Meiotic Recombination Hotspots of Fission Yeast Are Directed to Loci that Express Non-Coding RNA

Wayne P. Wahls1*, Eric R. Siegel2, Mari K. Davidson1
1 Department of Biochemistry and Molecular Biology, University of Arkansas for Medical Sciences, Little Rock, Arkansas, United States of America, 2 Department of Biostatistics, University of Arkansas for Medical Sciences, Little Rock, Arkansas, United States of America

Abstract

**Background:** Polyadenylated, mRNA-like transcripts with no coding potential are abundant in eukaryotes, but the functions of these long non-coding RNAs (ncRNAs) are enigmatic. In meiosis, Rec12 (Spo11) catalyzes the formation of dsDNA breaks (DSBs) that initiate homologous recombination. Most meiotic recombination is positioned at hotspots, but knowledge of the mechanisms is nebulous. In the fission yeast genome DSBs are located within 194 prominent peaks separated on average by 65-kbp intervals of DNA that are largely free of DSBs.

**Methodology/Principal Findings:** We compared the genome-wide distribution of DSB peaks to that of polyadenylated ncRNA molecules of the prl class. DSB peaks map to ncRNA loci that may be situated within ORFs, near the boundaries of ORFs and intergenic regions, or most often within intergenic regions. Unconditional statistical tests revealed that this colocalization is non-random and robust (P≤5.5×10−6). Furthermore, we tested and rejected the hypothesis that the ncRNA loci and DSB peaks localize preferentially, but independently, to a third entity on the chromosomes.

**Conclusions/Significance:** Meiotic DSB hotspots are directed to loci that express polyadenylated ncRNAs. This reveals an unexpected, possibly unitary mechanism for what directs meiotic recombination to hotspots. It also reveals a likely biological function for enigmatic ncRNAs. We propose specific mechanisms by which ncRNA molecules, or some aspect of RNA metabolism associated with ncRNA loci, help to position recombination protein complexes at DSB hotspots within chromosomes.

Introduction

Over the past decade it has become clear that non-coding RNA (ncRNA) molecules have a major role in a vast array of diverse cellular processes. Small (~20–30 nt) ncRNAs such as siRNA and miRNA can trigger the formation of euchromatin or heterochromatin, affect positively or negatively transcription, induce the deadenylation of mRNA, trigger the targeted degradation of mRNA, and regulate positively or negatively the translation of functional mRNAs (reviewed by [1,2]). Nucleotide sequence complementarity between the small ncRNAs and target molecules serves to guide various protein complexes to the appropriate targets within mRNA (e.g., for cleavage by Argonaute) [3] or within DNA of chromosomes (e.g., for heterochromatinization by RITS) [4].

Another class of ncRNA molecules, the larger mRNA-like transcripts with little or no coding potential, are ubiquitous in eukaryotes ranging from fission yeast to humans [5,6]. While the precise values are unknown and vary from organism to organism, the total number of polyadenylated ncRNAs may exceed the number of protein-coding mRNAs. With a few exceptions (e.g., [7–10]) the function of these long ncRNAs is completely obscure, but it seems almost certain that they (like small ncRNAs) will be found to play important roles in the cell [1,2]. We report here an unexpected, robust connection between such ncRNAs and meiotic chromosome dynamics.

In meiosis, a combination of crossover recombination structures (chiasmata) and sister chromatid cohesion distal to chiasmata help to align homologous chromosome pairs and ensure their proper segregation in the first meiotic division (Figure 1) [11]. Consequently, meiotic recombination is not distributed randomly, but is tightly regulated to ensure that each chromosome pair receives at least one chiasma. Furthermore, recombination is positioned preferentially at hotspots along each chromosome, but current knowledge of the mechanisms for this clustering is nebulous (reviewed by [12–14]). The DNA binding sites for some transcription factors are hotspots [15–22], but there is no obvious DNA sequence preference (i.e., specific consensus sequence) for hotspots across the genome [23,24]. The relative accessibility of DNA in regions of “open” chromatin associated with transcription might facilitate the entry of meiotic recombination enzymes [16,25–28], but open chromatin is insufficient to promote recombination [29] and some hotspots lack open chromatin [30,31]. The binding of certain transcription factors [16,22,32], chromatin remodeling by tran-
scription factors [25,33], and transcription levels [19,34,35] either influence or are essential for local hotspot activity. Despite all of these seemingly clear connections to transcription, and paradoxically, hotspots in diverse organisms tend to cluster preferentially in non-coding regions [23,24,30]. It is therefore likely that additional, yet-unidentified factors help to regulate where meiotic recombination occurs.

The meiotic recombination protein Spo11 (Rec12 in fission yeast) is a conserved, topoisomerase II-like enzyme which introduces dsDNA breaks (DSBs) that initiate recombination [36]. These meiotic DSBs have been well characterized in two highly-diverged [37] organisms, budding yeast and fission yeast, and likely reflect a common mechanism for the initiation of recombination in all eukaryotes [reviewed by [14]]. A recent analysis of DNA tiling microarrays revealed the genome-wide distribution of DSBs in fission yeast [23]. Most of the DSBs are clustered within 194 prominent peaks that are spaced on average about 65 kbp apart, and between the peaks there are few, if any, detectable DSBs. The DSB clusters are found in both coding regions (e.g., genes) and in intergenic regions (IGRs), but they localize preferentially to large IGRs. Other than this general bias towards IGRs, no discrete features which might be responsible for DSB hotspot clustering were identified [23]. We report here that the DSB hotspots are directed preferentially to loci that express polyadenylated ncRNAs. We propose specific mechanisms by which ncRNA molecules, or features of ncRNA loci, help to regulate the positioning of meiotic recombination.

**Results**

Loci that express polyadenylated, ncRNAs are embedded within DSB peaks

Our long-term interest is to define how meiotic homologous recombination becomes localized at hotspots. We therefore examined the genomic DNA sequences surrounding meiotic DSB peaks of fission yeast, and we discovered that several of the peaks encompass DNA sequences which express ncRNA molecules. For example, a prominent DSB peak within the rec7 gene encompasses three polyadenylated, transcript from opposite strand RNA molecules, tos1, tos2, and tos3 (Figure 2A). These non-coding tos transcripts are induced only in meiosis [38] and are therefore present when Rec12 (Spo11) catalyses the formation of DSBs. Similarly, some prominent DSB peaks contain prl transcripts (polyA-bearing RNA without long open reading frames [5]) (e.g., Figure 2B). Such non-coding prl transcripts are expressed in meiosis, are found within both coding regions and IGRs but localize preferentially to large IGRs, and some of them are spliced [5,39]. In other words, the distribution and developmental regulation of some polyadenylated, ncRNA molecules seem to coincide with those of prominent DSB peaks.

**Genome-wide analyses of association between prl loci and DSB peaks**

The fission yeast genome contains about 5,000 coding genes [40]. Among 987 cDNA clones of a random library derived from cells in mitosis and meiosis, 68 correspond to polyadenylated, ncRNAs of the prl class [5]. This sample is of sufficient size and complexity to be representative of the distribution and abundance of prl molecules expressed from the genome. We therefore compared the genome-wide distribution of DSB peaks obtained from microarray analysis [23] to that of genomic DNA sequences from which the representative prl molecules are transcribed [5] (Table 1). Since a subset of the DSB peaks were classified previously as being “prominent” (as opposed to “weak”) [23], we analyzed both the prominent peaks alone and all peaks together (prominent plus weak).

The average distance between prominent DSB peaks is 65 kbp and the average distance between the representative prl loci is 185 kbp. One might expect these prl loci to map, on average, about 33 kbp away from DSB peaks. However, the distribution of distances between prl loci and their neighboring DSB peaks is skewed markedly towards a much shorter distance (Figure 3A–B). Furthermore, there is a unexpectedly high frequency of perfect colocalization (Figure 2, Table 1): Of the 68 ncRNA molecules analyzed, 18 (26.5%) map entirely within prominent DSB peaks. Six additional ncRNAs (8.8%) map entirely within weak DSB peaks. Fourteen more (20.5%) fall on or near the flanks of DSB peaks (within 5 kbp). In total, about 56% of the representative ncRNA loci are associated with DSB peaks by these criteria. To analyze these data further, we calculated the unconditional genome-wide probabilities of perfect colocalization, which are the most conservative and rigorous criteria possible.

The 194 prominent DSB peaks encompass about 700 kbp (5.6%) of the genome sequence and those peaks are separated by regions essentially devoid of DSBs [23]. One can view the ncRNA molecules as having a binomial distribution with respect to landing successfully (yes or no) within a DSB peak. We therefore analyzed the data using the exact binomial test of a single proportion. The
probability of success under the null hypothesis of random placement anywhere within the genome would be 5.6\% per molecule. This yields an expected value of 3.8 out of 68 prl molecules falling within a prominent DSB peak, and 96.4\% confidence that the number of molecules within a prominent peak should be seven or less. The probability of observing 18 or more prl molecules out of 68 landing within a prominent DSB peak by chance is low (P = 2.5 \times 10^{-8}) (Figure 3C). We conclude that there is a positive, non-random association between loci which express the prl class of ncRNA molecules and the positions of prominent DSB clusters that initiate meiotic recombination.

The preceding analyses considered the 194 prominent DSB peaks that collectively contain most of the meiotic DSBs catalyzed by Rec12 (Spo11). However, there are an additional 159 “weak” DSB peaks that encompass less than 5\% of the genome and whose aggregate area integrals contribute 12\% of the total DSBs measured across the genome [23]. Within these weak peaks we found six additional ncRNA transcripts of the prl data set (e.g., Figure 2C, I). Therefore 35.3\% (24 of 68) of the prl molecules fall entirely within a DSB peak (weak or prominent). The complete set of all detectable DSB peaks (weak and prominent) map to about 10.6\% of the genome sequence [23]. Under a null-hypothesis success probability of 10.6\% per molecule, one would expect 7.2 of the 68 prl molecules to fall within a peak by chance, and have 97.4\% confidence that the number falling within a peak would be 12 or less. The probability of observing 24 or more occurrences out of 68 is low (P = 5.5 \times 10^{-8}) (Figure 3C). We conclude that there is a positive, non-random association between prl loci and DSB peaks for the entire contingent of all known DSB peaks (prominent and weak) across the genome.

Test of an alternative hypothesis: Independent localization to IGRs?

About 81\% of the prominent DSB peaks, 61\% of the weak DSB peaks, and 87\% of the representative prl loci are within IGRs [5,23]. This raises the possibility that the apparent non-random associations noted above arise simply from prl loci and DSB peaks localizing preferentially, but independently, to IGRs. To explore

Figure 2. Colocalization of DSB hotspots with ncRNA. The frequencies of DSBs for each oligonucleotide tile (circles) and DSB peaks (curves) were plotted as a function of distance and are shown relative to the positions of protein-coding ORFs (boxes) and long ncRNAs (tos and prl, arrowhead indicates poly-A tail). The ncRNAs can be found within “weak” (C, I) and “prominent” (A, B, D–H) DSB peaks. Peaks and ncRNAs can map together within ORFs (A), near ORF-IGR boundaries (D, E), or within IGRs (B, C, F–I). The DSB-associated, polyadenylated ncRNAs may be spliced (H, I) or not (A–G).

doi:10.1371/journal.pone.0002887.g002
Table 1. Genome-wide distribution of representative prl loci and their proximity to DSB peaks.

| Locus | Chr. | Position (bp) | Dir. | Prominent peaks | All peaks |
|-------|------|---------------|------|----------------|----------|
| prl19 | 1    | 0051158 to 0051583 | f    | Left 42.2 Right 13.5 | Left 42.2 Right 13.5 |
| prl37 | 1    | 0071982 to 0072380 | r    | Left 63.2 Right 7.3 | Left 3.4 Right 4.3 |
| prl03 | 1    | 0237074 to 0237848 | r    | Left 0.0 Right 0.0 | Left 0.0 Right 0.0 |
| prl27 | 1    | 0299464 to 0299610 | f    | Left 56.8 Right 78.7 | Left 42.1 Right 78.7 |
| prl05 | 1    | 0434104 to 0434843 | f    | Left 48.8 Right 2.5 | Left 2.7 Right 2.5 |
| prl47 | 1    | 0491816 to 0492201 | r    | Left 27.7 Right 165.6 | Left 27.7 Right 25.3 |
| prl50 | 1    | 0977075 to 0977568 | r    | Left 8.1 Right 169.6 | Left 8.1 Right 15.4 |
| prl58 | 1    | 1330382 to 1330528 | r    | Left 14.3 Right 85.8 | Left 14.3 Right 58.8 |
| prl57 | 1    | 232758 to 2327629 | R    | Left 0.0 Right 0.0 | Left 0.0 Right 0.0 |
| prl59 | 1    | 2362411 to 2362747 | f    | Left 16.0 Right 16.8 | Left 16.0 Right 5.0 |
| prl55 | 1    | 2416235 to 2416706 | f    | Left 16.0 Right 5.0 | Left 16.0 Right 5.0 |
| prl43 | 1    | 2432340 to 2432921 | r    | Left 5.2 Right 5.2 | Left 15.8 Right 15.8 |
| prl65 | 1    | 2472399 to 2472897 | r    | Left 18.8 Right 43.8 | Left 43.8 Right 0.0 |
| prl56 | 1    | 2885670 to 2886792 | f    | Left 6.6 Right 4.3 | Left 6.6 Right 4.3 |
| prl31 | 1    | 2975608 to 2976007 | r    | Left 0.0 Right 0.0 | Left 0.0 Right 0.0 |
| prl01 | 1    | 3003816 to 3004916 | f    | Left 27.6 Right 23.5 | Left 27.6 Right 23.5 |
| prl14 | 1    | 3190456 to 3191217 | f    | Left 28.5 Right 80.1 | Left 28.5 Right 80.1 |
| prl18 | 1    | 3699381 to 3700010 | r    | Left 118.0 Right 45.3 | Left 34.2 Right 45.3 |
| prl48 | 1    | 3942128 to 3942906 | R    | Left 0.1 Right 45.7 | Left 45.7 Right 0.1 |
| prl49 | 1    | 4008184 to 4008423 | f    | Left 0.0 Right 0.0 | Left 0.0 Right 0.0 |
| prl63 | 1    | 4008184 to 4008423 | r    | Left 0.0 Right 0.0 | Left 0.0 Right 0.0 |
| prl53 | 1    | 4008172 to 4008681 | f    | Left 0.0 Right 0.0 | Left 0.0 Right 0.0 |
| prl13 | 1    | 4259371 to 4259795 | r    | Left 13.2 Right 98.4 | Left 13.2 Right 98.4 |
| prl12 | 1    | 4439118 to 4439451 | f    | Left 0.0 Right 0.0 | Left 0.0 Right 0.0 |
| prl46 | 1    | 4651351 to 4651792 | r    | Left 211.2 Right 20.2 | Left 43.6 Right 20.2 |
| prl54 | 1    | 4820804 to 4821352 | f    | Left 29.5 Right 57.2 | Left 29.5 Right 20.7 |
| prl52 | 1    | 5023945 to 5024116 | f    | Left 108.7 Right 131.5 | Left 60.4 Right 38.3 |
| prl28 | 1    | 5441748 to 5442341 | r    | Left 0.0 Right 0.0 | Left 0.0 Right 0.0 |
| prl21 | 1    | 5454229 to 5454498 | r    | Left 2.7 Right 48.5 | Left 2.7 Right 10.9 |
| prl61 | 1    | 5546205 to 5546648 | f    | Left 39.4 Right 39.4 | n.a.² Right 39.4 |
| prl11 | 2    | 0028624 to 0028796 | f    | n.a.³ Right 97.0 | n.a.³ Right 97.0 |
| prl62 | 2    | 0227372 to 0227760 | f    | Left 0.0 Right 0.0 | Left 0.0 Right 0.0 |
| prl66 | 2    | 0547325 to 0547761 | f    | Left 18.7 Right 25.0 | Left 18.7 Right 25.0 |
| prl34 | 2    | 0835871 to 0836243 | f    | Left 0.0 Right 0.0 | Left 0.0 Right 0.0 |
| prl42 | 2    | 1032324 to 1032771 | f    | Left 64.4 Right 11.2 | Left 64.4 Right 11.2 |
| prl15 | 2    | 1329046 to 1329760 | r    | Left 57.6 Right 30.9 | Left 9.8 Right 30.9 |
| prl24 | 2    | 1497664 to 1498394 | r    | Left 38.8 Right 44.9 | Left 38.8 Right 44.9 |
| prl41 | 2    | 1652051 to 1652489 | f    | Left 76.9 Right 23.7 | Left 76.9 Right 23.7 |
| prl17 | 2    | 1777375 to 1777788 | r    | Left 23.6 Right 107.6 | Left 0.0 Right 0.0 |
| prl60 | 2    | 2005290 to 2005716 | r    | Left 50.7 Right 190.8 | Left 21.4 Right 13.6 |
| prl38 | 2    | 2186474 to 2186908 | r    | Left 232.0 Right 9.6 | Left 71.8 Right 9.6 |
| prl68 | 2    | 2204841 to 2205295 | r    | Left 2.2 Right 130.7 | Left 2.2 Right 60.6 |
| prl08 | 2    | 2485488 to 2486126 | r    | Left 10.7 Right 38.7 | Left 0.0 Right 0.0 |
| prl09 | 2    | 2485500 to 2486230 | r    | Left 10.7 Right 38.7 | Left 0.0 Right 0.0 |
| prl25 | 2    | 2567830 to 2568256 | f    | Left 35.9 Right 13.9 | Left 12.3 Right 5.2 |
| prl39 | 2    | 2804337 to 2804580 | r    | Left 0.0 Right 0.0 | Left 0.0 Right 0.0 |
this possibility further, we restricted the analysis to the 40% of the fission yeast genome comprised of IGRs, and we computed the corresponding conditional probabilities under the binomial distribution and null hypothesis of no association. DSB peaks (prominent and weak) cover 22% of DNA within IGRs, prominent peaks cover 14% of DNA within IGRs, and 59 prl loci are specific to IGRs. Twenty four of the 59 IGR-specific prl loci (41%) occur within a DSB peak (P = 9.6 × 10^{-4}), 18 of which (31%) occur within a prominent peak (P = 8.6 × 10^{-4}). We conclude that the intimate association between DSB peaks and prl loci does not arise from independently preferential localization of each to an IGR. Rather, the DSB peaks colocalize specifically with ncRNA loci that may be situated within ORFs, near ORF-IGR boundaries, or most often within IGRs (e.g., Figure 2).

**Discussion**

It is implausible, statistically, that the association between DSB peaks and ncRNAs of the prl class is coincidental (P = 5.5 × 10^{-8}, Figure 3C). The strength of this association was documented using the most strict criteria possible (perfect colocalization) and it takes into account all of those instances where DSB peaks and representative ncRNA loci do not colocalize precisely. It applies whether one examines only the prominent DSB peaks or all known DSB peaks. Moreover, we tested and rejected the hypothesis that prl loci and DSB peaks localize preferentially, but independently, to a third entity on the chromosomes. Together these findings support an unambiguous, overall conclusion—meiotic DSB hotspots of fission yeast are directed preferentially to loci that express long, polyadenylated ncRNA molecules of the prl class.

**DSB hotspots localize to ncRNA loci far more intimately than to any other known factors**

The genome-wide, preferential localization of DSB peaks to representative ncRNA loci is unconditional and robust (P ≤ 5.5 × 10^{-8}, Figure 3C). For comparison, three other factors are reported to correlate positively with the genome-wide distribution of DSB hotspots, at least under conditional parameters [23,24]. In both fission yeast and budding yeast, DSB peaks are associated with a slightly elevated G:C content (P > 0.05 to P ≤ 0.0001, depending upon window size and the other conditional parameters applied). In both yeasts there is also an association of DSB peaks with those IGRs located 5’ of protein-coding genes that are divergently transcribed (P = 0.001 to P ≤ 0.0001, from conditional assessment of IGRs alone). This correlation breaks...
down (P>0.05) in fission yeast when the size of the IGR is controlled for; such IGR size controls were not reported for budding yeast. And in fission yeast, the presumptive promoter regions of genes with DSB peaks are enriched for gene ontology (GO) terms associated with “interaction between organisms” or “transcription factor activity” (from conditional assessment of presumptive promoter regions alone, no P values reported). Among these various factors, the presence of loci expressing the polyadenylated prl class of ncRNA is by far the single best predictor of where meiotic DSB peaks localize in the genome (P=5.5x10^-8, Figure 3C).

Transcripts of the prl class are expressed in meiosis (many exclusively in meiosis) [5,38,39], so they are present at the developmental stage in which Rec12 (Spo11) catalyzes the formation of DSBs. We therefore propose that ncRNA molecules or loci of the prl class help to activate DSB hotspots. This process may be conserved, because in mice one ncRNA has been shown to map to a well-defined meiotic recombination hotspot [41] and at least one other has been implicated to do so [42].

Minimum and maximum estimates of potency

The vast majority of fission yeast ncRNAs remain undiscovered [5], so the presence and frequency of ncRNA-free DSB peaks is uninformative scientifically. We note, however, that there are enough predicted ncRNA loci [5] to populate each of the known DSB peaks. About 35% of the representative prl loci are embedded entirely within DSB peaks (Figures 2–3, Table 1), so if one excludes any possible function at distance the remaining 65% of prl loci would not be sufficient to promote recombination. However, two factors suggest that the fraction of potentially recombinogenic prl loci may be much greater than 35%. First, the distribution of DSB peaks in fission yeast was determined using rad50S strains [23], and in rad50S strains of budding yeast about half of all meiotic DSB peaks escape detection [43,44]. If this applies to fission yeast, then there would be about twice as many DSB peaks as reported. In that case, somewhere between at least 35% (the current observed value) and approximately 70% (an extrapolated value) of the representative prl loci would fall entirely within DSB peaks. Second, if one allows for possible function at even a very short distance (e.g., ≤5 kbp), then the fraction of prl loci associated with DSB peaks also increases dramatically (by 58%, Table 1). Such possible function at distance is indicated clearly by the data (means and modes, Figure 3A–B). For these reasons, the theoretical maximum value for functional association may approach unity.

Hypothetical mechanisms: Chromatin structure or guide RNA

How might long, polyadenylated ncRNA molecules or loci that express ncRNA molecules help to position the initiation of meiotic recombination catalyzed by Rec12 (Spo11)? We propose two, not mutually exclusive, hypotheses on mechanism.

A chromatin docking-site hypothesis for hotspot meiotic recombination at ncRNA loci. In fission yeast meiotically induced chromatin remodeling occurs at DSB hotspots [45,46] and where tested is seemingly required for hotspot activity [25]. Therefore one possibility is that prl-dependent chromatin remodeling, due either to some aspect of transcription at prl loci or mediated by the ncRNA molecules themselves, creates a preferential site for the nucleation of recombination protein complexes. This hypothesis fits comfortably within the prevailing orthodoxy, which posits that the relative accessibility of DNA in regions of “open” chromatin has a role in hotspot activity (reviewed by [12,14,47,48]).
A guide RNA hypothesis for hotspot meiotic recombination. Small ncRNAs (e.g., siRNA) base pair with their targets (RNA or DNA) and thereby deliver protein complexes to those targets [3,4]. Similarly, ncRNA-DNA hybrids (R-loops) formed by long ncRNAs are implicated to direct the positioning of class switch recombination during B cell maturation [10,49,50]. We propose that an analogous mechanism operates to direct the machinary of meiotic recombination to ncRNA loci within chromosomes. This may involve a homology search of DNA by protein-RNA complexes (e.g., as is implicated for the siRNA-containing RITS complex). Alternatively, it may involve the recognition of R-loop structures by protein complexes (e.g., as is implicated for class switch recombination). In either case, the base pairing between ncRNA molecules and homologous chromosomal DNA would guide recombination to hotspots.

We emphasize that our proposals are not mutually exclusive with other hypothetical mechanisms proposed previously (reviewed by [12–14]) and the models even overlap to some extent. For example, R-loops would impart changes in local nucleosome organization and alter the sensitivity of DNA within chromatin to DNA damage. The mechanism proposed here would involve the pairing between ncRNA molecules and homologous chromosomal DNA, but it is currently unknown how RNA sequences would enter the genome.

Implications and context

Our findings have two main implications. First, they reveal an unexpected, potentially unitary mechanism for which directs meiotic recombination to hotspots (expressed ncRNA loci). Second, they reveal a likely biological function for many of the enigmatic, polyadenylated ncRNA molecules that are so abundant in eukaryotes.

One of our two alternative hypotheses is that the polyadenylated ncRNA molecules help to position meiotic recombination. For context, this hypothesis is consistent with, and might explain mechanistically, some of the many connections between transcription and recombination documented previously (see Introduction). A few additional examples are provided here. Hotspot-activating RNA molecules would in principle confer differential DNA strand identity at hotspots, and hence could provide an underlying basis for the seemingly asymmetrical nature of all DSBs (directionality) [51] and for the preferential transfer of one DNA strand into heteroduplex DNA (strand identity) [52]. Similarly, our hypothesis may explain why the known RNA metabolism protein Ski8 (Rec14 of fission yeast) is an essential component of the Spo11 (Rec12) meiotic recombination protein complex in several organisms ([53–55] and our unpublished observations).

Materials and Methods

Relative proximity of DSB peaks to prl loci

Sixty-eight polyadenylated, ncRNA molecules of the prl class were identified among 987 random cDNA clones [5]. ORF [40], tos [38], and prl [5] locations were obtained from the fission yeast genome database at the Sanger institute [40]. Raw data on the genome-wide distribution of meiotic DSBs from dataset S1 [23] were analyzed with permission (Creative Commons Attribution License). Raw experimental data were averaged and divided by those of a negative, mitotic control to yield DSB intensity ratio, and peak curves for display were drawn using a sliding, 5-point average. Where available, we used the prl locations annotated in the Agilent microarray spreadsheet (Agilent Technologies). Seventeen prl loci were not annotated in that spreadsheet (prl17, prl19, prl11, prl13, prl17, prl18, prl22, prl25, prl27, prl29, prl30, prl49, prl56, prl61, prl63, prl64, prl67). These were added manually. One locus (prl52) was annotated twice (on chromosomes I and III). The position for chromosome I was used, as it is congruent with the location listed in the genome database. The distances between each prl locus and its two neighboring DSB peaks (left and right) were determined and tabulated in Excel 2005 [Microsoft Corporation, Redmond, WA]. Ends of prl molecules (loci) that map within DSB peaks were assigned distance values of zero. Because 88% of the observed DSBs (from area integrals) were classified previously as falling within prominent DSB peaks [23], we analyzed the data both for prominent DSB peaks and for all DSB peaks (prominent plus weak) (Table 1).

Statistical measures

Each locus expressing an ncRNA molecule can map either within or outside of a DSB peak. Thus, under random genome-wide placement with constant per-locus “success” probability of mapping within a peak (the null hypothesis), the set of ncRNA loci should approximate the binomial distribution with respect to DSB peaks. We therefore modeled the data using the Binomial distribution. Binomial proportion parameters (success probabilities per locus) were estimated as equal to the fraction of the genome encompassed by “all” and “prominent” DSB peaks. Conditional binomial proportion parameters (success probabilities per locus, given that the locus is within an IGR) were similarly set equal to the fraction of IGRs encompassed by “all” and “prominent” DSB peaks. For every parameter estimate, the V value for the observed number of mappings to DSB peaks was calculated using the exact binomial test of a single proportion. Exact binomial one-sided upper confidence limits were also calculated for the expected number of successes under the null hypothesis. Calculations were performed using the “BINOMDIST” function in a Microsoft Excel spreadsheet (Microsoft Corporation, Redmond, WA). The binomial-proportion estimates, sample sizes, observed success rates, null-hypothesis confidence limits, and calculated P values for each test scenario are given in the main text.

Acknowledgments

We thank Giulia Baldini, Reine Protacio, and Alan Tackett for their input.

Author Contributions

Conceived and designed the experiments: WPW MKD. Performed the experiments: WPW MKD. Analyzed the data: WPW ERS. Contributed reagents/materials/analysis tools: WPW ERS MKD. Performed the experiments: WPW MKD. Analyzed the data: WPW ERS MKD. Wrote the paper: WPW.

References

1. Nileen TW (2007) RNA 1997–2007: a remarkable decade of discovery. Mol Cell 28: 773–780.
2. Mattrick JS (2007) A new paradigm for developmental biology. J Exp Biol 210: 1526–1547.
3. Liu J, Carmell MA, Rivas FV, Thomson JM, et al. (2004) Argonaute2 is the catalytic engine of mammalian RNAi. Science 305: 1526–1547.
4. Griffiths-Jones S (2007) Annotating noncoding RNA genes. Annu Rev Genomics Hum Genet 8: 279–298.
5. Ng K, Pullirsch D, Leeb M, Wutz A (2007) Xist and the order of silencing. EMBO Rep 8: 34–39.
6. Kasama T, Shigehisa A, Hirata A, Saito TT, Tougan T, et al. (2006) Spo5/S. cerevisiae Sporulation-specific DNA repair. Mol Genet Genomics 275: 209–215.
7. Nilsen TW (2007) RNA 1997–2007: a remarkable decade of discovery. Mol Cell 28: 715–720.
8. Griffiths-Jones S (2007) Annotating noncoding RNA genes. Annu Rev Genomics Hum Genet 8: 279–298.
9. Watanabe Y, Yamamoto M (1994) S. pombe me24 encodes a RNA-binding protein essential for premeiotic DNA synthesis and meiosis I, which cooperates with a novel RNA species mrRNA. Cell 78: 487–498.
10. Yu K, Chedin F, Hirsch CL, Wilson TE, Lieber MR (2003) R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. Nat Immunol 4: 442–451.
11. Page SL, Hawley RS (2003) Chromosome choreography: the meiotic ballet. Science 301: 765–769.
12. Peters TD (2003) Meiotic recombination hot spots and cold spots. Nat Rev Genet 2: 360–369.
13. Wahl WP (1998) Meiotic recombination hotspots: Shaping the genome and insights into hypercaspary minisattelite DNA change. Curr Topics Dev Biol 37: 37–75.
14. Nishant KT, Rao MR (2006) Molecular features of meiotic recombination hot spots. Bioessays 28: 43–56.
15. Wahl WP, Smith GR (1994) A heteromorphic protein that binds to a meiotic homologous recombination hot spot: correlation of binding and hot spot activity. Genes Dev 8: 1693–1702.
16. Kon N, Krawchuk MD, Warren BG, Smith GR, Wahl WP (1997) Transcription factor Mox1/Mox2 (Atf1/Per1, Gdf7/Per1) activates the M26 meiotic recombination hotspot in S. pombe. Proc Natl Acad Sci USA 94: 13763–13770.
17. Gao J, Davidson MK, Wahls WP (2000) Distinct regions of ATP/CREB proteins ATF and Per1 control recombination hotspot ade6-M26 and the osmotic stress response. Nucleic Acids Res 28: 2830–2831.
18. Schuchert P, Langford M, Kaseln E, Kohli J (1991) A specific DNA sequence is required for high frequency of recombination in the ade6 gene of fission yeast. EMBO J 10: 2157–2163.
19. Schultes NP, Sozotov JK (1991) A poly(dA:dT) tract is a component of the recombination initiation site at the ARG4 locus in Schizosaccharomyces cerevisiae. Mol Cell Biol 11: 322–329.
20. Steiner WW, Smith GR (2005) Natural meiotic recombination hot spots in the Schizosaccharomyces pombe genome successfully predicted from the simple sequence motif M26. Mol Cell Biol 25: 9054–9062.
21. White MA, Wierfl M, Dedofil P, Peters TD (1991) DNA-binding protein RAP1 stimulates meiotic recombination at the HIS5 locus in yeast. Proc Natl Acad Sci USA 88: 9755–9759.
22. White MA, Dominka M, Peters TD (1993) Transcription factors are required for the meiotic recombination hotspot at the HIS5 locus in Schizosaccharomyces cerevisiae. Proc Natl Acad Sci USA 90: 6621–6622.
23. Cromie GA, Hyppa RW, Cam HP, Farah JA, Grewal SI, et al. (2007) A discrete class of intergenic DNA dictates meiotic DNA break hotspots in fission yeast. PLoS Genet 3: e141.
24. Gerton JL, DeRisi J, Shroff R, Licht M, Brown PO, et al. (2000) Inaugural article: global mapping of meiotic recombination hotspots and coldspots in the yeast Schizosaccharomyces cerevisiae. Proc Natl Acad Sci USA 97: 11383–11390.
25. Yamada T, Mizuno KI, Hirota K, Kon N, Wahl WP, et al. (2004) Roles of histone acetylation and chromatin remodelling factor in a meiotic recombination hotspot. EMBO J 23: 1792–1803.
26. Shenkar R, Shen MH, Arheim N (1991) DNase I-hypersensitive sites and transcription-factor-binding motifs within the M. betula meiotic recombination hot spot. Mol Cell Biol 11: 1813–1819.
27. Ohita K, Shibata T, Nicolas A (1994) Changes in chromatin structure at recombination initiation sites during yeast meiosis. EMBO J 13: 5754–5763.
28. Wu TC, Licht M (1994) Meiosis-induced double-strand break sites determined by yeast chromatin structure. Science 263: 533–534.
29. Fan QQ, Peters TD (1996) Relationship between nucleosome-hypersensitive sites and meiotic recombination hot spot activity at the HIS4 locus of Schizosaccharomyces cerevisiae. Mol Cell Biol 16: 2037–2043.
30. Myers S, Botolo L, Freeman C, McVean G, Donnelly P (2005) A fine-scale map of recombination rates and hotspots across the human genome. Science 310: 321–324.
31. Mizuno K, Koide T, Sagai T, Moriwaki K, Shiroishi T (1996) Molecular analysis of a recombinational hotspot adjacent to Lgat2 gene in the mouse MHC: fine location and chromatin structure. Mammalian Genome 7: 490–496.
32. Kon N, Schroeder SC, Krawchuk MD, Wahls WP (1998) Regulation of the Mox1/Mox2-dependent ade6-M26 meiotic recombination hotspot and development by the Spc1 mitogen-activated protein kinase of fission yeast. Mol Cell Biol 18: 7575–7583.
33. Mieczkowski PA, Dominka M, Buck MJ, Lieb JD, Peters TD (2007) Loss of a histone deacetylase dramatically alters the genomic distribution of Spo11p-catalyzed DNA breaks in Schizosaccharomyces pombe. Proc Natl Acad Sci USA 104: 3953–3960.
34. Griim C, Schaer P, Munz P, Kohli J (1991) The strong ADH1 promoter stimulates mitotic and meiotic recombination at the ade6 gene of Schizosaccharomyces pombe. Mol Cell Biol 11: 289–298.
35. Zahn-Zabal M, Lehrmann E, Kohli J (1993) Hot spots of recombination in fission yeast: inactivation of the M26 hot spot by deletion of the ade6 promoter and the novel hotspot ade4-aim. Genetics 140: 469–478.
36. Kon N, Schroeder SC, Krawchuk MD, Wahls WP (1997) Transcription factor Mox1/Mox2 (Atf1/Per1, Gdf7/Per1) activates the M26 meiotic recombination hotspot in S. pombe. Proc Natl Acad Sci USA 94: 13763–13770.
37. Wahls WP, Smith GR (1994) A heteromorphic protein that binds to a meiotic homologous recombination hot spot: correlation of binding and hot spot activity. Genes Dev 8: 1693–1702.
38. Kon N, Krawchuk MD, Warren BG, Smith GR, Wahl WP (1997) Transcription factor Mox1/Mox2 (Atf1/Per1, Gdf7/Per1) activates the M26 meiotic recombination hotspot in S. pombe. Proc Natl Acad Sci USA 94: 13763–13770.
39. Fan QQ, Peters TD (1996) Relationship between nucleosome-hypersensitive sites and meiotic recombination hot spot activity at the HIS4 locus of Schizosaccharomyces cerevisiae. Mol Cell Biol 16: 2037–2043.
40. Myerson J, Knoll M, Albermann S, Moraes CT, Schwob J, et al. (2005) A fine-scale map of recombination rates and hotspots across the human genome. Science 310: 321–324.
41. Mizuno K, Koide T, Sagai T, Moriwaki K, Shiroishi T (1996) Molecular analysis of a recombinational hotspot adjacent to Lgat2 gene in the mouse MHC: fine location and chromatin structure. Mammalian Genome 7: 490–496.
42. Kon N, Schroeder SC, Krawchuk MD, Wahls WP (1998) Regulation of the Mox1/Mox2-dependent ade6-M26 meiotic recombination hotspot and development by the Spc1 mitogen-activated protein kinase of fission yeast. Mol Cell Biol 18: 7575–7583.