Distribution of meiotic recombination sites

Bernard de Massy

Institut de Génétique Humaine, 141 rue de la Cardonille, 34396 Montpellier cedex 5, France

Meiotic recombination generates gene conversion and crossover events that are distributed heterogeneously in the genome. Studies in yeast show that initiation of recombination, which occurs by the formation of DNA double-strand breaks, determines the distribution of gene conversion and crossover events that take place in nearby intervals. Recent data in humans and mice also indicate the presence of highly localized initiation sites that promote crossovers clustered around the region of initiation and seem to share common features with sites in yeast. On a larger scale, chromosomal domains with various recombination rates have been identified from yeast to mammals. This indicates a higher level of regulation of recombination in the genome with potential consequences on genome structure.

During meiosis, recombination between homologous chromosomes generates gene conversion (nonreciprocal events) and crossing over (reciprocal events). Analyses in many organisms show that meiotic recombination events are not distributed randomly in the genome. Given that homologous recombination can, in principle, spread over large distances, the distribution of events depends on where recombination is initiated as well as how far it extends. The combination of genetic and molecular analysis in Saccharomyces cerevisiae and, more recently, in mammals shows that recombination events (gene conversion and crossover) cluster tightly in the vicinity of initiation sites. Therefore, the distribution of exchanges essentially reflects the distribution of initiation sites. Here, I review the properties of these sites and associated events in the yeasts S. cerevisiae, and Schizosaccharomyces pombe, mice and humans.

The current model for meiotic recombination

The current model of meiotic recombination combines the double-strand-break repair (DSBR) model [1] and the synthesis-dependent strand-annealing (SDSA) model (Box 1) (reviewed in [2]). The combination of these pathways has been proposed to explain extra-chromosomal recombination in mammalian cells [3] and human minisatellite rearrangements [4]. Molecular analysis in S. cerevisiae provides direct support for this model with respect to meiotic recombination [5,6].

Several specific features of this model are particularly relevant for the comparative analysis of events in yeast and mammals presented below. First, the chromatin that initiates and, therefore, where the DSB occurs, is the recipient of genetic information. This leads to conversion on the initiating chromatid. Conversion occurs by mismatch repair, and not by double-strand gap as originally proposed. Second, the strand exchange is initiated and extended from the 3’ ends, thus, heteroduplex forms initially in regions where degradation of 5’ ends occurs but can extend further by either strand displacement or branch migration of Holliday junctions. Repair of heteroduplex produces either gene conversion or restoration depending on the choice of the corrected strand. If Holliday junctions are resolved within the interval defined by the gene-conversion tract, the sites of crossing over map at the boarders of these tracts. However, branch migration of the Holliday junctions can lead to crossover away from this region.

Local distribution of recombination in yeasts

DSBs, which are detected by molecular analysis, are generated by Spo11 in conjunction with several other proteins at many sites in the genome (reviewed in [7]). Four levels of constraint have been observed with respect to the distribution of DSBs in yeast: (1) at the nucleotide level, DSBs are not distributed randomly, but no clear sequence motif has been identified; (2) DSBs are commonly located in accessible regions of the chromatin; (3) the formation of DSBs appears to depend on chromosome organization and to occur preferentially in chromatin loops away from DNA-axis associations that are mediated by cohesins and Red1 [8,9]; and (4) DSBs are clustered in chromosome domains and are reduced in telomeric regions and the vicinity of centromeres [10,11].

Local distribution of initiation sites

DSBs occur in highly localized regions and spread over 70–250 bp (reviewed in [7]). DNA sequence analysis reveals no unique conserved consensus sequences, although a degenerate 50-bp motif partly correlates with DSB sites. However, one common feature is that DSBs are located in accessible regions of the chromatin next to either promoters or binding sites for transcription factors (reviewed in [12–14]). Based on two studies, DSB activity does not correlate with local transcriptional activity, but depends on transcription-factor binding (HIS4 in S. cerevisiae and ade6-M26 in S. pombe). The constraint imposed by chromatin is also revealed by the activity of a Gal4–Spo11 fusion protein, which induces high DSB levels in a region that contains a Gal4-binding site and normally a low DSB.

Corresponding author: Bernard de Massy (bdemassy@igh.cnrs.fr).

http://tigs.trends.com 0168-9525/$ - see front matter © 2003 Elsevier Ltd. All rights reserved. doi:10.1016/S0168-9525(03)00201-4
However, some DSB sites do not seem to be associated with the binding of transcription factors, indicating that there might be alternative determinants for accessibility to the recombination machinery. It has been proposed that a chromatin-modification code contributes to the definition of DSB sites [13].

In the DSBR pathway, the D loop captures the second 3’ end (Fig. 1c). DNA synthesis and ligation generate a double Holliday junction [78–80] with heteroduplex DNA flanking the DSB site [81]. This recombination intermediate is resolved by the cleavage and ligation of strands of same polarities at identical positions, which generates crossover products. This pathway might also generate noncrossovers (not shown). Mismatch repair of heteroduplexes can lead to either gene conversion or restoration depending on the choice of the corrected strand. Here, repair towards gene conversion is shown on both sides of the initiation point. This is the major direction of repair for markers located near the DSB site, probably because of the preferential removal of strands with an available 3’ end (i.e. strands from the initiating duplex). Repair of heteroduplex DNA away from the DSB site might take place on the opposite strand and lead to restoration.

In the synthesis-dependent strand-annealing (SDSA) pathway, the D loop is disassembled by displacement of the newly synthesized strand, which anneals with the other DSB end (Fig. 1d). Repair of the break is completed by DNA synthesis and ligation. Mismatch repair of heteroduplex DNA generates gene conversion without crossover.

In both pathways, the process of mismatch repair is presented arbitrarily at the last step. This is probably an oversimplified view because mismatch repair itself might involve several pathways and act at more than one step during the process [80]. Various alternatives could be considered: for example, early repair of heteroduplex DNA at the stage of D-loop formation could lead to conversion events on both sides of the initiation event through the SDSA pathway. Little is known about how and when Holliday junctions are processed, and when mismatch repair takes place, although some data indicate that these two events might be coordinated [82]. In the case shown, Holliday junctions do not migrate from the region of initial heteroduplex DNA, therefore crossover molecules have exchange points at the ends of conversions tracts (shown as flags). More distal crossover points might form if one or both Holliday junctions migrated away from the initiation site.

Although most data that support this model result from studies in yeast, the conservation of the major proteins involved indicates that the main lines should be conserved in yeast and mammals [83].
on both sides of the initiation site. The molecular mechanism responsible for this decrease, referred to as a gradient of gene conversion, is debated. It could involve a regulation of the directionality of mismatch repair, which influences the relative frequencies of gene conversion and restoration as a function of the distance from initiation, or it might involve a regulation of heteroduplex length, or both (Box 1). Data from several loci are reported in Table 1. At each locus gene conversion frequencies of several markers were measured. In each case, meiotic DSBs were detected and the marker located closest to the DSB site was found to have the highest frequency of gene conversion. The relative frequencies of only the highest (normalized to 1) and lowest markers are indicated in the table.

When all markers tested are taken into account, the log of the frequencies of gene conversion follows a linear relationship as a function of distance with a 99% confidence interval in all cases tested except the HIS4 locus. This exponential relationship is best explained by assuming that repair events that lead to gene conversion travel from the initiation point with a fixed probability \( (1 - p) \) of stopping at a given position \( n \). The probability that a gene conversion tract extends over a length \( L \) to position \( n \) from initiation is: \( P(L = n) = p^L (1 - p) \). A linear relationship is seen if the number of gene-conversion events is not well fitted by an exponential law and the significance of the predicted mean is therefore questionable. This exponential relationship is best explained by assuming that repair events that lead to gene conversion travel from the initiation point with a fixed probability \( (1 - p) \) of stopping at a given position \( n \). The probability that a gene conversion tract extends over a length \( L \) to position \( n \) from initiation is: \( P(L = n) = p^L (1 - p) \). Examples of the distributions of gene-conversion frequencies and gene-conversion-tract lengths at the ARG4 locus are shown in Fig. 1. Applying this to other loci results in a fairly homogeneous range of gene-conversion-tract lengths in S. cerevisiae (except at the HIS4 locus) and S. pombe. The means range from 800–2000 bp at most loci.

The HIS4 locus, the distribution of gene conversion is not well fitted by an exponential law and the significance of the predicted mean is therefore questionable. This might reflect either some specific property of recombination in this region or that some of the events detected are not initiated from DSBs located in the HIS4 promoter region. The observation that a high proportion (~50%) of conversion events at a distal marker do not involve co-conversion with a proximal promoter marker indicates the presence of distal initiation events [17].

In agreement with the DSBR model and the processivity of heteroduplex repair, analysis at ARG4 and ade6 also show that most conversion tracts are continuous [18,19]. However, discontinuous conversion tracts are observed in other fungi, such as Neurospora crassa [20,21] and Ascobolus immersus [22], which indicates either alternative intermediates or alternative processing of intermediates by the mismatch-repair machinery in these species.

### Distribution of crossing over

The model of DSBR allows alternatives to how and where Holliday junctions could be resolved, and mapping gene-conversion events does not directly address this issue. There are at least two possibilities in terms of location: either Holliday junctions are resolved in the area covered by the conversion tracts or they migrate further away in flanking regions. It should be emphasized that the location of crossovers might be distinct from the actual point of Holliday-junction resolution. Two studies performed in S. cerevisiae, provide more information about this issue. Symington et al. [23,24], used polymorphisms to select and map crossovers in a 22 kb interval and two regions that contained high densities of crossover were identified.

### Table 1. Analysis of gene-conversion gradients in yeasts and of a crossover gradient in mice

<table>
<thead>
<tr>
<th>Saccharomyces cerevisiae</th>
<th>Schizosaccharomyces pombe</th>
<th>Mus musculus</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ARG4 (1)</strong></td>
<td><strong>ARG4 (2)</strong></td>
<td><strong>HIS2</strong></td>
</tr>
<tr>
<td>Highest marker*</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Lowest marker</td>
<td>0.02</td>
<td>0.12</td>
</tr>
<tr>
<td>Distanceb</td>
<td>1383</td>
<td>2100</td>
</tr>
<tr>
<td>P</td>
<td>0.9975</td>
<td>0.999</td>
</tr>
<tr>
<td>Linear correlation within 99% confidence (P &lt; 0.01)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean tract length (bp)</td>
<td>794</td>
<td>2078</td>
</tr>
<tr>
<td>95% confidence interval</td>
<td>667–981</td>
<td>1420–3860</td>
</tr>
</tbody>
</table>

*At each locus, the frequency of events for the highest frequency marker is normalized to 1.

bDistance in base pairs between the highest and lowest frequency markers. The frequency of events as a function of distance follows an exponential law (~p value). References: ARG4 (1), [86]; ARG4 (2), [18]; HIS2, [81]; HIS4 (1), [62]; HIS4 (2), M.F.F. Abdulla and R.H. Borts, pers. commun.; ade6, [19]; ura4-aim, J. Kohli, pers. comm.; Psmb9, [36].

1) and (2) refer to two independent experiments.

![Fig. 1. (a) Distribution of gene conversion frequencies at the ARG4 locus. Values are normalized to 1 for the marker with the highest frequency set at position 0. Experimental values for eight markers are shown in pink. The best fit of the linear correlation between the log of frequencies and distance is shown in blue (p = 0.99748). P(L = n) = p^L (1 - p). (b) Distribution of the gene-conversion tract length assuming a bidirectional processing on either sides of the initiation point and an exponential law of p = 0.99748. P(L = n|n + 1)p^L (1 - p)2. The mean of this distribution is 794 bp.](http://tigs.trends.com)
DSBs were analyzed. Subsequent DSB analysis in this region showed that the two regions with high crossover density also have high levels of DSBs [25]. From the sizes of the interval tested, it is likely that most crossovers occur within ~2–3 kb of their initiation site. Most gene-conversion tracts detected are <2 kb long and are continuous. A minority (8%) of long conversion events (4.5–12.0 kb) were also identified. Using a similar approach, Borts et al. mapped gene conversion and crossing-over association. As above, crossover mapped to the boarders of conversion tracts in most cases, but positions of the DSBs were not determined [26,27]. These data are consistent with a model in which the Holliday junctions do not migrate from the region of initial asymmetric heteroduplex (Box 1).

By contrast, a recent comparison of DSB and crossover density in large intervals in the S. pombe genome led Young et al. [28] to propose that Holliday junctions migrate and are resolved at long distances from their site of initiation in this organism. However, in a study of gene conversion and crossover association at the ade6 locus, crossover separated from selected conversion events were detected in one interval and found to represent only a minority of events [19]. Further analyses are therefore needed to determine the proportion of events that follow this pathway in S. pombe [29].

### Local distribution of recombination events in mammals

#### Local distribution of sites

Several recombination hot spots have been defined in humans and mice. These are based on the higher density of crossing over (relative to either adjacent regions or the genome average) identified from pedigree analysis in humans and screening recombinant in-crosses between laboratory mouse strains. More recent studies have also identified meiotic recombination hot spots and high-resolution mapping of recombination events through direct molecular analysis in humans and mice. The level of precision, therefore, differs according to the methods used. Table 2 summarizes those that are precisely located on the physical map. Large differences are observed between the activity of these hot spots, and their definition is obviously relative. For example, recombination at Dna1 is three-fold lower than the genome average but about seven-fold higher than surrounding regions.

In terms of location, of the 17 hot spots listed, nine are intergenic (5’ and 3’ of genes), three are within introns, and five are mapped at insufficient resolution to distinguish intronic and exonic locations. Although analysis of DNA sequences in these regions has identified no obvious motifs, several properties have been noted [30,31]. Even though not confined to transcription-promoter regions, determinants for hot-spot localization in mammals might

### Table 2. Hot spots in humans and mice

<table>
<thead>
<tr>
<th>Hot spot*</th>
<th>Frequency (cM)</th>
<th>Interval (Kb)</th>
<th>Fold above genome average</th>
<th>Location</th>
<th>Methods</th>
<th>Locus Link databasea</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Human</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MS32</td>
<td>0.04</td>
<td>1.5</td>
<td>30</td>
<td>Intergenic, next to minisatellite</td>
<td>Pedigree analysis and direct molecular detection</td>
<td></td>
</tr>
<tr>
<td>CEB1</td>
<td>0.16</td>
<td>2</td>
<td>⩾70</td>
<td>Intergenic, next to minisatellite</td>
<td>Pedigree analysis and direct molecular detection</td>
<td></td>
</tr>
<tr>
<td>β-Globin</td>
<td>0.9</td>
<td>11</td>
<td>90</td>
<td>Intergenic, 5′ of β globin</td>
<td>Pedigree analysis and single sperm analysis</td>
<td>HBB; locus ID, 64162</td>
</tr>
<tr>
<td>PGM1</td>
<td>1.7</td>
<td>58</td>
<td>30</td>
<td>Intragenic, intron or exon</td>
<td>Pedigree analysis</td>
<td>PGM1; locus ID, 5236</td>
</tr>
<tr>
<td>CMT1a</td>
<td>0.0015</td>
<td>1.4</td>
<td>1.2</td>
<td>Intergenic</td>
<td>Pedigree analysis and direct molecular detection</td>
<td>CMT1A; locus ID, 1248</td>
</tr>
<tr>
<td>DNA1</td>
<td>0.0005</td>
<td>1.9</td>
<td>0.3</td>
<td>Intergenic, 5′ of DNA</td>
<td>Pedigree analysis and direct molecular detection</td>
<td>HLA-DOA; locus ID, 3111</td>
</tr>
<tr>
<td>DNA2</td>
<td>0.0037</td>
<td>1.3</td>
<td>3</td>
<td>Intergenic, 3′ of DNA, next Alu</td>
<td>Pedigree analysis and direct molecular detection</td>
<td>HLA-DOA; locus ID, 3111</td>
</tr>
<tr>
<td>DNA3</td>
<td>0.13</td>
<td>1.2</td>
<td>120</td>
<td>Intergenic, 3′ of DNA, next Alu</td>
<td>Pedigree analysis and direct molecular detection</td>
<td>HLA-DOA; locus ID, 3111</td>
</tr>
<tr>
<td>DMB1</td>
<td>0.0031</td>
<td>1.8</td>
<td>2</td>
<td>Intragenic, intron or exon</td>
<td>Pedigree analysis and direct molecular detection</td>
<td>HLA-DMB; locus ID, 3109</td>
</tr>
<tr>
<td>DMB2</td>
<td>0.028</td>
<td>1.2</td>
<td>30</td>
<td>Intragenic, 3′ of DMB</td>
<td>Pedigree analysis and direct molecular detection</td>
<td>HLA-DMB; locus ID, 3109</td>
</tr>
<tr>
<td>TAP2</td>
<td>0.0058</td>
<td>1</td>
<td>10</td>
<td>Intragenic, intron</td>
<td>Pedigree analysis and direct molecular detection</td>
<td>Tap2; locus ID, 6891</td>
</tr>
<tr>
<td>PAR1</td>
<td>0.3</td>
<td>3.9</td>
<td>90</td>
<td>Intragenic, intron or exon</td>
<td>Pedigree analysis and direct molecular detection</td>
<td>Shox; locus ID, 6473</td>
</tr>
<tr>
<td><strong>Mouse</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ea</td>
<td>0.9</td>
<td>0.4</td>
<td>3750</td>
<td>Intragenic, intron</td>
<td>Genetic crosses</td>
<td>H2-Ea; locus ID, 14968</td>
</tr>
<tr>
<td>Eb</td>
<td>0.1</td>
<td>4</td>
<td>40</td>
<td>Intragenic, intron</td>
<td>Genetic crosses</td>
<td>H2-Eb1; locus ID, 14969</td>
</tr>
<tr>
<td>Pamb9</td>
<td>1</td>
<td>1.8</td>
<td>930</td>
<td>Intergenic, 3′ of Pamb9</td>
<td>Genetic crosses and direct molecular detection</td>
<td>Pamb9; locus ID, 16912</td>
</tr>
<tr>
<td>Gc7</td>
<td>1.5</td>
<td>5.3</td>
<td>Unknown</td>
<td>Intragenic, intron or exon</td>
<td>Genetic crosses</td>
<td>D17H6S5S6-3; locus ID, 27762</td>
</tr>
<tr>
<td>Pb</td>
<td>Unknown</td>
<td>2.4</td>
<td>Unknown</td>
<td>Intragenic, intron or exon</td>
<td>Genetic crosses</td>
<td>H2-Pb; locus ID, 15004</td>
</tr>
</tbody>
</table>

*For each hot spot, the frequency of crossing-over is given in cM in the interval that defines the hot spot. The increase relative to the genome average is calculated taking a value of 0.89 cM Mb−1 for human [63] and 0.6 cM Mb−1 for mouse [64]. These are average values for male and female recombination rates. The CMT1a hot spot is a preferred site for unequal exchanges. The methods used to define the locations and frequencies of the hot spots are indicated. References: MS32, [32]; CEB1, [4]; β-globin, [65–67]; Pgm1, [68]; Cmt1a, [69]; Dna1, Dna2, Dna3, Dmb1 and Dmb2, [33]; Tap2, [30]; Par1, [70]; Ea, [71]; Eb, [72,73]; Pamb9, [35,74,75]; Gc7, [76]; Pb [31].

share common features with those described in yeast, such as chromatin accessibility. To what extent gene-regulatory elements and transcription factors play a role in this localization is currently unknown.

Local distribution of events
Direct molecular detection of recombination events was developed by A. Jeffreys group to study minisatellite rearrangements that take place at a high rate during meiosis. This strategy allows the selective amplification of recombinant molecules by PCR. In minisatellite regions, the analysis of events either with or without the exchange of flanking markers indicated the existence of two pathways, the DSBR and SDSA pathways described in Box 1. In addition, extensive polymorphisms have allowed the mapping of minisatellite rearrangements that show the presence of gene conversion and crossover gradients [4,32]. The same approach has been used to study some crossover hot spots that are not associated with minisatellites in humans and mice (Table 2). Although the activities of these hot spots differ by up to several-hundred-fold, they share common features, such as clustering of crossovers in a small interval (1–3 kb) and nonrandom distribution with a central region of highest density. In addition, all recombinant molecules detected are simple exchanges with no mosaicism.

The distribution of crossovers at the mouse Psmb9 hot spot is presented in Fig. 2. The properties of Psmb9 are compatible with the presence of an initiation site that induces recombination events that are repaired according to the DSBR model, and where crossover points correspond to the borders of gene-conversion tracts (Box 1). The distribution of crossovers fits an exponential law on either side of the central region, corresponding to a mean value for tract sizes of 480 bp, which is slightly shorter than that deduced from gene-conversion events in yeast (Table 1). This observation of highly localized crossovers fits with the analysis of linkage disequilibrium near hot spots in humans. In the regions between hot spots [e.g. the 100 kb interval between Dmb and Tap2 (Fig. 3a)], there is a strong linkage, which indicates that few crossovers extend outside the hot-spot regions [33]. Two additional pieces of information strongly support these hot spots as sites of initiation. First, one prediction of the DSBR model is that when only one homologous chromosome initiates, of the two types of crossover molecules that can be recovered, the exchange point of one will always be on one side of the initiation site, and the other on the opposite side (Box 1). In some heterozygous situations, such a bias has been observed at the human Dna2 hot spot [34] and at the mouse Psmb9 hot spot (P. Baudat and B. de Massy, pers. commun.). Second, the presence of an initiation site predicts a high frequency of gene-conversion events that are not associated with a crossover. Recombinant molecules compatible with such events occur at the Psmb9 hot spot for a marker located in the central region. The sizes of the gene-conversion events (without crossover in the interval tested) recovered in this analysis were short (<500 bp), which is compatible with the analysis of crossover distribution discussed above [35].

Global distribution of recombination in yeasts
The extensive molecular analyses in yeasts reveal that DSB sites cluster into chromosomal domains. For instance, on S. cerevisiae chromosome III (300 kb long), two separate domains (~50 kb each) with high DSB levels can be identified [10,11]. The influence of domains is well illustrated by the correlation between the level of DSB of an inserted sequence and the global DSB level of the domain where it is inserted [36]. Components that define these domains are unknown. In addition, centromeric regions and telomeres (10–20 kb proximal) are repressed for DSB formation.

Global distribution of recombination in mammals
High-resolution analysis of a 200 KB interval in the major histocompatibility complex (MHC) class II region reveals six hot spots (Table 2), the activities of which varies from 0.3 to 110 cM Mb⁻¹. These hot spots cluster in three regions and are separated by 50–100 kb (Fig. 3a). In these intervals, the values of recombination frequencies are a rough estimate based on linkage disequilibrium (LD) analysis. The crossovers in the hot-spot regions, which cover <10% of the whole interval, account for >95% of crossover events. This analysis illustrates clearly the high discontinuity of distribution of recombination rates and a surprisingly high level of variation in recombination rates between hot spots (up to 400 fold).

At a lower resolution, crossover distribution in the whole MHC (3.5 Mb) was analyzed using pedigree and single-sperm analysis [37]. Markers 30–200 kb apart define 30 intervals in which crossover density varies from 0.15 to 2.50 cM Mb⁻¹ with an average of 0.49 cM Mb⁻¹. This is slightly lower than the male average (0.85 cM Mb⁻¹ from [38]) (Fig. 3b). Six regions are ‘hotter’
than average, one of which includes the strongest hot-spot defined at high resolution (Dna3). At still lower resolution, the human and mouse genetic maps reveal Mb fluctuations in recombination rates [38–41]. These variations are shown for human chromosome 6, which contains the MHC region (Fig. 3c). On this chromosome, recombination frequencies were measured by direct molecular detection in the six intervals showing peaks of recombination rates (Dna1, Dna2 and Dna3 near the DNA gene, Dmb1 and Dmb2 near the DMB gene and Tap2). These values are from male meiosis. Between these sites, recombination rates were approximated from linkage disequilibrium studies. (b) Mapping crossovers on human chromosome 6 (male and female). Data are plotted with 3 Mb windows, moved by 1 Mb steps. The positions of the MHC and the centromere (c) are shown. Data reproduced with permission from: (a), [33]; (b), [37]; and (c), [38].

the correlation observed at high resolution in yeast is, therefore, not universal.

Fig. 3. Recombination hot spots on human chromosome 6. (a) High-resolution mapping of crossover within 200 kb in the major histocompatibility complex (MHC) class II region. Recombination frequencies were measured by direct molecular detection in the six intervals showing peaks of recombination rates (Dna1, Dna2 and Dna3 near the DNA gene, Dmb1 and Dmb2 near the DMB gene and Tap2). These values are from male meiosis. Between these sites, recombination rates were approximated from linkage disequilibrium studies. (b) Mapping crossovers in the MHC region by sperm typing. Each point represents the centre of the intervals (average of 12 individuals). The position of the 200 kb interval of the class II genes in (a) is shown as a solid line, the remainder of this region is shown as a dotted line. (c) Mapping crossovers on human chromosome 6 (male and female). Data are plotted with 3 Mb windows, moved by 1 Mb steps. The positions of the MHC and the centromere (c) are shown. Data reproduced with permission from: (a), [33]; (b), [37]; and (c), [38].

GC content and the distribution of recombination

The analysis of DSB distribution in S. cerevisiae shows a correlation between DSB activity and GC content: domains with high levels of DSB correspond to GC-rich domains [10,11]. Furthermore, DSB activities of a DNA fragment inserted at various locations on yeast chromosome III correlate with local GC content [44]. In humans, two LD studies indicate a correlation between high GC content and elevated recombination rates [45,46]. However, estimation of recombination rates based on LD analysis is subject to several caveats and should be taken with caution [47]. Nevertheless, the correlation between recombination density and GC is observed from the complete human genetic map [48,49], and has also been deduced from immunolocalization of recombination nodules [50]. GC-rich regions are organized into chromosome domains, also identified as isochores [51], which correspond to gene-rich R bands. Many properties distinguish GC-rich from GC-poor domains, such as replication timing and the frequency of repeated sequences [52,53]. It is important to note that although the correlation is significant, the variation in GC content accounts for only a small part of the variation in recombination rates (17%, $R^2 = 0.17$) [38]. The detailed analysis of recombination in the human MHC shows no correlation between recombination rate and GC content: the sharp increase in the proportion of GC residues in class II compared to class III genes does not cause a difference in recombination rates on either side of this transition (0.66 cM Mb$^{-1}$ in the class II region, intervals Dpb1 to Notch4, and 0.46 cM Mb$^{-1}$ in the class III region, intervals Notch4 to MicB) [37]. Furthermore, in the MHC class II region, of the three crossover clusters that have been mapped at high resolution (Dna3, Dmb2 and Tap2), only one (Tap2) corresponds to a region with a higher than average GC content. In this region, three GC peaks are located in cold intervals for recombination (B. de Massy and L. Duret, unpublished). Given that these hot spots are in regions with relatively low GC content, one can conclude that a high level of recombination activity does not require a GC-rich domain or a high local GC content. The correlation observed at high resolution in yeast is, therefore, not universal.

Whatever the correlation observed, it is important to determine whether it is either a cause or a consequence
(or both) of the recombination activity. Several studies address this issue, taking into account several parameters of genome evolution and considering either selectionist or neutralist models (reviewed in [54]). Although selectionist models related to features such as gene organization and/or chromatin properties have not been tested, two observations favor the neutralist model. First, in yeast, the correlation between DSB levels and GC content is higher when the GC content at the third base of a codon (GC3) is used as parameter, rather than GC values as a whole [55]. Second, in mice, the recent translocation of a gene into the pseudo-autosomal region correlates with an increase in GC3 from 50 to 73%. This indicates strongly that elevated GC content is a consequence and not the cause of recombination activity in this case [56]. Two processes have, therefore, been proposed to explain a neutral effect: either a mutation bias or a gene-conversion bias [55,57,58]. The hypothesis that recombination contributes to changes in GC content also needs to account for the observation that recombination events are highly localized and, therefore, explain how, for instance, a local gene-conversion bias could affect the GC composition over large chromosome domains. Given the time-scale over which these changes are expected to take place, the validation of such a hypothesis might require understanding the distribution, evolution and possible movement of recombination sites in the genome. Indeed, recombination sites raise a paradoxical situation: given the mechanism of DSBR, sequences on the active chromosome will be replaced by that of the doxical situation: given the mechanism of DSBR, sequences in the genome. Indeed, recombination sites raise a parenthesis might require understanding the distribution, could affect the GC composition over large chromosome regions, crossovers are tightly clustered and follow a correlation between DSB levels and GC content is higher when the GC content at the third base of a codon (GC3) is used as parameter, rather than GC values as a whole [55]. Two processes have, therefore, been proposed to explain a neutral effect: either a mutation bias or a gene-conversion bias [55,57,58]. The hypothesis that recombination contributes to changes in GC content also needs to account for the observation that recombination events are highly localized and, therefore, explain how, for instance, a local gene-conversion bias could affect the GC composition over large chromosome domains. Given the time-scale over which these changes are expected to take place, the validation of such a hypothesis might require understanding the distribution, evolution and possible movement of recombination sites in the genome. Indeed, recombination sites raise a paradoxical situation: given the mechanism of DSBR, sequences on the active chromosome will be replaced by that of the less active one, which would be expected to lead to the loss of initiation sites [59]. However, given the complexity of determinants that influence recombination rates described above, and that many of these determinants do not reside in the area where recombination takes place, the preferential loss of sequences on the initiating chromosome might only marginally affect its recombination activity.

Conclusions

Many aspects of the mechanism of meiotic recombination have been identified in yeast, where the initiating lesion is a DSB. Repair by the DSBR pathway leads to gene conversion and crossing-over, and repair by the SDSA pathway leads to gene conversion without crossover. The main lines of this mechanism are thought to be conserved in mammals. In *S. cerevisiae*, analysis of DSBs shows that initiation takes place in open chromatin. Genetic analysis shows that both gene conversion and crossover cluster tightly within a few kb of the DSB. Comparison of genetic and DSB data in *S. pombe* indicates a different situation in which reciprocal exchanges take place far (>10 kb) away from initiation. Although DSBs have not yet been identified in mammals, several crossover hot-spot regions are likely to correspond to preferred initiation sites. In such regions, crossovers are tightly clustered and follow a distribution in agreement with the DSBR pathway, similar to that described in *S. cerevisiae*. The heterogeneous distribution of these hot spots creates large regions with low levels of recombination separated by small intervals with hot-spot activities and results in up to thousand-fold variations in recombination rates. At the chromosome level in both yeast and mammals, a second layer of regulation of recombination is apparent from the presence of large domains with either high or low levels of recombination. One of the proposed consequences of heterogeneous distribution of recombination, which is supported by data from yeast and mice, is the increase in GC content in regions with high levels of recombination. These interpretations provide new ideas and predictions for mechanistic aspects of meiotic recombination.

Acknowledgements

I thank F. Baudat, J. Buard, L. Duret and M. Lichten for critical reading of the manuscript, and the Centre National de la Recherche Scientifique, the Commissariat à l’Energie Atomique and the Association pour la Recherche contre le Cancer for supporting my work.

References

17 Porter, S.E. et al. (1995) Genetic evidence that the meiotic recombina- tion hotspot at the HIS4 locus of Saccharomyces cerevisiae does not represent a site for a symmetrically processed double-strand break. Genetics 134, 5–19.


Isobe, T. et al. (2002) Molecular characterization of the Pb recombination hotspot in the mouse major Histocompatibility Complexity class II region. Genomics 80, 229–235


Han, L.L. et al. (2000) Unequal exchange at the charcot-marie-tooth disease type 1A recombination hot-spot is not elevated above the genome average rate. Hum. Mol. Genet. 9, 1881–1889


Khamhata, S. et al. (1996) Es recombinational hotspot in the mouse major histocompatibility complex maps to the fourth intron of the Ea gene. Genome Res. 6, 195–201


77 Sun, H. et al. (1991) Extensive 3′-overhanging, single-stranded DNA associated with the meiosis-specific double-strand breaks at the arg4 recombination initiation site. Cell 64, 1155–1161

News & Features on BioMedNet

Start your day with BioMedNet’s own daily science news, features, research update articles and special reports. Every two weeks, enjoy BioMedNet Magazine, which contains free articles from Trends, Current Opinion, Cell and Current Biology. Plus, subscribe to Conference Reporter to get daily reports direct from major life science meetings.

http://news.bmn.com

Here is what you will find in News & Features:

Today’s News
Daily news and features for life scientists.
Sign up to receive weekly email alerts at http://news.bmn.com/alerts

Special Report
Special in-depth report on events of current importance in the world of the life sciences.

Research Update
Brief commentary on the latest hot papers from across the Life Sciences, written by laboratory researchers chosen by the editors of the Trends and Current Opinions journals, and a panel of key experts in their fields.
Sign up to receive Research Update email alerts on your chosen subject at http://update.bmn.com/alerts

BioMedNet Magazine
Sign up to receive weekly email alerts at http://news.bmn.com/alerts

Conference Reporter
BioMedNet’s expert science journalists cover dozens of sessions at major conferences, providing a quick but comprehensive report of what you might have missed. Far more informative than an ordinary conference overview, Conference Reporter’s easy-to-read summaries are updated daily throughout the meeting.
Sign up to receive email alerts at http://news.bmn.com/alerts