rich) and cold (DSB-poor) domains that span large portions of chromosomes (in the order of 100 kb)[17,27,28]. Within these large domains, DSBs form preferentially at localized hot spots[26,29,30]. Finally, within any given hot spot,

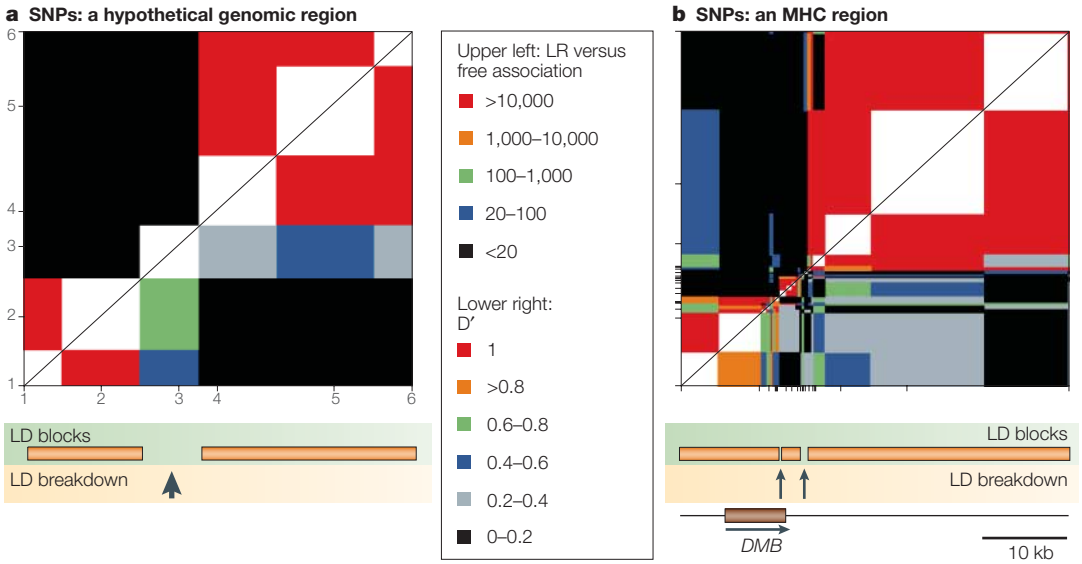

Figure 2 | **Graphical presentation of linkage-disequilibrium data. a** | Linkage disequilibrium (LD) in a hypothetical genomic region. In this example, there are 6 SNPs, numbered 1–6 (labelled below and to the left of the plot). LD is calculated for each pair of SNPs from genotypes of a panel of individuals. Each pairwise comparison is plotted as a rectangle centred on each SNP, and extending half way to adjacent markers. The |D′| values are plotted by colour code (see key) below the diagonal. The statistical significance of each |D′| value can be deduced from the LIKELIHOOD RATIO (LR) versus free association, which is shown above the diagonal. In this example, |D′| = 1 between markers 1 and 2, indicating complete LD. This value is supported by a high (>10,000) LR. By contrast, the |D′| value between markers 3 and 5 is 0.4–0.6, but this value is accompanied by poor odds (LR versus free association <20). A region of LD breakdown is apparent between markers 2 and 4 (arrow). **b** | LD in a portion of the human major histocompatibility complex (MHC). Patterns of LD near the *DMB* gene are shown (data from REF. 11). Only SNPs with a minor allele frequency ≥0.15 have been included. LD blocks are identified visually and are shown as orange boxes. There are two regions of LD breakdown (vertical arrows). Adapted with permission from REF. 11 © (2001) Macmillan Magazines Ltd.

specific phosphodiester bonds are cleaved at different frequencies[31–34]. This last level of spatial resolution is dictated at least in part by Spo11 itself because certain *spo11* missense mutations alter the distribution of cleavage sites within hot spots[35].

Detailed reviews of fungal hot spots have been provided elsewhere[3,36–39]; here we highlight those features that will probably provide insight into the molecular mechanisms that are behind nonrandom crossover distributions in mammals. It should be kept in mind that recombination per unit of physical distance in budding yeast (0.37 cM kb^{-1} for 12×10^6 bp total) is much higher than in humans (0.001 cM kb^{-1} for ~3×10^9 bp total)[38,40]. Therefore, individual hot spots in budding yeast are much more active in recombination: a 1–10% crossover frequency is not unusual for budding-yeast hot spots[3], whereas a 0.3% crossover frequency is the 'record' in humans so far (at the *SHOX* hot spot[12]).

The presence of DSB-hot and DSB-cold chromosomal domains has been amply documented across the budding-yeast genome[17,27,28]. On chromosome 3, for example, regions near the telomeres show few or no prominent DSBs. There are also few DSBs near the centromere, whereas interstitial regions in each of the chromosome arms show a high frequency of DSBs (FIG. 5a). This bias in the distribution of DSBs represents a large-scale position effect, because a recombination reporter that is placed at different positions takes on the properties of its location: insertions into cold regions give low DSB levels and insertions into hot regions give high DSB levels[41,42]. This phenomenon supports a model in which higher-order structures and/or chromosome dynamics contribute to the control of the frequency of recombination-initiation events, but at present the molecular basis of this process is not known. DSB-hot domains in budding yeast correspond to regions with relatively G+C-rich sequences[17,28,43]. Importantly, crossovers in mammals also form preferentially in chromosomal domains that correspond to similar G+C-rich ISOCHORES or 'R-BANDS'[40,44,45] (although at the local level, there is no evidence that mammalian hot spots and their immediate environments are unusually G+C rich). It therefore seems probable that hot and cold chromosomal domains are specified by similar processes in humans and in budding yeast.

Within these larger domains, DSBs in budding yeast form within localized hot spots. If appropriately analysed by Southern blotting of restriction digests of whole genomic DNA, these hot spots appear as clusters of frequently cleaved sequences, which typically span ~70–250 bp and are surrounded by sequences in which breaks form rarely, if at all[17,31–34] (see FIG. 5c). The punctate distribution of DSBs at this spatial resolution is highly reminiscent of the punctate distribution of the recombination events that were described above for the MHC and other regions in mammals. The similarity of these patterns in turn leads, in a straightforward way, to the idea that mammalian hot spots correspond to the sites at which DSBs form preferentially.

LIKELIHOOD RATIO
The relative likelihood of obtaining observed experimental data under two different models.

ISOCHORE
A region of genomic DNA sequence in which G+C compositions are relatively uniform.

R-BANDS
A chromosome banding pattern that is produced with various staining procedures; it is reciprocal to the pattern that is produced by Giemsa staining (G-Bands).

The factors that determine whether a given DNA sequence will be a DSB hot spot in budding yeast are not completely understood, but some general rules have emerged. One important determinant is the chromatin structure. Essentially, all known DSB hot spots are nuclease-hypersensitive in both mitotic and meiotic chromatin, and the results of several experiments indicate that an open chromatin configuration is necessary for DSB formation[26,29,30,46,47]. But chromatin structure cannot be the sole arbiter because not all nuclease-hypersensitive sites are DSB hot spots and there is no correlation between the degree of nuclease hypersensitivity and the frequency of DSB formation[26,29,30,41,42,47]. Most naturally occurring hot spots in budding yeast (but not in mammals) lie in promoter regions[17,36]. However, transcription *per se* is not required for DSB formation, because deletion of the TATA box at *HIS4* markedly diminishes transcription without affecting the recombination frequency[48], and because several artificial hot spots have been created that seem not to be transcriptionally active[41,49–52]. The tendency of DSBs to occur in promoter regions might reflect the fact that promoters tend to have an open chromatin configuration. Alternatively, sequence-specific binding proteins (such as transcription factors) might be important for targeting the DSB machinery, either by inducing a local chromatin structure that is permissive for DSB formation, or by recruiting the recombination-initiation machinery by direct protein–protein interactions[48,53–55]. Petes and colleagues have argued that there are in fact three classes of recombination hot spot: those that require transcription factors for targeting recombination complexes, those that are transcription-factor-independent because they are always in a permissive state for DSB formation and those that are hot because of a high G+C content, without the necessity of being in an open chromatin configuration[28,38,55].

Other sources of nonrandomness. If the number of early forming chromosome-associated complexes of the strand-exchange proteins Rad51 and Dmc1 is used to estimate the total number of recombination events, it is apparent that meiotic cells make a large excess of DSBs relative to the number of crossovers that eventually form (see, for example, REFS 56,57). Because every DSB must be repaired, it follows that the excess breaks must be processed to give a non-crossover configuration. The cytological data in mammals indicate that non-crossovers outnumber crossovers by approximately 10:1

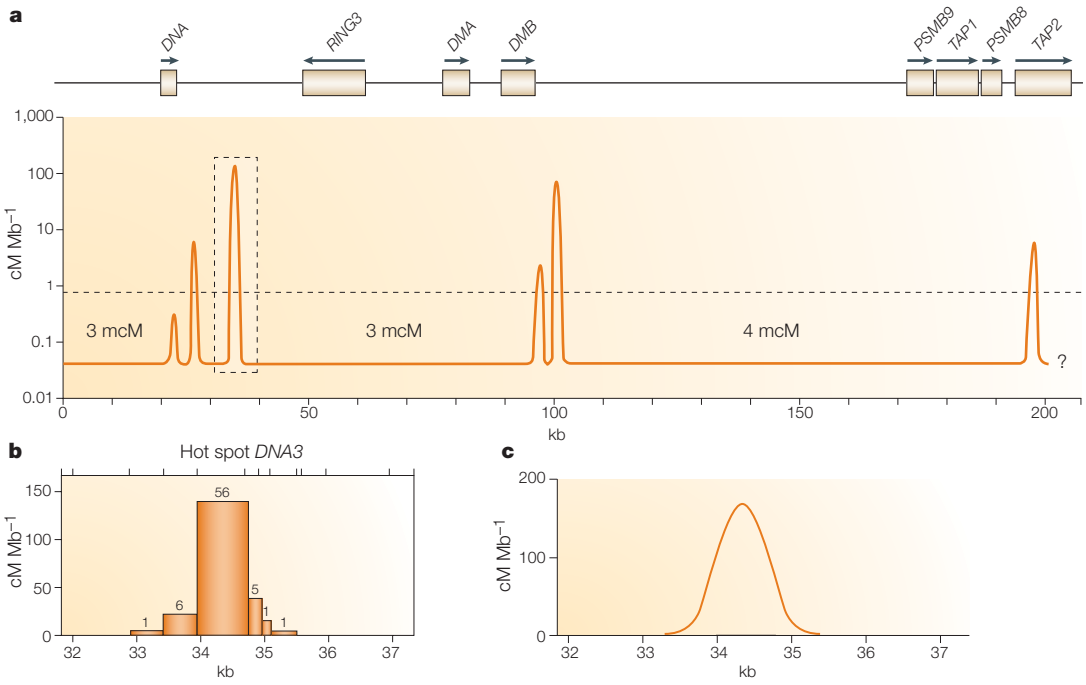

Figure 3 | **Highly localized recombination activity in the human major histocompatibility complex class II region.** The presence of peaks in crossover activity, which are embedded in 'cold' DNA, are reminiscent of recombination patterns in maize (for example, the 140-kb interval between the *anthocyanin1* and *shrunken2* genes, REF. 126), and might be typical of higher eukaryotes with complex genomes. **a** | Genes in the region are shown as boxes above the plot. The peaks in recombination activity represent the six crossover hot spots characterized by sperm analyses (see main text). The approximate recombination rate for inter-hot-spot DNA segments is given in millicentiMorgans (mcM, cMx10^{-3}). Mean male recombination rate (0.81 cM Mb^{-1}, REF. 40) is shown as a dashed line. The most active crossover hot spot, *DNA3* (shown in more detail in **b** and **c**), is indicated with a dashed box. **b** | Within the *DNA3* hot spot, crossovers cluster into a narrow (less than 2-kb) region. A total of 70 crossovers were recovered from 57,000 sperm DNA molecules from 1 man. Orange bars indicate the crossover activity in cM Mb^{-1} in each heterozygous SNP interval (SNPs shown as vertical bars above the plot). The number of crossovers that map to each interval is shown above the bars. **c** | Underlying morphology and crossover activity of the *DNA3* hot spot (based on the least-squares best-fit curve for the crossover data, assuming a normal distribution of crossover breakpoints). Peak crossover activity was determined from the mean rate in the six men tested. Adapted with permission from REF. 11 © (2001) Macmillan Magazines Ltd.

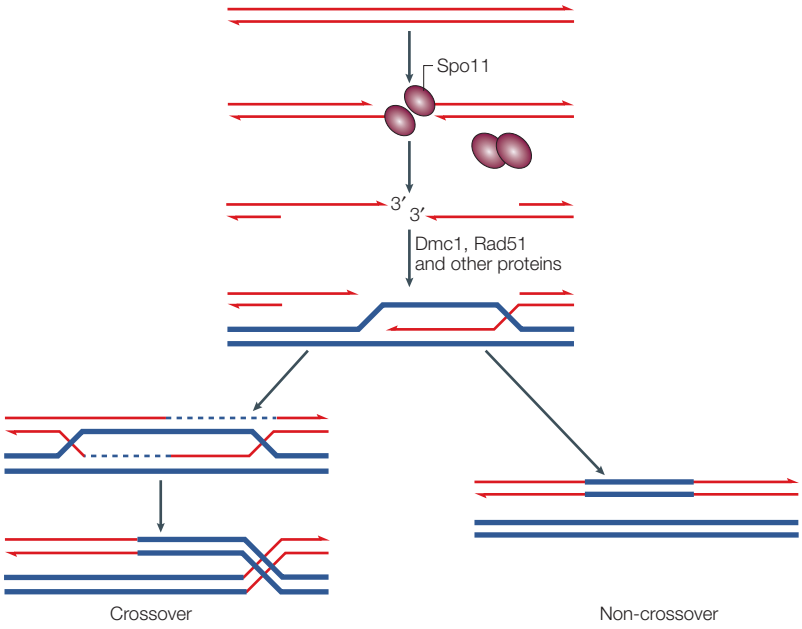

Figure 4 | **The meiotic recombination pathway in budding yeast.** Physical analysis of DNA isolated from meiotic cell cultures has defined the important steps along the recombination pathway, starting with the initiating double-strand breaks (DSBs)[49,127,128]. DSB formation is catalysed by Spo11, which seems to act through a topoisomerase-like reaction to generate a transient, covalent protein–DNA intermediate[129,130]. Spo11 is then removed from the DNA and the 5′-strand termini are nucleolytically resected to yield variable-length, 3′-single-strand tails[131,132]. In a series of reactions that are dependent on budding-yeast homologues of bacterial RecA[132,133], these tails undergo strand invasion of intact homologous duplexes, ultimately giving rise, in at least some cases, to double HOLLIDAY-JUNCTION intermediates, followed by mature recombinant products.

(REF. 58). This estimate applies to the genome as a whole; however, the bias in the outcome of DSB repair at any given hot spot or within a given chromosomal domain could differ substantially from this genome average. In principle, then, regional variation in crossover frequencies could reflect not only variation in DSB frequency, but also variation in how probable it is that a DSB will give rise to a crossover instead of a non-crossover outcome.

There are several potential examples of this idea. In budding yeast, the average crossover/non-crossover ratio is ~1:2 (measured as the fraction of gene conversions that are associated with the crossing over of flanking markers), but ratios at different loci have been reported to range from ~1:4 to ~2:1 (REF. 59). In mammals, the most striking possible example of this phenomenon is the obligate crossover within the *PAR1* region of the XY-chromosome pair. As judged by immunostaining for recombination proteins, DSB formation is distributed along the lengths of the X and Y chromosomes, forming in both the *PAR* and the non-homologous chromosome-arm regions[58]. The formation of a crossover specifically within *PAR1* in every male meiosis might therefore reflect a regional bias towards a crossover outcome rather than (or in addition to) an increase in DSB frequency relative to other parts of the genome (presumably, the DSBs outside the *PAR* are repaired by recombination between sister chromatids or between repeated sequences that are dispersed on the Y chromosome). Another possible example of region-specific bias in the crossover/non-crossover ratio is the tendency of crossovers to be relatively telomere-proximal in human males[60,61]. Alterations in the crossover/non-crossover ratio might also explain why some organisms show changes in the frequency of chiasmata that are formed adjacent to heterozygous configurations of large-scale chromosome rearrangements such as translocations or inversions (see, for example, REFS 62,63). Finally, crossover frequencies are suppressed when the recombining DNA molecules contain extensive sequence differences, as has been demonstrated in interspecific crosses in bacteria and budding yeast[64,65]. Although it is not understood precisely which stage in the recombination process is affected by sequence heterology, the involvement of the cellular mismatch-repair machinery in inhibiting recombination makes it probable that this effect occurs during or after strand exchange.

Why is the distribution of recombination events nonrandom? Because nonrandom crossover distributions are so widespread in evolution, the question arises as to what purpose is served by controlling the spatial organization of recombination events. One possibility is that the cell must balance two opposing needs: recombination is necessary to ensure proper chromosome segregation at the first meiotic division, but it must also be reined in to minimize the breaking up of favourable linkage groups and to foster genome stability. Perhaps controlling the spatial distribution of crossovers provides a way to achieve this balance. However, comparison of mouse and human MHC class II hot-spot distributions shows that although gene order in the mouse and human MHC is conserved, crossover hot-spot locations do not seem to be conserved (FIG. 6). This would indicate first, that hot-spot location is not dictated by 'linkage groups' of particular genes, or at least that the requirement for genes contained within a haplotype is different in the two species. Second, the comparison shows that recombination hot spots are relatively transient, with a life span that is at least shorter than the time since mouse–human divergence. Evidence from LD studies that the human β-globin hot spot is not active in rhesus macaques (or if it is, it has not left its imprint on haplotype diversity) is consistent with this[66].

Another possibility is that restrictions on crossover position facilitate the optimal mechanical/biochemical function of chiasmata in chromosome segregation. In the light of this, it is noteworthy that errors in chromosome segregation are correlated with unusual crossover positions in many cases of maternally derived trisomy 21 (REFS 67,68). What is 'optimal' might vary significantly from organism to organism and even between sexes in the same species or among chromosomes in the same cell.

Finally, deliberately inflicting DSBs on its own genome is a dangerous game for the organism to play: a single unrepaired break will probably be lethal, yet hundreds of breaks are made in each meiotic cell. Perhaps restrictions on the spatial organization of DSB formation promote the efficient repair of DSBs by ensuring that breaks form in regions of the genome that are conducive

HOLLIDAY JUNCTION
A point at which the strands of two double-stranded DNA molecules exchange partners, which occurs as an intermediate in crossing-over.

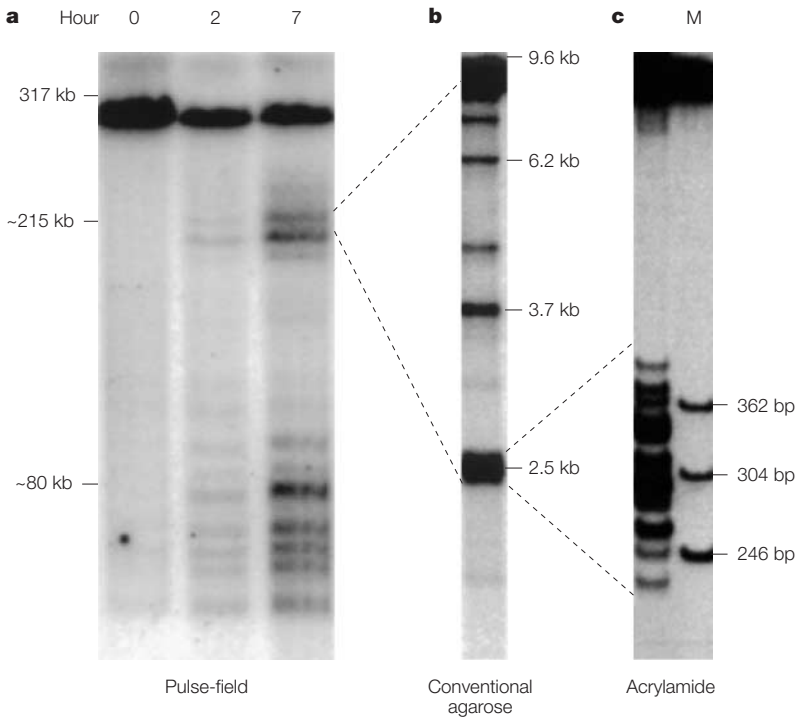

Figure 5 | **Nonrandom distribution of double-strand breaks in budding yeast, revealed by different resolutions of gel analysis. a** | The double-strand break (DSB) distribution across the entire length of a budding yeast chromosome (chromosome 3) is revealed by electrophoresis followed by Southern blotting and indirect end-labelling with a probe directed to sequences near the end of the chromosome. Samples were prepared from a synchronous cell culture at various times after induction of meiosis. Immediately before entry into meiosis (0 h), only full-length chromosomes are seen. As the cells progress into meiosis, faster migrating species are observed, corresponding to chromosomes that have suffered one or more DSBs. Approximate sizes of prominent bands are indicated. **b** | At higher resolution, a single band on a pulse-field gel can be resolved into several discrete regions of DSB formation (hot spots) that are separated by regions with few or no DSBs. In this example, genomic DNA was prepared 7 h into meiosis, digested with a restriction enzyme, separated and analysed by Southern blotting and indirect end-labelling with a probe derived from sequences in the *ARE1* gene. The full-length parental band is 9.6 kb; each of the DSB bands maps to the promoter of a gene in this region. Approximate sizes of some of the bands are indicated. **c** | At even higher resolution, DSBs within a hot spot occur at particular phosphodiester bonds that are distributed nonrandomly across ~150 bp. In this example, digested genomic DNA was analysed by Southern blotting after electrophoresis. Sizes of molecular weight standards are shown (lane M). For all three panels, samples were prepared from strains carrying a mutation (*rad50S*) in which DSBs form but are not processed further. Adapted with permission from REF. 17 © (1997) National Academy of Science and REF. 33 © (1995) Macmillan Magazines Ltd.

to proper assembly of DSB-repair complexes, and by ensuring the appropriate coordination of recombination with the formation of higher-order chromosome structures.

Effects on genome structure and evolution

In every case in which mammalian recombination events have been analysed at sufficiently high resolution, crossovers and non-crossover conversions cluster within small regions that are surrounded by large stretches of recombinationally suppressed DNA. Because such high-resolution data are so far available for only a few (and perhaps atypical) loci in mice and humans, extrapolating this pattern to the genome as a whole warrants caution. Nevertheless, this highly punctuated distribution is similar to that documented across large portions of the genome in budding yeast. So, it seems probable that this pattern will turn out to be the rule rather than the exception for mammals as well, with most recombination events occurring within such localized hot spots in every meiosis. If the average spacing between hot spots or hot-spot clusters within the human MHC class II region (~65 kb on average) is representative of spacing in the 3.2×10^9-bp genome, then there should be ~50,000 recombination hot spots. Such a punctate crossover distribution would have several consequences for genome structure and evolution.

LD blocks in the human genome. Recently, it has emerged that a large part of the human genome consists of so-called LD, or haplotype, blocks[11,69–72]. Within each block, polymorphic markers are in association with each other, with the net effect that there is a limited repertoire of haplotypes (reviewed in REF. 73). This block structure and limited haplotype diversity provides an important tool to simplify the search for genetic determinants of common disease (BOX 3).

On the basis of the precedent set by the MHC (see above), it has been suggested that the block structure in the rest of the genome is mainly a consequence of highly localized recombination hot spots being separated by recombinationally 'cold' DNA (the cold sequences would therefore make up the blocks themselves)[69,71]. It is important to note, however, that recombination distributions probably do not explain all aspects of haplotype-block structure. Genetic drift can generate haplotype blocks in the absence of recombination hot spots, as shown by various simulation studies[72,74–76]. There is also some experimental evidence that highly active recombination hot spots do not necessarily break down LD fully (A.J.J., unpublished observations), and that not all regions of LD breakdown contain (sperm) recombination hot spots (L.K. and A.J.J., manuscript in preparation). Recently, hot-spot locations and crossover activity have been estimated using COALESCENT APPROACHES, and examples so far seem to agree remarkably well with sperm-crossover data[77,78]. Such approaches might prove more useful for predicting hot spots than LD/haplotype structure alone.

In addition to contributions to the block-like haplotype structure, the occurrence of punctate recombination distributions imposes important considerations for the evolutionary dynamics of haplotype diversity. First, most gene-conversion events should only be able to create novel allelic combinations at a localized level within hot spots, and therefore should have little, if any, effect on haplotypes that are contained within long LD blocks (there is, however, some evidence that rare gene conversion outside hot spots might contribute to haplotype diversification within LD blocks[79]). Second, the existence of LD blocks results in the correlated behaviour of multiple SNPs. Specifically, within a recombinationally suppressed region of DNA, the fixation or extinction of alleles will not occur independently; instead, loss of a haplotype from a population will result in the loss of all variants that are restricted to that haplotype.

COALESCENT APPROACHES
A means of investigating the shared genealogical history of genes. A genealogy is constructed backwards in time starting with the present-day sample. Lineages coalesce when they have a common ancestor.

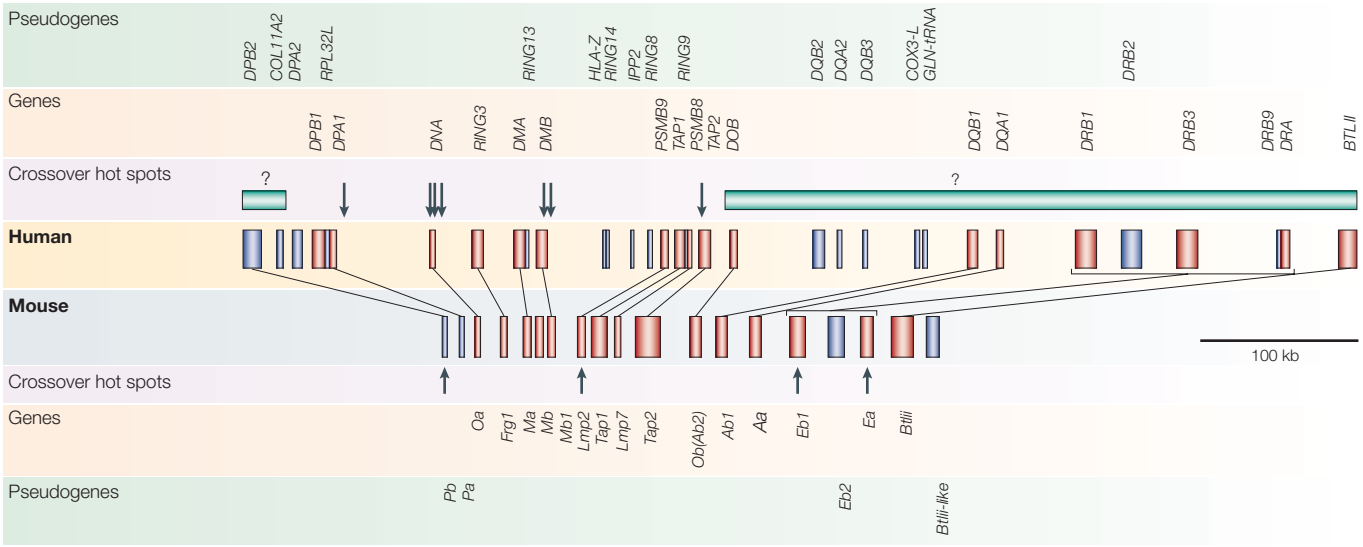

Figure 6 | **Comparison of crossover hot-spot locations in the human and mouse major histocompatibility complex class II region.** Genes are shown as red boxes, and pseudogenes as blue boxes. Orthologous genes are connected with lines (adapted from REF. 134). Known hot spots (REFS 10,11,14,15,109–114,116, 135 and L. K. and A.J.J., manuscript in preparation) are indicated with arrows. Mouse hot spots were, however, identified using pedigrees and it is probable that there are further weaker hot spots in this region of the mouse major histocompatibility complex (MHC). Regions in the human MHC that have not yet been targeted for high-resolution linkage disequilibrium and/or sperm-crossover analyses are marked with a thick green line.

Meiotic drive. The DSB-repair model for meiotic recombination[80] (FIG. 4) predicts that the chromatid on which the DSB is made is the recipient of information. This prediction is borne out in studies in budding yeast (see, for example, REF. 81). If DSB formation at a particular hot spot has an equal chance of occurring on either homologue, then within a population of recombinant molecules, the crossover breakpoints should map to the same locations for the 2 reciprocal crossover configurations, and alleles at any polymorphic sequences within the hot spot itself should be transmitted in a 1:1 ratio (see REF. 82 for a detailed discussion). On the other hand, if one homologue is preferentially used for DSB formation, crossover asymmetry is observed: alleles near the hot spot centre that are on the non-initiating chromosome are over-transmitted to crossover molecules, which results in different crossover-breakpoint locations for the reciprocal crossover configurations. Such over-transmission is evident at a human MHC hot spot, in which DSBs are preferentially initiated on the chromosome that carries the A allele of a G/A SNP near the centre of the hot spot, leading to an approximately 87:13 over-transmission of the non-initiating G allele to crossover progeny[82]. The result is weak but significant MEIOTIC DRIVE in favour of the fixation of the non-initiating G allele. Crossover asymmetry (disparity) has also been found at other hot spots in humans and mice[7,15], which is consistent with the 'hot-spot conversion paradox'[83] whereby hot spots should self-destruct through meiotic drive consistently in favour of recombination-suppressing variants.

MEIOTIC DRIVE
Distortion of meiotic inheritance such that one allele at a heterozygous site is recovered in greater than half of the gametes (as opposed to the expected 50:50 Mendelian segregation).

Box 3 | Using haplotype-block structure in genetic mapping

Gray *et al.*[123], and a little later Johnson *et al.*[124], pointed out that for a non-recombining haplotype block, a lower density of (SNP) markers is sufficient for disease-mapping purposes. Such SNPs were called 'haplotype-tag SNPs' (htSNPs)[124]. An htSNP could capture the information of its haplotype block, and therefore cut down the cost and labour of having to analyse a much higher number of SNPs. Simplifying the search for genetic determinants of common disease immediately sparked huge interest. htSNPs should reduce the problem of disease mapping to two steps: finding haplotype blocks, and then, within the patient cohort, finding the haplotype in the block with the strongest association to disease. The HapMap Project, launched in 2001, aims to map haplotype-block structure in DNA samples from Nigeria, Japan, China and the United States, and to define htSNPs for each block[125]. Although this will yield no direct information on the cause of haplotype blocks, it will generate data on how long and how 'universal' haplotype blocks are on average.

Some words of caution are needed. The definition of haplotype blocks is always dependent on marker density — it should be kept in mind that for some genomic regions or in certain populations, high-frequency SNPs, or indeed clear block boundaries, might simply not exist. As yet, it is unclear what statistical thresholds are appropriate to define block boundaries. Depending on the (arbitrary) stringency of block definition, a large LD block might be broken into shorter sub-blocks, each of lower diversity. It is also possible that an aetiological variant lies inside a recombination hot spot and would be impossible to identify using a haplotype-tagging strategy.

PCR COLONY METHOD
A method in which individual DNA molecules are immobilized in a matrix and then subjected to PCR amplification *in situ*. For (haploid) sperm DNA, polymorphic sites can be typed within amplification products to yield haplotypes of chromosomal regions, allowing, in principle, the identification of recombinant molecules.

Nucleotide diversity. An inverse correlation between nucleotide diversity and LD has been noted[84], and there seems to be a strong positive correlation between nucleotide variability and recombination rate at the megabase level[85]. One way to account for these observations is to postulate that recombination might be mutagenic. Indeed, homologous recombination that is induced by a DSB in mitotically growing budding-yeast cells does seem to be mutagenic[86], although it is not known whether this is true of meiotic recombination as well. Alternatively, the localized nature of meiotic recombination events predicts that SNPs within and near a recombination hot spot might be able to escape correlated shifts in allele frequencies that occur within the LD blocks that flank the hot spot. Specifically, alleles within hot spots could avoid rapid correlated extinctions by being constantly reshuffled onto different haplotypes. This idea is supported by the relatively stable levels of nucleotide diversity that are seen in hot spots, but not LD blocks, across populations[79].

Conclusions

For the MHC, LD data and limited numbers of familial crossovers[87] are in good agreement with sperm-crossover data, which indicates that the same hot spots also function in females but at different levels of activity. So, the most intense MHC hot spot (*DNA3*) shows similar levels of activity in males and females, whereas a hot spot in the *TAP2* gene seems to be some 30-fold 'hotter' in women[10]. Therefore, the same genomic locations might be potential recombination hot spots in all individuals, but there could be some mechanism that sets the crossover activity to a different level in the two sexes or for different haplotypes. Variation in recombination rates between individuals has been noted at the genome-wide level, as well as for relatively large chromosomal domains and at localized hot spots (see, for example, REFS 7,9,15,40,82,88,89).

When considering recombination in the context of selection, it is remarkable how, in the human MHC at least, the distribution of sperm crossovers agrees with patterns of LD, which is a population-based measure of recombination. The former involves direct detection of recombinants that have not been subjected to natural selection (apart from successfully completing meiosis), whereas the latter consists of a collection of different haplotypes that have succeeded in surviving in the population in the face of selective forces. This indicates that in this segment of the MHC, selection has had little effect on the length of haplotype blocks and supports the idea that recombination hot spots, rather than population-genetic factors, are key players in shaping diversity. Of course, it is possible that the location of the hot spots is dictated by selection in the first place. In the light of this, it would be interesting to see if hot-spot locations are conserved between humans and evolutionarily much closer species, such as the great apes.

Future challenges include identifying and characterizing more examples of mammalian hot spots to begin to understand the factors that govern their distribution. The HapMap Project (BOX 3) should reveal how general the block-like patterns of LD in humans are and how long, on average, these blocks are. Once a large number of regions of LD breakdown have been identified, they could in theory be systematically targeted for sperm-crossover assays. The drawback of sperm PCR methods lies in the laborious fine-tuning of the assays, which makes them almost impossible to scale up. Alternative approaches are therefore necessary to analyse crossovers on the whole-genome scale; haplotyping individual DNA molecules across putative crossover hot spots using single-molecule imaging or the PCR COLONY METHOD[90] might be one way to do this.

1. Petronczki, M., Siomos, M. F. & Nasmyth, K. Un menage a quatre: the molecular biology of chromosome segregation in meiosis. *Cell* **112**, 423–440 (2003).
2. Moore, D. P. & Orr-Weaver, T. L. Chromosome segregation during meiosis: building an unambivalent bivalent. *Curr. Top. Dev. Biol.* **37**, 263–299 (1998).
3. Lichten, M. & Goldman, A. S. Meiotic recombination hotspots. *Annu. Rev. Genet.* **29**, 423–444 (1995).
A lucid and detailed review of fungal and other recombination hot spots in meiosis.
4. Sherry, S. T. *et al.* dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* **29**, 308–311 (2001).
5. Hubert, R., MacDonald, M., Gusella, J. & Arnheim, N. High resolution localization of recombination hot spots using sperm typing. *Nature Genet.* **7**, 420–424 (1994).
6. Arnheim, N., Calabrese, P. & Nordborg, M. Hot and cold spots of recombination in the human genome: the reason we should find them and how this can be achieved. *Am. J. Hum. Genet.* **73**, 5–16 (2003).
7. Jeffreys, A. J., Murray, J. & Neumann, R. High-resolution mapping of crossovers in human sperm defines a minisatellite-associated recombination hotspot. *Mol. Cell* **2**, 267–273 (1998).
8. Schneider, J. A., Peto, T. E., Boone, R. A., Boyce, A. J. & Clegg, J. B. Direct measurement of the male recombination fraction in the human β-globin hot spot. *Hum. Mol. Genet.* **11**, 207–215 (2002).
9. Cullen, M., Perfetto, S. P., Klitz, W., Nelson, G. & Carrington, M. High-resolution patterns of meiotic recombination across the human major histocompatibility complex. *Am. J. Hum. Genet.* **71**, 759–776 (2002).
10. Jeffreys, A. J., Ritchie, A. & Neumann, R. High resolution analysis of haplotype diversity and meiotic crossover in the human TAP2 recombination hotspot. *Hum. Mol. Genet.* **9**, 725–733 (2000).
11. Jeffreys, A. J., Kauppi, L. & Neumann, R. Intensely punctate meiotic recombination in the class II region of the major histocompatibility complex. *Nature Genet.* **29**, 217–222 (2001).
Demonstrates the coincidence of linkage-disequilibrium breaks and sperm-crossover hot spots in the human major histocompatibility complex, and shows that most recombination events are highly localized.
12. May, C. A., Shone, A. C., Kalaydjieva, L., Sajantila, A. & Jeffreys, A. J. Crossover clustering and rapid decay of linkage disequilibrium in the Xp/Yp pseudoautosomal gene *SHOX*. *Nature Genet.* **31**, 272–275 (2002).
13. Lien, S., Szyda, J., Schechinger, B., Rappold, G. & Arnheim, N. Evidence for heterogeneity in recombination in the human pseudoautosomal region: high resolution analysis by sperm typing and radiation-hybrid mapping. *Am. J. Hum. Genet.* **66**, 557–566 (2000).
14. Guillon, H. & de Massy, B. An initiation site for meiotic crossing-over and gene conversion in the mouse. *Nature Genet.* **32**, 296–299 (2002).
15. Yauk, C. L., Bois, P. R. & Jeffreys, A. J. High-resolution sperm typing of meiotic recombination in the mouse MHC *E β* gene. *EMBO J.* **22**, 1389–1397 (2003).
16. Thuriaux, P. Is recombination confined to structural genes on the eukaryotic genome? *Nature* **268**, 460–462 (1977).
17. Baudat, F. & Nicolas, A. Clustering of meiotic double-strand breaks on yeast chromosome III. *Proc. Natl Acad. Sci. USA* **94**, 5213–5218 (1997).
With reference 28, this paper demonstrates the nonrandom distribution of double-strand breaks across large regions of the yeast genome.
18. 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).
This paper and reference 14 demonstrate that mammalian sperm-crossover hot spots are also hot spots for non-crossover gene conversion.
19. Jeffreys, A. J. & May, C. A. DNA enrichment by allele-specific hybridization (DEASH): a novel method for haplotyping and for detecting low-frequency base substitutional variants and recombinant DNA molecules. *Genome Res.* **13**, 2316–2324 (2003).
20. Qin, J., Richardson, L. L., Jasin, M., Handel, M. A. & Arnheim, N. Mouse strains with an active *H2-Ea* meiotic recombination hot spot exhibit increased levels of *H2-Ea*-specific DNA breaks in testicular germ cells. *Mol. Cell. Biol.* **24**, 1655–1666 (2004).
21. Keeney, S. Mechanism and control of meiotic recombination initiation. *Curr. Top. Dev. Biol.* **52**, 1–53 (2001).
22. Paques, F. & Haber, J. E. Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisiae*. *Microbiol. Mol. Biol. Rev.* **63**, 349–404 (1999).
23. Smith, K. N. & Nicolas, A. Recombination at work for meiosis. *Curr. Opin. Genet. Dev.* **8**, 200–211 (1998).

24. Baudat, F., Manova, K., Yuen, J. P., Jasin, M. & Keeney, S. Chromosome synapsis defects and sexually dimorphic meiotic progression in mice lacking Spo11. *Mol. Cell* **6**, 989–998 (2000).
25. Romanienko, P. J. & Camerini-Otero, R. D. The mouse *Spo11* gene is required for meiotic chromosome synapsis. *Mol. Cell* **6**, 975–987 (2000).
26. Wu, T.-C. & Lichten, M. Meiosis-induced double-strand break sites determined by yeast chromatin structure. *Science* **263**, 515–518 (1994).
This work provides one of the first and clearest descriptions of the link between double-strand break formation and local chromatin structure in yeast.
27. Zenvirth, D. *et al.* Multiple sites for double-strand breaks in whole meiotic chromosomes of *Saccharomyces cerevisiae*. *EMBO J.* **11**, 3441–3447 (1992).
28. 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).
29. Fan, Q. Q. & Petes, T. D. Relationship between nuclease-hypersensitive sites and meiotic recombination hot spot activity at the *HIS4* locus of *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **16**, 2037–2043 (1996).
30. Ohta, K., Shibata, T. & Nicolas, A. Changes in chromatin structure at recombination initiation sites during yeast meiosis. *EMBO J.* **13**, 5754–5763 (1994).
31. Xu, L. & Kleckner, N. Sequence nonspecific double-strand breaks and interhomolog interactions prior to double-strand break formation at a meiotic recombination hot spot in yeast. *EMBO J.* **14**, 5115–5128 (1995).
32. Xu, F. & Petes, T. D. Fine-structure mapping of meiosis-specific double-strand DNA breaks at a recombination hotspot associated with an insertion of telomeric sequences upstream of the *HIS4* locus in yeast. *Genetics* **143**, 1115–1125 (1996).
33. Liu, J., Wu, T.-C. & Lichten, M. The location and structure of double-strand DNA breaks induced during yeast meiosis: evidence for a covalently linked DNA-protein intermediate. *EMBO J.* **14**, 4599–4608 (1995).
34. de Massy, B., Rocco, V. & Nicolas, A. The nucleotide mapping of DNA double-strand breaks at the *CYS3* initiation site of meiotic recombination in *Saccharomyces cerevisiae*. *EMBO J.* **14**, 4589–4598 (1995).
35. Diaz, R. L., Alcid, A. D., Berger, J. M. & Keeney, S. Identification of residues in yeast Spo11p critical for meiotic DNA double-strand break formation. *Mol. Cell. Biol.* **22**, 1106–1115 (2002).
36. Nicolas, A. Relationship between transcription and initiation of meiotic recombination: toward chromatin accessibility. *Proc. Natl Acad. Sci. USA* **95**, 87–89 (1998).
37. Wahls, W. P. Meiotic recombination hotspots: shaping the genome and insights into hypervariable minisatellite DNA change. *Curr. Top. Dev. Biol.* **37**, 37–75 (1998).
38. Petes, T. D. Meiotic recombination hot spots and cold spots. *Nature Rev. Genet.* **2**, 360–369 (2001).
39. de Massy, B. Distribution of meiotic recombination sites. *Trends Genet.* **19**, 514–522 (2003).
40. Kong, A. *et al.* A high-resolution recombination map of the human genome. *Nature Genet.* **31**, 241–247 (2002).
41. Wu, T.-C. & Lichten, M. Factors that affect the location and frequency of meiosis-induced double-strand breaks in *Saccharomyces cerevisiae*. *Genetics* **140**, 55–66 (1995).
42. Borde, V., Wu, T.-C. & Lichten, M. Use of a recombination reporter insert to define meiotic recombination domains on chromosome III of *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **19**, 4832–4842 (1999).
Demonstrates that recombination activity is partially a function of position on the chromosome.
43. Blat, Y., Protacio, R. U., Hunter, N. & Kleckner, N. Physical and functional interactions among basic chromosome organizational features govern early steps of meiotic chiasma formation. *Cell* **111**, 791–802 (2002).
44. Ashley, T. G-band position effects on meiotic synapsis and crossing over. *Genetics* **118**, 307–317 (1988).
45. Holmquist, G. P. Chromosome bands, their chromatin flavors, and their functional features. *Am. J. Hum. Genet.* **51**, 17–37 (1992).
46. Kirkpatrick, D. T., Wang, Y. H., Dominska, M., Griffith, J. D. & Petes, T. D. Control of meiotic recombination and gene expression in yeast by a simple repetitive DNA sequence that excludes nucleosomes. *Mol. Cell. Biol.* **19**, 7661–7671 (1999).
47. Keeney, S. & Kleckner, N. Communication between homologous chromosomes: genetic alterations at a nuclease-hypersensitive site can alter mitotic chromatin structure at that site both *in cis* and *in trans*. *Genes Cells* **1**, 475–489 (1996).
48. White, M. A., Detloff, P., Strand, M. & Petes, T. D. A promoter deletion reduces the rate of mitotic, but not meiotic, recombination at the *HIS4* locus in yeast. *Curr. Genet.* **21**, 109–116 (1992).
49. Cao, L., Alani, E. & Kleckner, N. A pathway for generation and processing of double-strand breaks during meiotic recombination in *S. cerevisiae*. *Cell* **61**, 1089–1101 (1990).
50. Nasar, F., Jankowski, C. & Nag, D. K. Long palindromic sequences induce double-strand breaks during meiosis in yeast. *Mol. Cell. Biol.* **20**, 3449–3458 (2000).
51. Jankowski, C., Nasar, F. & Nag, D. K. Meiotic instability of CAG repeat tracts occurs by double-strand break repair in yeast. *Proc. Natl Acad. Sci. USA* **97**, 2134–2139 (2000).
52. Nag, D. K. & Kurst, A. A 140-bp-long palindromic sequence induces double-strand breaks during meiosis in the yeast *Saccharomyces cerevisiae*. *Genetics* **146**, 835–847 (1997).
53. White, M. A., Dominska, M. & Petes, T. D. Transcription factors are required for the meiotic recombination hotspot at the *HIS4* locus in *Saccharomyces cerevisiae*. *Proc. Natl Acad. Sci. USA* **90**, 6621–6625 (1993).
54. Fan, Q., Xu, F. & Petes, T. D. Meiosis-specific double-strand DNA breaks at the *HIS4* recombination hot spot in the yeast *Saccharomyces cerevisiae*: control in *cis* and *trans*. *Mol. Cell. Biol.* **15**, 1679–1688 (1995).
55. Kirkpatrick, D. T., Fan, Q. & Petes, T. D. Maximal stimulation of meiotic recombination by a yeast transcription factor requires the transcription activation domain and a DNA-binding domain. *Genetics* **152**, 101–115 (1999).
56. Moens, P. B. *et al.* Rad51 immunocytology in rat and mouse spermatocytes and oocytes. *Chromosoma* **106**, 207–215 (1997).
57. Terasawa, M., Shinohara, A., Hotta, Y., Ogawa, H. & Ogawa, T. Localization of RecA-like recombination proteins on chromosomes of the lily at various meiotic stages. *Genes Dev.* **9**, 925–934 (1995).
58. Moens, P. B. *et al.* The time course and chromosomal localization of recombination-related proteins at meiosis in the mouse are compatible with models that can resolve the early DNA–DNA interactions without reciprocal recombination. *J. Cell Sci.* **115**, 1611–1622 (2002).
59. Fogel, S., Mortimer, R. K. & Lusnak, K. in *The Molecular Biology of the Yeast Saccharomyces* (eds. Strathern, J. N., Jones, E. W. & Broach, J. R.) 289–339 (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1981).
60. Mohrenweiser, H. W., Tsujimoto, S., Gordon, L. & Olsen, A. S. Regions of sex-specific hypo- and hyper-recombination identified through integration of 180 genetic markers into the metric physical map of human chromosome 19. *Genomics* **47**, 153–162 (1998).
61. Yu, A. *et al.* Comparison of human genetic and sequence-based physical maps. *Nature* **409**, 951–953 (2001).
62. Roberts, P. A. Differences in synaptic affinity of chromosome arms of *Drosophila melanogaster* revealed by differential sensitivity to translocation heterozygosity. *Genetics* **71**, 401–415 (1972).
63. Parker, J. S., Palmer, R. W., Whitehorn, M. A. F. & Edgar, L. A. Chiasma frequency effects of structural chromosome change. *Chromosoma* **85**, 673–686 (1982).
64. Hunter, N., Chambers, S. R., Louis, E. J. & Borts, R. H. The mismatch repair system contributes to meiotic sterility in an interspecific yeast hybrid. *EMBO J.* **15**, 1726–1733 (1996).
65. Rayssiguier, C., Thaler, D. S. & Radman, M. The barrier to recombination between *Escherichia coli* and *Salmonella typhimurium* is disrupted in mismatch-repair mutants. *Nature* **342**, 396–401 (1989).
66. Wall, J. D., Frisse, L. A., Hudson, R. R. & Di Rienzo, A. Comparative linkage-disequilibrium analysis of the β-globin hotspot in primates. *Am. J. Hum. Genet.* **73**, 1330–1340 (2003).
67. Lamb, N. E. *et al.* Characterization of susceptible chiasma configurations that increase the risk for maternal nondisjunction of chromosome 21. *Hum. Mol. Genet.* **6**, 1391–1399 (1997).
68. Lamb, N. E. *et al.* Susceptible chiasmate configurations of chromosome 21 predispose to non-disjunction in both maternal meiosis I and meiosis II. *Nature Genet.* **14**, 400–405 (1996).
69. Daly, M. J., Rioux, J. D., Schaffner, S. F., Hudson, T. J. & Lander, E. S. High-resolution haplotype structure in the human genome. *Nature Genet.* **29**, 229–232 (2001).
70. Patil, N. *et al.* Blocks of limited haplotype diversity revealed by high-resolution scanning of human chromosome 21. *Science* **294**, 1719–1723 (2001).
71. Gabriel, S. B. *et al.* The structure of haplotype blocks in the human genome. *Science* **296**, 2225–2229 (2002).
With reference 70, this paper demonstrates that a large portion of the human genome is contained within haplotype blocks and that these blocks show limited diversity.
72. Phillips, M. S. *et al.* Chromosome-wide distribution of haplotype blocks and the role of recombination hot spots. *Nature Genet.* **33**, 382–387 (2003).
73. Wall, J. D. & Pritchard, J. K. Haplotype blocks and linkage disequilibrium in the human genome. *Nature Rev. Genet.* **4**, 587–597 (2003).
74. Wang, N., Akey, J. M., Zhang, K., Chakraborty, R. & Jin, L. Distribution of recombination crossovers and the origin of haplotype blocks: the interplay of population history, recombination, and mutation. *Am. J. Hum. Genet.* **71**, 1227–1234 (2002).
75. Zhang, K. *et al.* Randomly distributed crossovers may generate block-like patterns of linkage disequilibrium: an act of genetic drift. *Hum. Genet.* **113**, 51–59 (2003).
76. Anderson, E. C. & Slatkin, M. Population-genetic basis of haplotype blocks in the 5q31 region. *Am. J. Hum. Genet.* **74**, 40–49 (2004).
77. Li, N. & Stephens, M. Modeling linkage disequilibrium and identifying recombination hotspots using single-nucleotide polymorphism data. *Genetics* **165**, 2213–2233 (2003).
78. Stumpf, M. P. & McVean, G. A. Estimating recombination rates from population-genetic data. *Nature Rev. Genet.* **4**, 959–968 (2003).
79. Kauppi, L., Sajantila, A. & Jeffreys, A. J. Recombination hotspots rather than population history dominate linkage disequilibrium in the MHC class II region. *Hum. Mol. Genet.* **12**, 33–40 (2003).
80. Szostak, J. W., Orr-Weaver, T. L., Rothstein, R. J. & Stahl, F. W. The double-strand-break repair model for recombination. *Cell* **33**, 25–35 (1983).
81. Nicolas, A., Treco, D., Schultes, N. P. & Szostak, J. W. An initiation site for meiotic gene conversion in the yeast *Saccharomyces cerevisiae*. *Nature* **338**, 35–39 (1989).
This is a reader-friendly description of a meiotic conversion hot spot in yeast that is accessible to the non-specialist.
82. Jeffreys, A. J. & Neumann, R. Reciprocal crossover asymmetry and meiotic drive in a human recombination hot spot. *Nature Genet.* **31**, 267–271 (2002).
83. Boulton, A., Myers, R. S. & Redfield, R. J. The hotspot conversion paradox and the evolution of meiotic recombination. *Proc. Natl Acad. Sci. USA* **94**, 8058–8063 (1997).
84. Ardlie, K. G., Kruglyak, L. & Seielstad, M. Patterns of linkage disequilibrium in the human genome. *Nature Rev. Genet.* **3**, 299–309 (2002).
85. Nachman, M. W. Single nucleotide polymorphisms and recombination rate in humans. *Trends Genet.* **17**, 481–485 (2001).
86. Strathern, J. N., Shafer, B. K. & McGill, C. B. DNA synthesis errors associated with double-strand-break repair. *Genetics* **140**, 965–972 (1995).
87. Cullen, M. *et al.* Characterization of recombination in the *HLA* class II region. *Am. J. Hum. Genet.* **60**, 397–407 (1997).
88. Tease, C., Hartshorne, G. M. & Hulten, M. A. Patterns of meiotic recombination in human fetal oocytes. *Am. J. Hum. Genet.* **70**, 1469–1479 (2002).
89. Lynn, A. *et al.* Covariation of synaptonemal complex length and mammalian meiotic exchange rates. *Science* **296**, 2222–2225 (2002).
90. Mitra, R. D. *et al.* Digital genotyping and haplotyping with polymerase colonies. *Proc. Natl Acad. Sci. USA* **100**, 5926–5931 (2003).
91. Hattori, M. *et al.* The DNA sequence of human chromosome 21. *Nature* **405**, 311–319 (2000).
92. Broman, K. W., Murray, J. C., Sheffield, V. C., White, R. L. & Weber, J. L. Comprehensive human genetic maps: individual and sex-specific variation in recombination. *Am. J. Hum. Genet.* **63**, 861–869 (1998).
93. Lynn, A. *et al.* Patterns of meiotic recombination on the long arm of human chromosome 21. *Genome Res.* **10**, 1319–1332 (2000).
94. Anderson, L. K., Reeves, A., Webb, L. M. & Ashley, T. Distribution of crossing over on mouse synaptonemal complexes using immunofluorescent localization of MLH1 protein. *Genetics* **151**, 1569–1579 (1999).
95. Barlow, A. L. & Hulten, M. A. Crossing over analysis at pachytene in man. *Eur. J. Hum. Genet.* **6**, 350–358 (1998).
96. Froenicke, L., Anderson, L. K., Wienberg, J. & Ashley, T. Male mouse recombination maps for each autosome identified by chromosome painting. *Am. J. Hum. Genet.* **71**, 1353–1368 (2002).
97. Sun, F. *et al.* Human male recombination maps for individual chromosomes. *Am. J. Hum. Genet.* **74**, 521–531 (2004).
98. Laurie, D. A. & Hulten, M. A. Further studies on chiasma distribution and interference in the human male. *Ann. Hum. Genet.* **49**, 203–214 (1985).
99. Laurie, D. A. & Hulten, M. A. Further studies on bivalent chiasma frequency in human males with normal karyotypes. *Ann. Hum. Genet.* **49**, 189–201 (1985).
100. Lawrie, N. M., Tease, C. & Hulten, M. A. Chiasma frequency, distribution and interference maps of mouse autosomes. *Chromosoma* **104**, 308–314 (1995).
101. Shiroishi, T., Koide, T., Yoshino, M., Sagai, T. & Moriwaki, K. Hotspots of homologous recombination in mouse meiosis. *Adv. Biophys.* **31**, 119–132 (1995).
102. Hedrick, P. W. Inference of recombinational hotspots using gametic disequilibrium values. *Heredity* **60**, 435–438 (1988).
103. Chakravarti, A. *et al.* Nonuniform recombination within the human β-globin gene cluster. *Am. J. Hum. Genet.* **36**, 1239–1258 (1984).

104. Smith, R. A., Ho, P. J., Clegg, J. B., Kidd, J. R. & Thein, S. L. Recombination breakpoints in the human β-globin gene cluster. *Blood* **92**, 4415–4421 (1998).
105. Yip, S. P., Lovegrove, J. U., Rana, N. A., Hopkinson, D. A. & Whitehouse, D. B. Mapping recombination hotspots in human phosphoglucomutase (*PGM1*). *Hum. Mol. Genet.* **8**, 1699–1706 (1999).
106. Cruciani, F. *et al.* Linkage disequilibrium analysis of the human adenosine deaminase (*ada*) gene provides evidence for a lack of correlation between hot spots of equal and unequal homologous recombination. *Genomics* **82**, 20–33 (2003).
107. Twells, R. C. *et al.* Haplotype structure, LD blocks, and uneven recombination within the *LRP5* gene. *Genome Res.* **13**, 845–855 (2003).
108. Cullen, M., Erlich, H., Klitz, W. & Carrington, M. Molecular mapping of a recombination hotspot located in the second intron of the human *TAP2* locus. *Am. J. Hum. Genet.* **56**, 1350–1358 (1995).
109. Bryda, E. C., DePari, J. A., Sant'Angelo, D. B., Murphy, D. B. & Passmore, H. C. Multiple sites of crossing over within the Eb recombinational hotspot in the mouse. *Mamm. Genome* **2**, 123–129 (1992).
110. Lafuse, W. P. & David, C. S. Recombination hot spots within the I region of the mouse *H-2* complex map to the *E* β and *E* α genes. *Immunogenetics* **24**, 352–360 (1986).
111. Zimmerer, E. J. & Passmore, H. C. Structural and genetic properties of the *Eb* recombinational hotspot in the mouse. *Immunogenetics* **33**, 132–140 (1991).
112. Uematsu, Y., Kiefer, H., Schulze, R., Fischer-Lindahl, K. & Steinmetz, M. Molecular characterization of a meiotic recombinational hotspot enhancing homologous equal crossing-over. *EMBO J.* **5**, 2123–2129 (1986).
113. Steinmetz, M., Stephan, D. & Fischer Lindahl, K. Gene organization and recombinational hotspots in the murine major histocompatibility complex. *Cell* **44**, 895–904 (1986). **Provides the first account of localized crossover clustering in mammalian pedigrees.**
114. Shiroishi, T. *et al.* Recombinational hotspot specific to female meiosis in the mouse major histocompatibility complex. *Immunogenetics* **31**, 79–88 (1990).
115. Snoek, M., Teuscher, C. & van Vugt, H. Molecular analysis of the major MHC recombinational hot spot located within the *G7c* gene of the murine class III region that is involved in disease susceptibility. *J. Immunol.* **160**, 266–272 (1998).
116. Mizuno, K., Koide, T., Sagai, T., Moriwaki, K. & Shiroishi, T. Molecular analysis of a recombinational hotspot adjacent to *Lmp2* gene in the mouse MHC: fine location and chromatin structure. *Mamm. Genome* **7**, 490–496 (1996).
117. Shiroishi, T., Sagai, T., Hanzawa, N., Gotoh, H. & Moriwaki, K. Genetic control of sex-dependent meiotic recombination in the major histocompatibility complex of the mouse. *EMBO J.* **10**, 681–686 (1991). **Demonstrates the effect of modifying factors on crossing-over at a localized hot spot.**
118. Laan, M. & Paabo, S. Demographic history and linkage disequilibrium in human populations. *Nature Genet.* **17**, 435–438 (1997).
119. Zavattari, P. *et al.* Major factors influencing linkage disequilibrium by analysis of different chromosome regions in distinct populations: demography, chromosome recombination frequency and selection. *Hum. Mol. Genet.* **9**, 2947–2957 (2000).
120. Hedrick, P. W. Gametic disequilibrium measures: proceed with caution. *Genetics* **117**, 331–341 (1987).
121. Lewontin, R. C. The interaction of selection and linkage. II. Optimum models. *Genetics* **50**, 757–782 (1964).
122. Weiss, K. M. & Clark, A. G. Linkage disequilibrium and the mapping of complex human traits. *Trends Genet.* **18**, 19–24 (2002).
123. Gray, I. C., Campbell, D. A. & Spurr, N. K. Single nucleotide polymorphisms as tools in human genetics. *Hum. Mol. Genet.* **9**, 2403–2408 (2000).
124. Johnson, G. C. *et al.* Haplotype tagging for the identification of common disease genes. *Nature Genet.* **29**, 233–237 (2001).
125. National Human Genome Research Institute. International consortium launches genetic variation mapping project. *NIH News Advisory* [online], <http://genome.gov/10005336> (2002).
126. Yao, H. *et al.* Molecular characterization of meiotic recombination across the 140-kb multigenic a1-sh2 interval of maize. *Proc. Natl Acad. Sci. USA* **99**, 6157–6162 (2002).
127. Game, J. C., Sitney, K. C., Cook, V. E. & Mortimer, R. K. Use of a ring chromosome and pulsed-field gels to study interhomolog recombination, double-strand breaks, and sister-chromatid exchange in yeast. *Genetics* **123**, 695–713 (1989).
128. Sun, H., Treco, D., Schultes, N. P. & Szostak, J. W. Double-strand breaks at an initiation site for meiotic gene conversion. *Nature* **338**, 87–90 (1989).
129. Keeney, S., Giroux, C. N. & Kleckner, N. Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family. *Cell* **88**, 375–384 (1997). **With reference 130, this paper identifies Spo11 as the catalytic subunit of the meiotic double-strand break machinery, and provides evidence that it cuts DNA through a topoisomerase-like transesterase reaction.**
130. Bergerat, A. *et al.* An atypical topoisomerase II from Archaea with implications for meiotic recombination. *Nature* **386**, 414–417 (1997).
131. Sun, H., Treco, D. & Szostak, J. W. Extensive 3′-overhanging, single-stranded DNA associated with the meiosis-specific double-strand breaks at the *ARG4* recombination initiation site. *Cell* **64**, 1155–1161 (1991).
132. Bishop, D. K., Park, D., Xu, L. & Kleckner, N. *DMC1*: a meiosis-specific yeast homolog of *E. coli recA* required for recombination, synaptonemal complex formation, and cell cycle progression. *Cell* **69**, 439–456 (1992).
133. Shinohara, A., Ogawa, H. & Ogawa, T. Rad51 protein involved in repair and recombination in *S. cerevisiae* is a RecA-like protein. *Cell* **69**, 457–470 (1992).
134. Yuhki, N. *et al.* Comparative genome organization of human, murine, and feline MHC class II region. *Genome Res.* **13**, 1169–1179 (2003).
135. Isobe, T. *et al.* Molecular characterization of the *Pb* recombination hotspot in the mouse major histocompatibility complex class II region. *Genomics* **80**, 229–235 (2002).

Acknowledgements

L.K. was supported in part by the Instrumentarium Science Foundation, the Finnish Cultural Foundation and the Osk. Huttunen Foundation. A.J.J. is supported by grants from the Royal Society, the Medical Research Council and the Wellcome Trust (UK). S.K. is supported by grants from the US National Institutes of Health and the Byrne Fund.

Competing interests statement

The authors declare that they have no competing financial interests.

Online links

DATABASES

The following terms in this article are linked online to:
Entrez: http://www.ncbi.nih.gov/Entrez
ADA | *ARE1* | *HIS4* | *Lmp2* | *LRP5* | *PGM1* | *SHOX* | Spo11 | *TAP2*

FURTHER INFORMATION

Alec Jeffreys' laboratory: http://www.le.ac.uk/ge/ajj/index.html
GermOnline (a web database for information about germ cell biology, including meiosis): http://germonline.unibas.ch
International HapMap Project: http://www.hapmap.org
Scott Keeney's laboratory: http://www.ski.edu/keeney
Access to this links box is available online.