Genome Biology

RESEARCH HIGHLIGHT

PRDM9 points the zinc finger at meiotic recombination hotspots

Matthew J Neale*

Abstract

Meiotic recombination events are spread nonrandomly across eukaryotic genomes in 'hotspots'. Recent work shows that a unique histone methyltransferase, PRDM9, determines their distribution.

Recombination hotspots

In most organisms, a central feature of meiosis is the induction of homologous recombination. In meiosis, homologous recombination involves searching for homologous DNA sequences on homologous chromosomes, not on sister chromatids as in somatic recombination. The result is the effective alignment of chromosome pairs (a prerequisite for their subsequent accurate segregation into separate gametes), and the reciprocal exchange of chromosomal regions between the two homologs to create new allele combinations (recombination). These rearrangements mean that each gamete contains a unique mixture of the parental alleles.

In mammals, multiple lines of investigation (studies of pedigree, linkage disequilibrium and sperm typing) show that meiotic recombination events are not uniformly distributed across the genome. Instead, large areas with low median recombination rate are punctuated by discrete 1-2 kb regions called 'hotspots' that have a much higher recombination rate (reviewed in [1]). How this distribution arose and how it is maintained are topics of great interest to genome biologists.

The punctate hotspot distribution observed in mammals is directly comparable to the detailed recombination maps also described in the budding yeast, *Saccharomyces cerevisiae*, where the molecular steps of meiotic recombination have been characterized most clearly [1,2]. In one way this similarity is unsurprising. All organisms seem to induce meiotic recombination by the same evolutionarily conserved process: DNA double-strand break (DSB) formation catalyzed by the topoisomerase-like protein Spo11 (reviewed in [1]). Thus, all meiotic recombination hotspots are also Spo11-DSB hotspots. However, what is it about Spo11-induced recombination that causes such discrete hotspots to arise?

The chromatin connection

In budding yeast, DSB hotspots generally map to short 50-200 bp regions of open chromatin found almost exclusively adjacent to transcription promoters, but with no obvious DNA sequence motif (reviewed in [1]). In humans, most hotspots are also excluded from coding regions, but rather than residing at or close to the promoter, there is a trend for human hotspots to map more distantly (about 30 kb) from the nearest transcription start site [3].

Hotspot designation appears to require post-translational chromatin modification, with trimethylation of histone H3 on lysine residue 4 (H3K4me3) being a robust identifier of hotspot activity in mouse [4]. Furthermore, disruption of Set1 (the only known H3K4 methyltransferase activity in yeast) causes dramatic changes in the pattern of hotspot usage across the yeast genome [5], suggesting that H3K4me3 is an evolutionarily conserved regulator of hotspot distribution. Precisely how the H3K4me3 mark associates with meiotic hotspots - and what its role is at these sites - are two important issues that need to be resolved.

Fascinating observations that begin to explain this complex association between the H3K4me3 mark and hotspot distribution have recently emerged from groundbreaking work recently published in *Science* [6-8].

Pointing the zinc finger

Working in mouse, the groups of Bernard de Massey and Kenneth Paigen independently identified a region on chromosome 17 that functioned *in trans* to activate hotspots in distant locations [6,7]. This *trans*-activating locus was narrowed to a tiny 181 kb region containing just four genes, of which one, *Prdm9* (also known as *Meisetz*), makes a striking candidate for a gene involved in mammalian hotspot regulation through chromatin

*Correspondence: m.neale@sussex.ac.uk
MRC Genome Damage and Stability Centre, University of Sussex, Brighton BN1 9RQ, UK

modification: *Prdm9* is expressed during early meiosis; its disruption blocks prophase progression leading to sterility; and, importantly, PRDM9 contains a conserved central domain with H3K4 methyltransferase activity [9].

However, the most intriguing feature of PRDM9 resides in its carboxy-terminal domain, which comprises a set of C2H2-type zinc-finger repeats. Within such an array, each sequential zinc finger is predicted to bind a sequential trinucleotide on a target DNA molecule, suggesting that PRDM9 could bind DNA with sequence specificity and thus influence recombination hotspot usage by local H3K4 trimethylation.

The sequence encoding the zinc-finger array of PRDM9 has a minisatellite-like structure. Each finger is encoded by 28 amino acids, with residue positions 6, 9 and 12 (-1, +3 and +6 relative to the zinc-finger alpha helix) predicted to specify DNA contacts. Sequencing a panel of *Prdm9* alleles from 20 mouse strains reveals a surprising degree of variation [7]. Not only is repeat number variable (11 to 14 repeats), but amino acid differences between the repeats are also extremely frequent. Specifically, of the 24 amino acid differences that distinguish inactive and active *Prdm9 trans*-activating alleles, 23 are found in the zinc-finger repeat, and 21 of these map to residues that are likely to control DNA binding specificity [6]. Equivalent analyses of human *Prdm9* using DNA libraries derived from multiple sources and spanning multiple ethnicities reveals a similar situation to mouse [6,7]: repeat number is highly variable (8 to 16), with most differences between repeats restricted to within the penultimate five repeats and comprising only the amino acids involved in putative DNA contacts. The extent of sequence variability and its restriction to the DNA-binding residues of the zinc-fingers is remarkable. Could these differences therefore specify hotspot usage in humans?

At the same time, a third grouping of researchers had taken an entirely different approach to investigate what designates a human recombination hotspot [8,10]. By systematically searching the Phase 2 HapMap for sequence motifs present at hotspot-associated regions, Myers and colleagues [10] identified a 13 bp degenerate motif (CCNCCNTNNCCNC) that is predicted to be critical in defining recombination activity at 40% of all known meiotic hotspots. On the basis of the wider (30-40 bp) context flanking the 13 bp motif, and of its apparent 3 bp periodicity, these authors proposed that the motif was likely to be bound by a zinc-finger protein with at least 12 fingers. Subsequent computational searching identified five candidate zinc-finger proteins, of which only one stands up to the challenge of tolerating degeneracy at positions 3, 6, 8, 9 and 12 of the target site, but not elsewhere. As you might have suspected, the identified protein is PRDM9 [8].

If the zinc-finger repeats of PRDM9 truly specify hotspot usage, then differing PRDM9 alleles should bind hotspot motifs differentially, and individuals with differing *Prdm9* alleles should show differential recombination activity across the genome. Baudat and colleagues [6] used the three *Prdm9* alleles (A, B and I) present in members of the Hutterite founder population to show that these predictions can be satisfied. (The Hutterites are a human population who went through a bottleneck in the 18th century but expanded rapidly thereafter.) Specifically, AB and AI heterozygous individuals both used significantly different hotspots from those used by the AA homozygous individuals. Indeed, it is estimated that at least 18% of the variation in hotspot usage observed among the Hutterite population can be attributed solely to PRDM9 diversity [6]. Finally, recombinant PRDM9 proteins of A or I variant bind *in vitro* to differing 13 bp motifs with the expected specificity [6].

Hotspot evolution

Humans and chimpanzees share about 99% sequence identity at aligned bases yet seem to share very few, if any, hotspot locations (discussed in [8]). This observation suggests that recombination hotspots are evolving far more rapidly than are the underlying sequence determinants. Indeed, of 22 inferred human hotspot loci at which there is also conservation of the 13 bp motif in both species, only one revealed evidence for conserved usage in the chimpanzee [8]. Given that PRDM9 is thought to target recombination specifically to the 13 bp motif via its zinc-finger array, it is logical to ask whether chimpanzee PRDM9 is really expected to bind the same motif. In fact, compared with human PRDM9, all but the first of the repeats from chimpanzee differ at amino acid positions critical to DNA binding [8,11]. Thus, chimpanzee PRDM9 is indeed expected to bind a DNA motif unrelated to that of humans [8]. A critical question remains as to whether or not this chimpanzee-specific motif associates with chimpanzee hotspots.

Comparisons of the repeat structure preserved between other metazoans indicate that accelerated evolution of *Prdm9* is a universal feature [11]. It is interesting to consider what is driving such rapid evolution of the *Prdm9* repeat. The very nature of the repeat structure (it is a coding minisatellite) may make it unstable and prone to alteration via slippage of the replication machinery. This, however, cannot explain the positive selection for amino acid changes that confer differential DNA binding specificities. One intriguing idea is that because hotspot motifs are prone to loss via biased gene conversion, there may be selective pressure for mechanisms that generate new hotspot activities (see [6,8,11] for these and alternative considerations).

A dramatic feature exposed by these studies [6-8] is of the relative fluidity with which recombination distributions can be altered by combining in one protein (PRDM9) an epigenetic marking activity (H3K4me3) with a rapidly diverging DNA binding domain. Yet many interesting questions are unresolved. What is the molecular function of the H3K4me3 mark? What is the significance of the DNA binding specificity of PRDM9, and why is it evolving so rapidly? Does PRDM9 specify all or just some hotspots? And, what effect would expression of a generic or 'zinc-fingerless' *Prdm9* allele have on recombination distributions? The answers to these questions will significantly advance our understanding of recombination hotspots and how they have evolved.

Acknowledgements

MJN is supported by a University Research Fellowship from the Royal Society and a New Investigator Grant from the Medical Research Council.

Published: 26 February 2010

References

1. Kauppi L, Jeffreys AJ, Keeney S: **Where the crossovers are: recombination distributions in mammals.** *Nat Rev Genet* 2004, **5:**413-424.
2. Roig I, Keeney S: **Probing meiotic recombination decisions.** *Dev Cell* 2008, **15:**331-332.
3. Myers S, Bottolo L, Freeman C, McVean G, Donnelly P: **A fine-scale map of recombination rates and hotspots across the human genome.** *Science* 2005, **310:**321-324.
4. Buard J, Barthes P, Grey C, de Massy B: **Distinct histone modifications define initiation and repair of meiotic recombination in the mouse.** *EMBO J* 2009, **28:**2616-2624.
5. Borde V, Robine N, Lin W, Bonfils S, Geli V, Nicolas A: **Histone H3 lysine 4 trimethylation marks meiotic recombination initiation sites.** *EMBO J* 2009, **28:**99-111.
6. Baudat F, Buard J, Grey C, Fledel-Alon A, Ober C, Przeworski M, Coop G, de Massy B: **PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice.** *Science* 2009, **327:**836-840.
7. Parvanov ED, Petkov PM, Paigen K: **Prdm9 controls activation of mammalian recombination hotspots.** *Science* 2009, **327:**835.
8. Myers S, Bowden R, Tumian A, Bontrop RE, Freeman C, Macfie TS, McVean G, Donnelly P: **Drive against hotspot motifs in primates implicates the PRDM9 gene in meiotic recombination.** *Science* 2009, **327:**876-879.
9. Hayashi K, Yoshida K, Matsui Y: **A histone H3 methyltransferase controls epigenetic events required for meiotic prophase.** *Nature* 2005, **438:**374-378.
10. Myers S, Freeman C, Auton A, Donnelly P, McVean G: **A common sequence motif associated with recombination hot spots and genome instability in humans.** *Nat Genet* 2008, **40:**1124-1129.
11. Oliver PL, Goodstadt L, Bayes JJ, Birtle Z, Roach KC, Phadnis N, Beatson SA, Lunter G, Malik HS, Ponting CP: **Accelerated evolution of the Prdm9 speciation gene across diverse metazoan taxa.** *PLoS Genet* 2009, **5:**e1000753.

doi:10.1186/gb-2010-11-2-104
Cite this article as: Neale MJ: **PRDM9 points the zinc finger at meiotic recombination hotspots.** *Genome Biology* 2010, **11:**104.