DNA Sequence-Mediated, Evolutionarily Rapid Redistribution of Meiotic Recombination Hotspots

Commentary on Genetics 182: 459–469 and Genetics 187: 385–396

Wayne P. Wahls[1] and Mari K. Davidson
Department of Biochemistry and Molecular Biology, University of Arkansas for Medical Sciences, Little Rock, Arkansas 72205-7199

ABSTRACT Hotspots regulate the position and frequency of Spo11 (Rec12)-initiated meiotic recombination, but paradoxically they are suicidal and are somehow resurrected elsewhere in the genome. After the DNA sequence-dependent activation of hotspots was discovered in fission yeast, nearly two decades elapsed before the key realizations that (A) DNA site-dependent regulation is broadly conserved and (B) individual eukaryotes have multiple different DNA sequence motifs that activate hotspots. From our perspective, such findings provide a conceptually straightforward solution to the hotspot paradox and can explain other, seemingly complex features of meiotic recombination. We describe how a small number of single-base-pair substitutions can generate hotspots *de novo* and dramatically alter their distribution in the genome. This model also shows how equilibrium rate kinetics could maintain the presence of hotspots over evolutionary timescales, without strong selective pressures invoked previously, and explains why hotspots localize preferentially to intergenic regions and introns. The model is robust enough to account for all hotspots of humans and chimpanzees repositioned since their divergence from the latest common ancestor.

> The conflict between the evolutionary persistence of hotspots and the instability intrinsic to their mode of action implies a deep flaw in our understanding of the mechanism of meiotic recombination (Pineda-Krch and Redfield 2005, p. 2321).

> ...Homologous recombination may be regulated primarily by a finite number of discrete DNA sites and proteins that interact with those sites (Wahls and Smith 1994, abstract).

Meiotic Recombination Hotspots

In meiosis, crossover recombination structures help to align paired homologous chromosomes on the metaphase plate of the first meiotic division, and this alignment is required for the faithful segregation of homologs (Gerton and Hawley 2005). Meiotic recombination is clustered at hotspots that regulate its frequency and distribution along chromosomes. Jürg Kohli's laboratory discovered DNA sequence-dependent activation of recombination hotspots in fission yeast ~20 years ago (Schuchert *et al.* 1991). At about the same time, Tom Petes' laboratory provided evidence for such regulation in budding yeast (White *et al.* 1991, 1993; Fan *et al.* 1995), which is highly diverged from fission yeast. The yeast paradigms long stood alone, but we now know that DNA sequence elements also help to position meiotic recombination at hotspots in mammals (Myers *et al.* 2008; Baudat *et al.* 2010). In each case, it appears that sequence-specific DNA-binding proteins trigger epigenetic modifications of chromatin structure that help to regulate the initiation of recombination by Spo11 (Rec12) (Kon *et al.* 1997; Yamada *et al.* 2004; Hirota *et al.* 2007; Buard *et al.* 2009; Baudat *et al.* 2010; Myers *et al.* 2010; Parvanov *et al.* 2010).

Three approaches have been used to discover regulatory (hotspot) DNA sequence motifs. The first approach has been to map hotspot locations genetically and then, by using scanning base-pair substitution mutagenesis in the genome,

Copyright © 2011 by the Genetics Society of America
doi: 10.1534/genetics.111.134130
Available freely online through the author-supported open access option.
Supporting information is available online at http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.111.134130/-/DC1.
[1]Corresponding author: Department of Biochemistry and Molecular Biology, University of Arkansas for Medical Sciences, 4301 West Markham St. (slot 516), Little Rock, AR 72205-7199. E-mail: wahlswaynep@uams.edu

to define the DNA sequence(s) required for activity (Gutz 1971; Szankasi *et al.* 1988; Schuchert *et al.* 1991). This approach is laborious and is practical only in model organisms such as yeast. Indeed, the scanning mutational, "gold standard" approach for documenting unambiguously the DNA sequence dependence of hotspots has, to our knowledge, been carried out only in fission yeast (Schuchert *et al.* 1991; Steiner *et al.* 2009, 2011).

The second approach, made possible by high-resolution mapping of hotspot positions across sequenced genomes (*e.g.*, Gerton *et al.* 2000; Myers *et al.* 2005; Ptak *et al.* 2005; Winckler *et al.* 2005; Cromie *et al.* 2007), has been to search computationally for motifs that are nonrandomly associated with hotspots. This method identified a consensus motif that is present at a subset of meiotic crossover hotspots (COH) and at hotspots of nonallelic homologous recombination (NAHR) in humans (Myers *et al.* 2008). Nucleotide polymorphisms within the motif correlate with attenuated hotspot activity, providing evidence that the motif is indeed recombinogenic.[1] Closely related hotspot-associated motifs were subsequently detected in humans and mice (Baudat *et al.* 2010; Kong *et al.* 2010; Myers *et al.* 2010). One caveat is that correlation does not demonstrate causation, and there is at least one example where a hotspot motif predicted computationally (Blumental-Perry *et al.* 2000) was dispensable for hotspot activity when tested experimentally (Haring *et al.* 2004). Computational searches are prone to false-negative results, too, and can miss (Cromie *et al.* 2007) motifs that are known to be recombinogenic and that are associated with >20% of hotspots throughout the genome (Schuchert *et al.* 1991; Wahls and Davidson 2010).

Recognizing that hotspot motifs are likely prevalent but elusive, Walter Steiner's group developed a third, "brute force" biological approach for motif discovery (Steiner *et al.* 2009, 2011). They screened individually the frequency of meiotic recombination in ~46,000 fission yeast strains harboring short, randomized nucleotide sequences within a test locus. A subset of candidate motifs were subsequently tested again using the rigorous criteria base-pair substitution mutagenesis. The authors showed that at least five distinct DNA sequence motifs activate hotspots, and they provided compelling evidence that there are many more regulatory motifs yet to be discovered.

The GENETICS articles by Steiner *et al.* (2009, 2011) deserve special mention, both for the insightful experimental approach and for the implications of the findings. The multiplicity of *cis*-acting regulatory elements is striking, as is the fact that very different motifs and binding proteins can function redundantly to promote recombination. Such findings support the idea that discrete DNA sites regulate much, perhaps most, meiotic recombination. They also provide a fresh new perspective with which to consider published data and long-extant puzzles, as described below.

[1]Experimental data support the association data. Tandem copies of a hypervariable minisatellite (SAT) sequence, and its binding proteins, promote homologous recombination in cultured cells (Wahls *et al.* 1990, 1991; Wahls and Moore 1998). If one allows for a 1-bp gap, there is perfect identity between the SAT sequence (5′-CCACC–TGCCCACCTCT-3′) and the conserved positions within the COH/NAHR consensus motif (5′-CCNCCNTNNCCNC-3′). Given circular permutation of tandem repeats, additional alignments can be made.

Hotspot Paradox

About 40 years ago Herbert Gutz (1971) described the fundamental characteristics of meiotic recombination hotspots. First, hotspots are regulated in *cis* because they are allele specific and display Mendelian inheritance. Second, a given hotspot allele promotes recombination in only a subset of meioses. Third, a chromosome that harbors an activated hotspot serves preferentially as the recipient of genetic information from the homologous chromosome (gene conversion), and a subset of gene conversions is accompanied by crossing over (Gutz 1971; Schuchert and Kohli 1988). Consequently, when heterozygous, the chromosome region harboring the hotspot is preferentially converted into a hotspot-inactive state (Figure 1A). The conversion rate varies according to hotspot, with at least 1% of meiotic products being converted at highly active hotspots of mice and yeast (Grimm *et al.* 1994; Cromie *et al.* 2005; Guillon *et al.* 2005). Hotspots therefore seed their own destruction and on the evolutionary timescale should be lost from the population.

The genomic distribution of hotspots varies markedly between closely related taxa (Ptak *et al.* 2005; Winckler *et al.* 2005), by ~50% or more between species of the same genus (Tsai *et al.* 2010) and, to a lesser extent, even between members of the same species (Kong *et al.* 2010), illustrating the evolutionary transience of hotspots. These changes occur rapidly, and even in humans one can chart the eventual death of individual hotspots (Jeffreys and Neumann 2009). Nevertheless, recombination hotspots remain abundant in sexually active eukaryotes. Therein lies the "hotspot paradox" (Boulton *et al.* 1997). Individual hotspots are suicidal, but, collectively, hotspots are somehow maintained. Moreover, the dynamic, evolutionarily rapid redistribution of hotspots requires that the mechanisms for replacement be facile and relatively plastic with regard to chromosomal location. We suggest that the mechanisms are coupled to, and can be explained fully by, the DNA sequence-dependent regulation of recombination hotspots.

Equilibrium Dynamics of Mutations and Gene Conversion: A Model

Two mechanisms have been shown experimentally to remove and add recombination hotspots in the genome. These are gene conversion (Gutz 1971) and base-pair substitutions (Szankasi *et al.* 1988; Schuchert *et al.* 1991). Below we describe evidence and models for how these mechanisms participate in a self-regulating, dynamic equilibrium that helps to maintain and reposition hotspots in the genome over time.

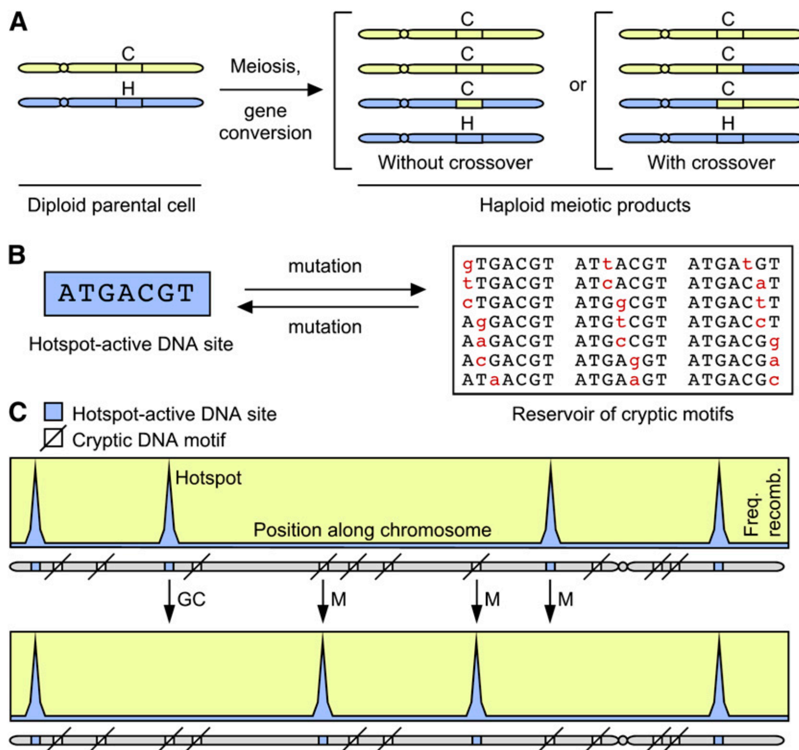

Figure 1 Model for the evolutionarily rapid redistribution of meiotic recombination hotspots. (A) Hotspots ("H") act as recipients of genetic information during gene conversion, leading preferentially to loss of the hotspot ("C," cold). (B) For every recombination-promoting DNA sequence motif [the *M26* DNA site of fission yeast is illustrated (Schuchert *et al.* 1991)], there is a reservoir of cryptic motifs. (C) Over time, DNA sequence-dependent hotspots are rendered inactive by mutation ("M") or gene conversion ("GC"). Hotspots arise *de novo* when mutations change cryptic DNA sequence motifs into hotspot motifs. Consequently, a small number of base-pair changes can dramatically alter the distribution of hotspots in the genome.

There are now at least 10 different DNA sequence motifs demonstrated by base-pair mutagenesis (5 in fission yeast), implicated by deletion studies (3 in budding yeast), or inferred from association studies (1 class each in humans and mice) to regulate hotspot activity (Schuchert *et al.* 1991; White *et al.* 1993; Myers *et al.* 2008; Steiner *et al.* 2009; Baudat *et al.* 2010). Each DNA sequence motif is short, commensurate with the molecular determinants for binding of a hotspot-activating protein or complex. For example, only 7 bp are required for hotspot activity of the *M26* DNA site in fission yeast (Schuchert *et al.* 1991), and only 8 bp of the 13-bp human crossover consensus motif are conserved (Myers *et al.* 2008). Notably, base-pair substitutions that create or ablate discrete DNA sites can generate and abolish hotspot activity, respectively (Schuchert *et al.* 1991; Steiner and Smith 2005; Steiner *et al.* 2009, 2011). Thus, within the genome is a collection of hotspot-active DNA sites and a reservoir of "cryptic" DNA sequence motifs that can be rendered active by as little as a single-base-pair substitution (Figure 1B).

A model for evolutionarily rapid redistribution of meiotic recombination hotspots is presented in Figure 1C. Individual, DNA sequence-dependent hotspots are inactivated by mutations within the DNA site or by gene conversion in meiosis. Opposing this trend is the *de novo* generation of recombinogenic DNA sites by mutations within cryptic DNA sequence motifs. In contrast, previous models suggested that *cis*-acting mutations cannot compensate successfully for the loss of hotspots by gene conversion (Boulton *et al.* 1997; Pineda-Krch and Redfield 2005; Coop and Myers 2007; Peters 2008). The distinction (and major conceptual shift) between models lies in the nature and density of DNA sequence motifs within the genome, which are factors not considered in the prior reports. (In a subsequent section, "*As Easy as A-G-C-T?*", we describe a third class of model that, like ours, involves DNA sequence motifs.)

The frequency with which spontaneous mutations generate hotspots *de novo* is likely high enough to support rapid evolutionary change because the reservoir of cryptic DNA sequence motifs is vast. For example, every regulatory DNA site 7 bp in length has 21 cryptic permutations that might be rendered hotspot active by a single-base-pair substitution (Figure 1B). These occur on average once every 780 nucleotides along each strand of DNA in the genome ($4^7 \div 21$, assuming random sequence DNA). In fission yeast (Schuchert *et al.* 1991; Steiner *et al.* 2009), in budding yeast (White *et al.* 1991, 1993; Fan *et al.* 1995), and likely in humans (Myers *et al.* 2008; Berg *et al.* 2010), multiple different DNA sequence motifs are recombinogenic, and each of those motifs has a corresponding cryptic reservoir (*e.g.*, Table S1). If one considers the recombinogenic DNA sequence motifs already defined experimentally in fission yeast (Schuchert *et al.* 1991; Steiner *et al.* 2009, 2011), there is on average a cryptic, single-base-pair variant DNA element about every 194 bp along the genome (Table S1). In other words, ~0.17% (1/582) of spontaneous mutations will generate a DNA sequence motif already known to be recombinogenic. This value calculated from the experimentally defined motifs sets the lower bound because there are additional recombination-promoting DNA sequences of fission yeast whose functional motifs remain to be defined by base-pair mutagenesis—and still more are predicted

statistically (Steiner *et al.* 2009). Given the surprisingly high frequency with which hotspot motifs are created (at least 0.17% per mutation), and that there are many intervening mitoses for each meiosis, the rate at which hotspots are created *de novo* by mutations might offset the rate at which they are lost through meiotic gene conversion. Further evidence supporting this model and its applicability to primates and to evolutionary timescales are described below.

Why Do Hotspots Avoid ORFs?

Another long-standing enigma is why recombination hotspots are located preferentially within intergenic regions (IGRs) and occur much less frequently within the ORFs of protein-coding genes (*e.g.*, Nicolas *et al.* 1989; Baudat and Nicolas 1997; Gerton *et al.* 2000; Buhler *et al.* 2007; Cromie *et al.* 2007; Frazer *et al.* 2007; Robine *et al.* 2007). This is apparently not due to avoidance of genes, but rather reflects avoidance of protein-coding regions. Hotspots do occur within genes, but when they do so, they are more prevalent within introns than within exons (Kong *et al.* 2010). Similarly, while hotspots are underrepresented in transcribed regions coding for proteins, they are abundant in transcribed regions that produce long, polyadenylated, noncoding RNAs (Wahls *et al.* 2008). We suggest that the regulation of hotspots by short DNA sequence motifs provides a mechanism for these phenomena.

The natural inclination is to ask, what mechanisms direct hotspots preferentially toward noncoding regions? From this perspective our DNA sequence-dependent model might seem unsatisfactory, because if mutations stochastically "sprinkle" hotspot motifs into the genome over time, then one would expect hotspots to arise with equal probability in coding and noncoding regions. One possibility is that natural selection favors newly arising hotspot motifs in noncoding regions (see below). However, one can ask the same question in a different way: what mechanisms might direct hotspots preferentially *away* from coding regions? This simple change in perspective, applied to existing data, revealed a causal link between molecular mechanisms for hotspot genesis and molecular mechanisms for negative selective forces (reduced organismal fitness) that help to drive the localization of hotspots (Figure 2).

Mutation of cryptic DNA sequence motifs within ORFs can generate hotspot-active DNA sites (Schuchert *et al.* 1991; Virgin *et al.* 1995; Steiner *et al.* 2009), but more often than not such mutations will also alter the sequence of the encoded protein or lead to premature termination of protein synthesis. For example, the single-base-pair substitution that created an *M26* hotspot DNA site in the *ade6* gene of fission yeast (Schuchert *et al.* 1991) also introduced a stop codon (Szankasi *et al.* 1988). This conferred a decrease in fitness because the cells can no longer grow unless adenine is added to the culture medium (Gutz 1971). Therefore natural selection, against the mutated proteins, would disfavor or eliminate the majority of hotspot-proficient DNA sites that arise within ORFs. In contrast, mutations that create hotspot-active DNA sites within IGRs or introns would not trigger negative selection due to mutated proteins. As DNA sequence-regulated hotspots arise *de novo* in the genome via mutation (a stochastic process), differential selective pressures would subsequently drive their localization away from coding regions and hence toward IGRs and introns. Indeed, the majority of DNA sequence motifs known to activate hotspots are found preferentially within noncoding regions (Steiner *et al.* 2011), providing evidence that such a drive operates across the genome and supporting our model. Interestingly, the bias is greater for motifs that are active than for those that are inactive, suggesting that additional factors influence the function of hotspot motifs, the

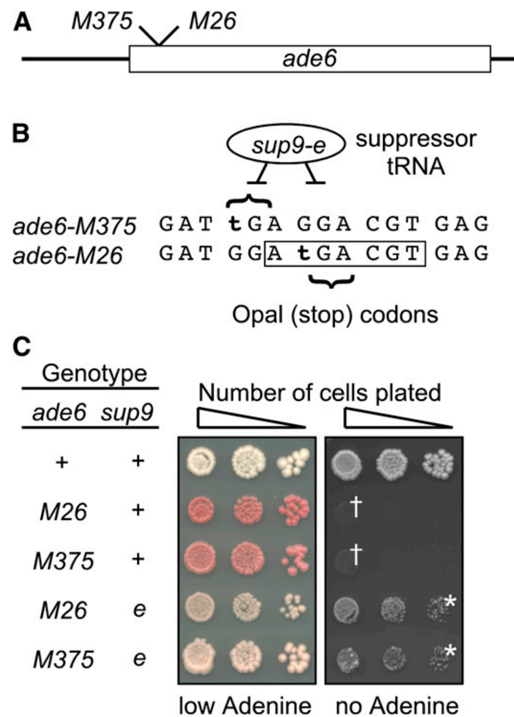

Figure 2 Hotspot-activating mutations within ORFs can, coincidentally, decrease the overall fitness of the organism. (A) Diagram of the fission yeast *ade6* locus with positions of alleles in the ORF. (B) The nonsense mutation (lowercase letters) that created the *ade6-M26* allele (Szankasi *et al.* 1988) simultaneously created a 7-bp DNA site (box) that promotes recombination (Schuchert *et al.* 1991). A similar mutation that created the *ade6-M375* allele generated neither a hotspot motif nor hotspot activity. (C) To illustrate published findings (Gutz 1971), we plated serial dilutions of cells on medium with and without adenine. The observed decrease in organismal fitness (single dagger) is due to the opal (stop) codons because it can be alleviated by a suppressor tRNA (asterisk). Furthermore, the decrease in fitness is unrelated to the hotspot motif or to hotspot activity because the hotspot is active in the presence of the suppressor tRNA (Goldman and Smallets 1979). Thus, while the decrease in fitness and the hotspot share a common origin, they are in fact independent (coincidental) consequences of the single-base-pair substitution. And because the *de novo* hotspot motif is tightly linked to the stop codon, negative selective forces that operate due to and upon the stop codon will also affect the hotspot motif. By these molecular mechanisms, natural selection disfavors most hotspots that arise by mutations in protein-coding regions, relative to those that arise in noncoding regions.

selective forces that shape their dynamic repositioning over time, or both.

Our model predicts that a fraction of hotspot-generating mutations within protein-coding regions, namely translationally silent mutations and occasionally missense mutations, would be tolerated. Experimental and correlative data are consistent with these predictions. First, the negative selective forces elicited by a nonsense mutation (coincident with generation of a hotspot motif) can be uncoupled from hotspot activity by a nonsense suppressor tRNA (Figure 2) (Gutz 1971; Goldman and Smallets 1979). Second, while DNA sequence-dependent hotspots are found preferentially in noncoding regions, they are also present in coding regions (Steiner and Smith 2005; Wahls and Davidson 2010; Steiner *et al.* 2011).

We note that negative selection against *de novo* hotspot motifs (Figure 2) need not be restricted to protein-coding regions. In principle, any noncoding region of the genome whose primary DNA sequence is important functionally would be similarly constrained as a target for the evolutionary retention of *de novo*, sequence-dependent hotspots. Well-defined examples of noncoding, sequence-constrained features include centromeres, telomeres, and silent mating-type loci, each of which is depleted for recombination (Choo 1998; Petes 2001). Parenthetically, regional variation of hotspot motif density is not the only factor that enhances or attenuates recombination regionally. Additional factors, such as centromeric heterochromatin, can actively suppress the initiation of recombination (Robine *et al.* 2007; Ellermeier *et al.* 2010). As another example, the linear element protein Rec10 (orthologous to synaptonemal complex protein Red1) can suppress the function of some hotspot motifs (Pryce *et al.* 2005), perhaps by sequestering motifs from their binding proteins, or from the recombination machinery, or both.

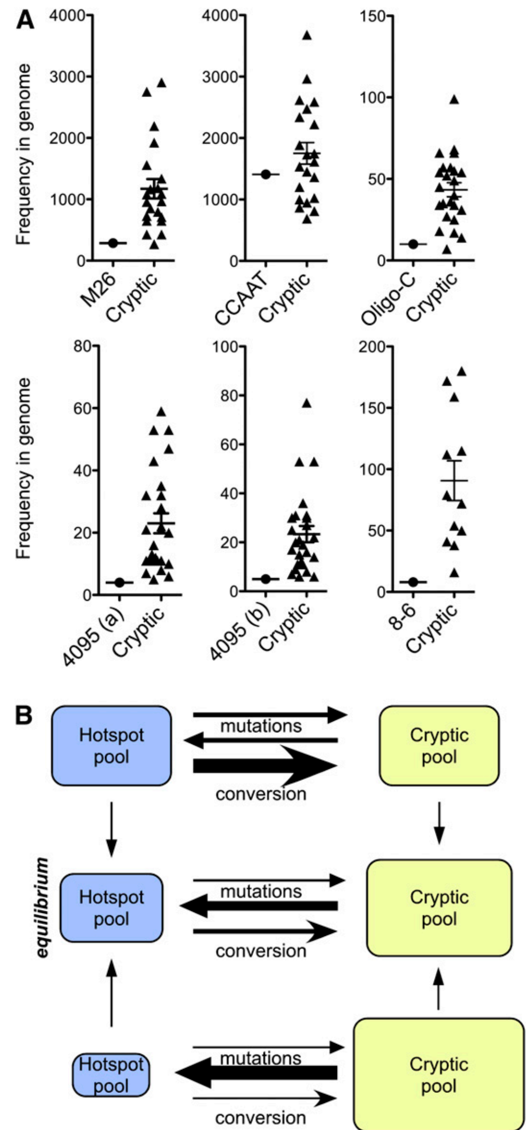

Figure 3 Equilibrium kinetics of DNA sequence-dependent hotspots. (A) Plots show frequencies of hotspot DNA sequence motifs and each single-base-pair variant ("Cryptic") motif in the fission yeast genome (Table S1). Bars indicate mean ± SEM. (B) Hotspot and cryptic motifs are interchanged by mutations and gene conversion. The ratio of motif frequencies (indicated by box sizes) affects rate vectors, driving the system to equilibrium.

Snapshot in Evolutionary Time

Further support of the models can be found in the sequence of the fission yeast genome, which reflects the sum total of dynamic changes in all preceding mitoses and meioses. Mutations are stochastic, so the rates at which any given DNA site is created or ablated by mutation should be equivalent. However, DNA sites that activate recombination hotspots are removed preferentially by gene conversion when heterozygous (Gutz 1971; Schuchert *et al.* 1991) and hence should be lost from the genome over successive generations. This process is evident because all of the recombinogenic DNA sequence motifs are underrepresented in the genome, relative to the mean frequencies of corresponding single-base-pair variants (cryptic motifs) (Figure 3A). There are several implications.

First, on the laboratory-experimental (Schuchert *et al.* 1991; Steiner *et al.* 2009) and inferred-historical (Figure 3A) timescales, multiple, distinct DNA sequence motifs of fission yeast promote meiotic recombination and are suicidal. DNA sequence-dependent hotspots of mammals, inferred from association studies, are likewise suicidal (Jeffreys and Neumann 2009; Myers *et al.* 2010).

Second, despite their suicidal tendencies, recombination-promoting DNA sequence motifs remain present (Figure 3A) and active (Schuchert *et al.* 1991; Steiner and Smith 2005; Steiner *et al.* 2009; Wahls and Davidson 2010) in the fission yeast genome. Such motifs must have arisen, and presumably continue to arise over time, by mutations (Figure 1). This rationale applies equally well for motif-dependent, suicidal hotspots of other eukaryotes.

Third, the persistence of hotspot motifs in the face of experimentally documented loss and gain requires, necessarily, a balance between rates of loss and gain. A mechanism to

regulate this balance (Figure 3B) is described in the next paragraph.

Fourth, it has been assumed *a priori* that strong positive selection is required to maintain the presence of hotspots in the genome (Boulton *et al.* 1997; Pineda-Krch and Redfield 2005; Coop and Myers 2007; Peters 2008; Ubeda and Wilkins 2011). However, the need to invoke such selection is attenuated if one considers that short DNA sequence motifs regulate recombination. Gene conversion removes hotspots over time, so the frequency of any given hotspot motif in the genome is lower than the average frequency of corresponding cryptic motifs (Figure 3A). Rate kinetics come into play. When the frequency ratio of hotspot-to-cryptic motifs (average frequency) is high (*e.g.*, 1:1), then mutations would have a negligible net effect on motif frequencies and conversion will preferentially reduce the frequency of hotspot motifs (Figure 3B, top). When the ratio of hotspot-to-cryptic motifs is low (*e.g.*, 1:10), then mutations will preferentially increase the frequency of hotspot motifs (Figure 3B, bottom). Thus, substrate concentration-dependent rate kinetics of conversion and mutation would drive the system to equilibrium. Even if a hotspot motif is eliminated from the genome, it would ultimately be resurrected from the vast pool of cryptic motifs by the inexorable stochastic process of mutation (Figures 1B and 3B). Speculatively, first-order rate kinetics could provide the primary force for retention (and commensurate repositioning) of DNA sequence-regulated hotspots over evolutionary timescales.

Selective Pressures and Population Genetics

We do not mean to imply, in the preceding section, that natural selection has no role in the positioning or evolutionary maintenance of hotspots. Indeed, negative selection demonstrably impinges upon some *de novo* hotspots (Figure 2). Additional forces, selective and nonselective, likely contribute to hotspot dynamics and a subset of such forces is described here.

Meiotic recombination is broadly conserved and, with few exceptions, crossover recombination is required for the faithful segregation of homologs in the first meiotic division (Gerton and Hawley 2005). However, while natural selection operates to maintain recombination, mathematical modeling indicated that the known benefits of recombination (on fertility and viability) are insufficient to maintain hotspots in the face of their loss by gene conversion (Boulton *et al.* 1997; Pineda-Krch and Redfield 2005; Peters 2008). Furthermore, in the absence of crossover interference recombination rates can be titrated down by >10-fold, to approximately one crossover per chromosome pair, before there is a perceptible decrease in fitness due to aberrant chromosome segregation (Kan *et al.* 2011). Two possibilities that might explain this exist. First, selection for meiotic recombination *per se* has little or no role in maintaining hotspots. Second, selection for recombination has a key role in maintaining hotspots, but we have not yet identified the benefits of having multiple recombination events (active hotspots) on each chromosome in each meiosis. Somewhere between these extremes, natural selection might operate through a combination of *cis*- and *trans*-acting factors (Peters 2008; Ubeda and Wilkins 2011), including recombinogenic DNA sites and their binding proteins.

The DNA-binding protein Prdm9 (Meisetz) is implicated to be a chromatin-remodeling transcription factor (Hayashi *et al.* 2005), a "species-incompatibility" protein involved in hybrid sterility (Mihola *et al.* 2009), and a hotspot-activating protein (Baudat *et al.* 2010; Myers *et al.* 2010; Parvanov *et al.* 2010). Interestingly, its DNA-binding domain (and hence its hotspot motif selectivity) is evolving rapidly, and there is apparently positive selection for newly arising variants, at least in some taxa (Oliver *et al.* 2009; Thomas *et al.* 2009; Ponting 2011). Therefore, selective pressures that drive the rapid evolution of Prdm9 likely help to shape the recombination landscape by changing where Prdm9 promotes recombination, without necessarily changing the overall number of Prdm9-dependent hotspots. It has been suggested, conversely, that selection for the recombination-promoting functions of Prdm9 might help to drive its rapid evolution (Ponting 2011). This idea seems plausible but tentative, particularly in the context of points raised in the preceding paragraph. At issue is what the actual selective forces are and whether any of them operate via recombination.

Boulton *et al.* (1997) pointed out that hotspots might be maintained in part by selection for aspects of cellular physiology other than recombination. This is illustrated well by the Atf1-Pcr1 heterodimer that, like Prdm9, is both a hotspot-activating factor and a transcription factor (Wahls and Smith 1994; Shiozaki and Russell 1996; Wilkinson *et al.* 1996; Kon *et al.* 1997). In its latter role, the Atf1-Pcr1 heterodimer regulates the induced transcription of core environmental stress response genes required for cells to survive under a wide variety of different stress conditions (Chen *et al.* 2003; Davidson *et al.* 2004). The decrease in fitness observed in mutants lacking this protein complex is attributable to defects in transcription. Notably, the recombination-promoting activity of the Atf1-Pcr1 heterodimer maps to a different domain of Atf1 than that required for fitness under stress (Gao *et al.* 2008). Such findings support mechanistically the insight of Boulton *et al.* (1997). One is left with the question of whether there is any selection for the recombination-promoting activities of proteins such as Prdm9 and the Atf1-Pcr1 heterodimer. Intuition would suggest that such forces exist, even though they have so far eluded detection.

All sequence-specific DNA-binding proteins known or implicated to activate hotspots are also transcription factors (White *et al.* 1991, 1993; Wahls and Smith 1994; Kon *et al.* 1997; Steiner *et al.* 2009, 2011; Baudat *et al.* 2010; Myers *et al.* 2010; Parvanov *et al.* 2010). Discussions of selective forces that might operate upon their DNA-binding sites are beyond the scope of this article, but can be found elsewhere

(*e.g.*, Hahn *et al.* 2003; Doniger and Fay 2007; Babbitt 2010; He *et al.* 2011). It is sufficient to say that natural selection, which can drive newly arising transcription-factor-binding sites toward some locations of the genome and away from others (*e.g.*, Figure 2), helps coincidentally to drive hotspots preferentially to IGRs and promoter-containing regions. As for hotspot-activating proteins, natural selection upon the DNA sequence motifs might be largely or entirely distinct from natural selection upon recombination hotspot activity itself. Each possible scenario is fully compatible with our model for the stochastic generation of hotspot motifs from cryptic motifs.

Last but not least, population genetics can markedly influence the distribution of hotspots over evolutionary time frames. For example, simulation modeling indicated that allelic drift can affect hotspot positioning in humans due to small effective population sizes and bottlenecks (Coop and Myers 2007). Similarly, population genetics likely had a key role in the positioning of hotspots in closely related species of the genus *Saccharomyces* (Tsai *et al.* 2010). The population-genetic influences are not restricted to *cis*-acting determinants because allelic variation of *trans*-acting factors (*e.g.*, Prdm9) also has a role in specifying the positions of hotspots (Berg *et al.* 2010).

For context, two molecular mechanisms are known to ablate and create hotspots. These are gene conversion (Gutz 1971) and base-pair substitutions (Szankasi *et al.* 1988; Schuchert *et al.* 1991). These primary determinants of change likely operate together in a dynamic equilibrium to help maintain and reposition DNA sequence-dependent hotspots (Figures 1 and 3). Superimposed are other forces—selective and nonselective—that help to shape the recombination landscape over time (*e.g.*, Figure 2). Summarized metaphorically, base-pair substitutions can seed the field, and additional forces can subsequently do the weeding.

As Easy as A-G-C-T?

DNA sequence motifs that activate meiotic recombination hotspots have been exceptionally difficult to identify, due mainly to their short lengths, their context-variable penetrance,[2] and their functional redundancy (Wahls and Davidson 2010). However, the absence of evidence is not evidence for absence. Paradigms established long ago in fission yeast and budding yeast (Schuchert *et al.* 1991; White *et al.* 1991, 1993) have recently been confirmed or implicated in metazoans (Myers *et al.* 2008; Baudat *et al.* 2010) and protozoa (Jiang *et al.* 2011). Furthermore, individual species demonstrably have (Steiner *et al.* 2009, 2011) or likely have (Myers *et al.* 2005, 2008; Berg *et al.* 2010; Jiang *et al.* 2011) multiple, different hotspot-activating motifs. And to the extent tested, each motif apparently helps to regulate as much as 20–41% of recombination in the genome on the basis of frequency distributions of double-stranded DNA breaks and crossovers, respectively (Myers *et al.* 2008; Wahls and Davidson 2010). Such findings render into theory the hypothesis that "a significant fraction of recombination may be regulated by a finite number of discrete [DNA] sites such as *M26*" (Wahls and Smith 1994, p. 1699).

We now suggest, on the basis of the experimental evidence discussed in preceding sections, that most *de novo* hotspots arise from mutations that create recombination-promoting DNA sequence motifs (Figures 1 and 3). Mutations can also alter the DNA-binding-site specificity of hotspot-activating proteins, such as Prdm9, and hence relocate a subset of hotspots (Baudat *et al.* 2010). However, the frequency of such "shifts" is likely many orders of magnitude lower than the frequency with which mutations change cryptic motifs into hotspot motifs. [Consider a "mutational target" density of a few per genome *vs.* $>6 \times 10^4$ per genome (Table S1)] And once a shift has occurred, the newly chosen hotspot motif would become subject to the concentration-dependent rate kinetics of conversion and mutation that drive the motif to dynamic equilibrium in the genome (Figure 3). Indeed, both the shift and subsequent motif-specific drive can be inferred from comparing Prdm9-associated motifs of humans to those of chimpanzees (Myers *et al.* 2010).

The "equilibrium dynamics" model for the overall maintenance of hotspot numbers and the Prdm9 "shift" model are mechanistically distinct and mutually complementary. The shift model provides a way to relocate, in one fell swoop, a subset of hotspots. Such punctuated changes, which for rapidly evolving Prdm9 have probably occurred several times during human evolution (Oliver *et al.* 2009; Berg *et al.* 2010), substitute one set of motifs with another. The occasional shifts *a priori* would not substantially change the total number of hotspot motifs in the genome, so it is difficult to envision how the process could successfully counteract the relentless loss of motifs by gene conversion. The equilibrium dynamics model, on the other hand, does not explain punctuated shifts, but it does provide a way to replace continuously those hotspot motifs lost to conversion (Figure 3) and to progressively move hotspots throughout the genome (Figure 1). Notably, this model applies to all DNA sequence-dependent hotspots, not only to those whose binding proteins undergo atypically rapid evolution of DNA-binding-site specificity.[3] Together, the two models, each based on the theory that discrete DNA sequence motifs help

[2]We propose the term "context-variable penetrance" to describe the fact that individual, discrete DNA sequences motifs known to be recombinogenic exhibit variable levels of hotspot activity at different locations in the genome.

[3]Ten different sequence-specific DNA-binding proteins are known or implicated to help activate hotspots. These are Atf1-Pcr1 heterodimer, Php2-Php3-Php5 complex, Bas1, Bas2, Rap1, Rst2, and Prdm9 (White *et al.* 1991, 1993; Wahls and Smith 1994; Kon *et al.* 1997; Steiner *et al.* 2009, 2011; Baudat *et al.* 2010; Myers *et al.* 2010; Parvanov *et al.* 2010). Prdm9 is the only one whose DNA-binding specificity is known to change rapidly over time. Moreover, Prdm9 is absent from many eukaryotes, including some vertebrate taxa, and it has lost some or all of its functions in other taxa due to mutations (Oliver *et al.* 2009).

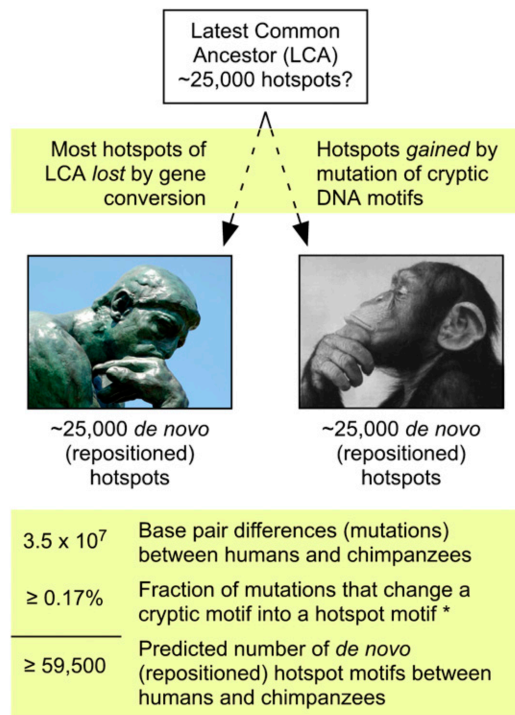

Figure 4 Hotspots created by mutation could account for hotspots repositioned over evolutionary timescales. Gene conversion preferentially removes hotspots from the genome (Gutz 1971; Boulton *et al.* 1997; Jeffreys and Neumann 2009), and this driving force is thought to have ablated most of the hotspots that were present in the latest common ancestor of humans and chimpanzees (Coop and Myers 2007). A second driving force is base-pair mutations, which can generate and remove hotspots (Szankasi *et al.* 1988; Schuchert *et al.* 1991; Fox *et al.* 1997; Steiner *et al.* 2009, 2011). These (and potentially other) opposing forces must operate in a dynamic equilibrium (*e.g.*, Figure 3) for hotspots to be maintained in the genome over evolutionary timescales. We suggest that most hotspots arise from mutations that create hotspot motifs and that this process could account for hotspots repositioned during species divergence. The calculations, which assume a density of regulatory motifs similar to that of fission yeast (asterisk), illustrate this point. Additional mechanisms, such as shifts in the DNA-binding site specificity of hotspot-activating proteins, could also relocate hotspots.

to position meiotic recombination, can explain many features of hotspot biology.

Accounting for Hotspots That Move

Humans and chimpanzees share almost 99% DNA sequence identity (Chimpanzee Sequencing and Analysis Consortium 2005) but few hotspot positions (Ptak *et al.* 2005; Winckler *et al.* 2005). By inference, most hotspot positions of the latest common ancestor have been ablated (converted away) during species divergence (Coop and Myers 2007). Could the hotspot motif model (Figure 1) explain all of the newly positioned hotspots? Yes, if one assumes that the density of cryptic motifs in primate genomes is similar to that documented in fission yeast (Figure 4, Table S1). There are $\sim 3.5 \times 10^7$ single-base-pair differences (mutations) between humans and chimpanzees (Chimpanzee Sequencing and Analysis Consortium 2005). If at least 0.17% of mutations change a cryptic motif into a hotspot motif, then together humans and chimpanzees would have at least 59,500 *de novo* hotspot motifs, relative to the latest common ancestor. At first approximation, these would be sufficient to account for all hotspots in each organism [$\sim$25,000 (Myers *et al.* 2005)]. Thus our hotspot motif model (Figure 1) is robust enough to account for the repositioning of hotspots and the maintenance of hotspot numbers over evolutionary time scales, with or without additional factors such as shifts. We view this as a provisional conclusion, pending a more systematic and comprehensive identification of the nature and density of hotspot motifs within primate (and other) genomes.

Conclusions

Occasionally, a change in one's perspective yields a clear solution to a seemingly intractable problem. One such problem is explaining, mechanistically, the distribution and dynamics of meiotic recombination hotspots. The GENETICS articles by Walter Steiner *et al.* (2009, 2011) revealed that many different DNA sequence motifs of the same organism are recombinogenic, providing an important piece for the puzzle. Our realization that cryptic motifs are densely packed in the genome and can be changed easily into hotspot motifs provided a fresh perspective for the interpretation of existing data. The resulting model, which applies to all DNA sequence-dependent hotspots, describes a mechanism for the dynamic, evolutionarily rapid redistribution of hotspots in the genome. The model also shows how hotspot numbers could be maintained over time without strong selective pressures. It explains why hotspots localize preferentially to IGRs and introns. And it can explain existing data on hotspot repositioning during species divergence.

There are many interesting questions. For example, what regulates the context-variable penetrance of hotspot motifs? What mechanisms underlie sex-specific differences in the distribution of recombination? What are the relative contributions of mutation-conversion equilibria and shifts to hotspot dynamics? The answers to such questions probably lie within the constellation of recombinogenic DNA sequence motifs and the regulation of proteins that bind to those DNA sites. We are, as revealed by the insightful recent work of Steiner *et al.* (2009, 2011), looking at the "tip of the iceberg" for such regulation.

Acknowledgments

We thank Reine Protacio for the construction and plating of strains (Figure 2); Giulia Baldini, Jun Gao, Fengling Kan, and Reine Protacio for helpful discussions; Adam Wilkins and anonymous reviewers for constructive suggestions; and the National Institute of General Medical Sciences at the National Institutes of Health for research support (grant GM81766).

Literature Cited

Babbitt, G. A., 2010 Relaxed selection against accidental binding of transcription factors with conserved chromatin contexts. Gene 466: 43–48.

Baudat, F., and A. Nicolas, 1997 Clustering of meiotic double-strand breaks on yeast chromosome III. Proc. Natl. Acad. Sci. USA 94: 5213–5218.

Baudat, F., J. Buard, C. Grey, A. Fledel-Alon, C. Ober *et al.*, 2010 PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice. Science 327: 836–840.

Buard, J., P. Barthes, C. Grey, and B. de Massy, 2009 Distinct histone modifications define initiation and repair of meiotic recombination in the mouse. EMBO J. 28: 2616–2624.

Berg, I. L., R. Neumann, K. W. Lam, S. Sarbajna, L. Odenthal-Hesse *et al.*, 2010 PRDM9 variation strongly influences recombination hot-spot activity and meiotic instability in humans. Nat. Genet. 42: 859–863.

Blumental-Perry, A., D. Zenvirth, S. Klein, I. Onn, and G. Simchen, 2000 DNA motif associated with meiotic double-strand break regions in *Saccharomyces cerevisiae*. EMBO Rep. 1: 232–238.

Boulton, A., R. S. Myers, and R. J. Redfield, 1997 The hotspot conversion paradox and the evolution of meiotic recombination. Proc. Natl. Acad. Sci. USA 94: 8058–8063.

Buhler, C., V. Borde, and M. Lichten, 2007 Mapping meiotic single-strand DNA reveals a new landscape of DNA double-strand breaks in *Saccharomyces cerevisiae*. PLoS Biol. 5: e324.

Chen, D., W. M. Toone, J. Mata, R. Lyne, G. Burns *et al.*, 2003 Global transcriptional responses of fission yeast to environmental stress. Mol. Biol. Cell 14: 214–229.

Chimpanzee Sequencing and Analysis Consortium, 2005 Initial sequence of the chimpanzee genome and comparison with the human genome. Nature 437: 69–87.

Choo, K. H., 1998 Why is the centromere so cold? Genome Res. 8: 81–82.

Coop, G., and S. R. Myers, 2007 Live hot, die young: transmission distortion in recombination hotspots. PLoS Genet. 3: e35.

Cromie, G. A., C. A. Rubio, R. W. Hyppa, and G. R. Smith, 2005 A natural meiotic DNA break site in *Schizosaccharomyces pombe* is a hotspot of gene conversion, highly associated with crossing over. Genetics 169: 595–605.

Cromie, G. A., R. W. Hyppa, H. P. Cam, J. A. Farah, S. I. Grewal *et al.*, 2007 A discrete class of intergenic DNA dictates meiotic DNA break hotspots in fission yeast. PLoS Genet. 3: e141.

Davidson, M. K., H. K. Shandilya, K. Hirota, K. Ohta, and W. P. Wahls, 2004 Atf1-Pcr1-*M26* complex links stress-activated MAPK and cAMP-dependent protein kinase pathways via chromatin remodeling of *cgs2+*. J. Biol. Chem. 279: 50857–50863.

Doniger, S. W., and J. C. Fay, 2007 Frequent gain and loss of functional transcription factor binding sites. PLoS Comput. Biol. 3: e99.

Ellermeier, C., E. C. Higuchi, N. Phadnis, L. Holm, J. L. Geelhood *et al.*, 2010 RNAi and heterochromatin repress centromeric meiotic recombination. Proc. Natl. Acad. Sci. USA 107: 8701–8705.

Fan, Q., F. Xu, and T. D. Petes, 1995 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.

Fox, M. E., J. B. Virgin, J. Metzger, and G. R. Smith, 1997 Position- and orientation-independent activity of the *Schizosaccharomyces pombe* meiotic recombination hot spot *M26*. Proc. Natl. Acad. Sci. USA 94: 7446–7451.

Frazer, K. A., D. G. Ballinger, D. R. Cox, D. A. Hinds, L. L. Stuve *et al.*, 2007 A second generation human haplotype map of over 3.1 million SNPs. Nature 449: 851–861.

Gao, J., M. K. Davidson, and W. P. Wahls, 2008 Distinct regions of ATF/CREB proteins Atf1 and Pcr1 control recombination hotspot *ade6–M26* and the osmotic stress response. Nucleic Acids Res. 36: 2838–2851.

Gerton, J. L., and R. S. Hawley, 2005 Homologous chromosome interactions in meiosis: diversity amidst conservation. Nat. Rev. Genet. 6: 477–487.

Gerton, J. L., J. DeRisi, R. Shroff, M. Lichten, P. O. Brown *et al.*, 2000 Inaugural article: global mapping of meiotic recombination hotspots and coldspots in the yeast *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA 97: 11383–11390.

Goldman, S. L., and S. Smallets, 1979 Site specific induction of gene conversion: the effects of homozygosity of the *ade6* mutant *M26* of *Schizosaccharomyces pombe* on meiotic gene conversion. Mol. Gen. Genet. 173: 221–225.

Grimm, C., J. Bahler, and J. Kohli, 1994 *M26* recombinational hotspot and physical conversion tract analysis in the *ade6* gene of *Schizosaccharomyces pombe*. Genetics 136: 41–51.

Guillon, H., F. Baudat, C. Grey, R. M. Liskay, and B. de Massy, 2005 Crossover and noncrossover pathways in mouse meiosis. Mol. Cell 20: 563–573.

Gutz, H., 1971 Site specific induction of gene conversion in *Schizosaccharomyces pombe*. Genetics 69: 331–337.

Hahn, M. W., J. E. Stajich, and G. A. Wray, 2003 The effects of selection against spurious transcription factor binding sites. Mol. Biol. Evol. 20: 901–906.

Haring, S. J., L. J. Lautner, J. M. Comeron, and R. E. Malone, 2004 A test of the CoHR motif associated with meiotic double-strand breaks in *Saccharomyces cerevisiae*. EMBO Rep. 5: 41–46.

Hayashi, K., K. Yoshida, and Y. Matsui, 2005 A histone H3 methyltransferase controls epigenetic events required for meiotic prophase. Nature 438: 374–378.

He, B. Z., A. K. Holloway, S. J. Maerkl, and M. Kreitman, 2011 Does positive selection drive transcription factor binding site turnover? A test with *Drosophila cis*-regulatory modules. PLoS Genet. 7: e1002053.

Hirota, K., W. W. Steiner, T. Shibata, and K. Ohta, 2007 Multiple modes of chromatin configuration at natural meiotic recombination hot spots in fission yeast. Eukaryot. Cell 6: 2072–2080.

Jeffreys, A. J., and R. Neumann, 2009 The rise and fall of a human recombination hot spot. Nat. Genet. 41: 625–629.

Jiang, H., N. Li, V. Gopalan, M. M. Zilversmit, S. Varma *et al.*, 2011 High recombination rates and hotspots in a *Plasmodium falciparum* genetic cross. Genome Biol. 12: R33.

Kan, F., M. K. Davidson, and W. P. Wahls, 2011 Meiotic recombination protein Rec12: functional conservation, crossover homeostasis and early crossover/non-crossover decision. Nucleic Acids Res. 39: 1460–1472.

Kon, N., M. D. Krawchuk, B. G. Warren, G. R. Smith, and W. P. Wahls, 1997 Transcription factor Mts1/Mts2 (Atf1/Pcr1, Gad7/Pcr1) activates the *M26* meiotic recombination hotspot in *Schizosaccharomyces pombe*. Proc. Natl. Acad. Sci. USA 94: 13765–13770.

Kong, A., G. Thorleifsson, D. F. Gudbjartsson, G. Masson, A. Sigurdsson *et al.*, 2010 Fine-scale recombination rate differences between sexes, populations and individuals. Nature 467: 1099–1103.

Mihola, O., Z. Trachtulec, C. Vlcek, J. C. Schimenti, and J. Forejt, 2009 A mouse speciation gene encodes a meiotic histone H3 methyltransferase. Science 323: 373–375.

Myers, S., L. Bottolo, C. Freeman, G. McVean, and P. Donnelly, 2005 A fine-scale map of recombination rates and hotspots across the human genome. Science 310: 321–324.

Myers, S., C. Freeman, A. Auton, P. Donnelly, and G. McVean, 2008 A common sequence motif associated with recombina-

tion hot spots and genome instability in humans. Nat. Genet. 40: 1124–1129.
Myers, S., R. Bowden, A. Tumian, R. E. Bontrop, C. Freeman *et al.*, 2010 Drive against hotspot motifs in primates implicates the *PRDM9* gene in meiotic recombination. Science 327: 876–879.
Nicolas, A., D. Treco, N. P. Schultes, and J. W. Szostak, 1989 An initiation site for meiotic gene conversion in the yeast *Saccharomyces cerevisiae*. Nature 338: 35–39.
Oliver, P. L., L. Goodstadt, J. J. Bayes, Z. Birtle, K. C. Roach *et al.*, 2009 Accelerated evolution of the *Prdm9* speciation gene across diverse metazoan taxa. PLoS Genet. 5: e1000753.
Parvanov, E. D., P. M. Petkov, and K. Paigen, 2010 Prdm9 controls activation of mammalian recombination hotspots. Science 327: 835.
Peters, A. D., 2008 A combination of *cis* and *trans* control can solve the hotspot conversion paradox. Genetics 178: 1579–1593.
Petes, T. D., 2001 Meiotic recombination hot spots and cold spots. Nat. Rev. Genet. 2: 360–369.
Pineda-Krch, M., and R. J. Redfield, 2005 Persistence and loss of meiotic recombination hotspots. Genetics 169: 2319–2333.
Ponting, C. P., 2011 What are the genomic drivers of the rapid evolution of PRDM9? Trends Genet. 27: 165–171.
Pryce, D. W., A. Lorenz, J. B. Smirnova, J. Loidl, and R. J. McFarlane, 2005 Differential activation of *M26*-containing meiotic recombination hot spots in *Schizosaccharomyces pombe*. Genetics 170: 95–106.
Ptak, S. E., D. A. Hinds, K. Koehler, B. Nickel, N. Patil *et al.*, 2005 Fine-scale recombination patterns differ between chimpanzees and humans. Nat. Genet. 37: 429–434.
Robine, N., N. Uematsu, F. Amiot, X. Gidrol, E. Barillot *et al.*, 2007 Genome-wide redistribution of meiotic double-strand breaks in *Saccharomyces cerevisiae*. Mol. Cell. Biol. 27: 1868–1880.
Schuchert, P., and J. Kohli, 1988 The *ade6–M26* mutation of *Schizosaccharomyces pombe* increases the frequency of crossing over. Genetics 119: 507–515.
Schuchert, P., M. Langsford, E. Kaslin, and J. Kohli, 1991 A specific DNA sequence is required for high frequency of recombination in the *ade6* gene of fission yeast. EMBO J. 10: 2157–2163.
Shiozaki, K., and P. Russell, 1996 Conjugation, meiosis, and the osmotic stress response are regulated by Spc1 kinase through Atf1 transcription factor in fission yeast. Genes Dev. 10: 2276–2288.
Steiner, W. W., and G. R. Smith, 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.
Steiner, W. W., E. M. Steiner, A. R. Girvin, and L. E. Plewik, 2009 Novel nucleotide sequence motifs that produce hotspots of meiotic recombination in *Schizosaccharomyces pombe*. Genetics 182: 459–469.
Steiner, W. W., P. A. Davidow, and A. T. Bagshaw, 2011 Important characteristics of sequence-specific recombination hotspots in *Schizosaccharomyces pombe*. Genetics 187: 385–396.
Szankasi, P., W. D. Heyer, P. Schuchert, and J. Kohli, 1988 DNA sequence analysis of the *ade6* gene of *Schizosaccharomyces pombe*: wild-type and mutant alleles including the recombination hot spot allele *ade6–M26*. J. Mol. Biol. 204: 917–925.
Thomas, J. H., R. O. Emerson, and J. Shendure, 2009 Extraordinary molecular evolution in the *PRDM9* fertility gene. PLoS ONE 4: e8505.
Tsai, I. J., A. Burt, and V. Koufopanou, 2010 Conservation of recombination hotspots in yeast. Proc. Natl. Acad. Sci. USA 107: 7847–7852.
Ubeda, F., and J. F. Wilkins, 2011 The Red Queen theory of recombination hotspots. J. Evol. Biol. 24: 541–553.
Virgin, J. B., J. Metzger, and G. R. Smith, 1995 Active and inactive transplacement of the *M26* recombination hotspot of *Schizosaccharomyces pombe*. Genetics 141: 33–48.
Wahls, W. P., and M. K. Davidson, 2010 Discrete DNA sites regulate global distribution of meiotic recombination. Trends Genet. 26: 202–208.
Wahls, W. P., and P. D. Moore, 1998 Recombination hotspot activity of hypervariable minisatellite DNA requires minisatellite DNA binding proteins. Somat. Cell Mol. Genet. 24: 41–51.
Wahls, W. P., and G. R. Smith, 1994 A heteromeric protein that binds to a meiotic homologous recombination hot spot: correlation of binding and hot spot activity. Genes Dev. 8: 1693–1702.
Wahls, W. P., L. J. Wallace, and P. D. Moore, 1990 Hypervariable minisatellite DNA is a hotspot for homologous recombination in human cells. Cell 60: 95–103.
Wahls, W. P., G. Swenson, and P. D. Moore, 1991 Two hypervariable minisatellite DNA binding proteins. Nucleic Acids Res. 19: 3269–3274.
Wahls, W. P., E. R. Siegel, and M. K. Davidson, 2008 Meiotic recombination hotspots of fission yeast are directed to loci that express non-coding RNA. PLoS ONE 3: e2887.
White, M. A., M. Wierdl, P. Detloff, and T. D. Petes, 1991 DNA-binding protein RAP1 stimulates meiotic recombination at the *HIS4* locus in yeast. Proc. Natl. Acad. Sci. USA 88: 9755–9759.
White, M. A., M. Dominska, and T. D. Petes, 1993 Transcription factors are required for the meiotic recombination hotspot at the *HIS4* locus in *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA 90: 6621–6625.
Wilkinson, M. G., M. Samuels, T. Takeda, W. M. Toone, J. C. Shieh *et al.*, 1996 The Atf1 transcription factor is a target for the Sty1 stress-activated MAP kinase pathway in fission yeast. Genes Dev. 10: 2289–2301.
Winckler, W., S. R. Myers, D. J. Richter, R. C. Onofrio, G. J. McDonald *et al.*, 2005 Comparison of fine-scale recombination rates in humans and chimpanzees. Science 308: 107–111.
Yamada, T., K. Mizuno, K. Hirota, N. Kon, W. P. Wahls *et al.*, 2004 Roles of histone acetylation and chromatin remodeling factor in a meiotic recombination hotspot. EMBO J. 23: 1792–1803.

GENETICS

Supporting Information
http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.111.134130/-/DC1

DNA Sequence-Mediated, Evolutionarily Rapid Redistribution of Meiotic Recombination Hotspots

Commentary on Genetics 182: 459–469 and Genetics 187: 385–396

Wayne P. Wahls and Mari K. Davidson

Copyright © 2011 by the Genetics Society of America
DOI: 10.1534/genetics.111.134130

TABLE S1 Frequencies of hotspot and cryptic DNA sequence motifs in the *S. pombe* genome

Motif	Sequence	Frequency
M26	**ATGACGT**	**285**
M26 cryptic 1	gTGACGT	267
M26 cryptic 2	tTGACGT	1162
M26 cryptic 3	cTGACGT	426
M26 cryptic 4	AgGACGT	649
M26 cryptic 5	AaGACGT	955
M26 cryptic 6	AcGACGT	711
M26 cryptic 7	ATaACGT	1559
M26 cryptic 8	ATtACGT	1077
M26 cryptic 9	ATcACGT	850
M26 cryptic 10	ATGgCGT	795
M26 cryptic 11	ATGtCGT	1180
M26 cryptic 12	ATGcCGT	726
M26 cryptic 13	ATGAgGT	1086
M26 cryptic 14	ATGAaGT	2905
M26 cryptic 15	ATGAtGT	2192
M26 cryptic 16	ATGACaT	1923
M26 cryptic 17	ATGACtT	2752
M26 cryptic 18	ATGACcT	966
M26 cryptic 19	ATGACGg	649
M26 cryptic 20	ATGACGa	1335
M26 cryptic 21	ATGACGc	425
CCAAT	**CCAATCA**	**1408**
CCAAT cryptic 1	gCAATCA	1885
CCAAT cryptic 2	aCAATCA	2619
CCAAT cryptic 3	tCAATCA	2966

CCAAT cryptic 4	CgAATCA	1202
CCAAT cryptic 5	CaAATCA	3681
CCAAT cryptic 6	CtAATCA	1536
CCAAT cryptic 7	CCgATCA	689
CCAAT cryptic 8	CCtATCA	999
CCAAT cryptic 9	CCcATCA	1025
CCAAT cryptic 10	CCAgTCA	949
CCAAT cryptic 11	CCAtTCA	2340
CCAAT cryptic 12	CCAcTCA	811
CCAAT cryptic 13	CCAAgCA	1731
CCAAT cryptic 14	CCAAaCA	2589
CCAAT cryptic 15	CCAAcCA	1454
CCAAT cryptic 16	CCAATgA	1747
CCAAT cryptic 17	CCAATaA	2475
CCAAT cryptic 18	CCAATtA	2223
CCAAT cryptic 19	CCAATCg	869
CCAAT cryptic 20	CCAATCt	1616
CCAAT cryptic 21	CCAATCc	1363
Oligo-C	**CCCCGCAC**	**10**
Oligo-C cryptic 1	gCCCGCAC	35
Oligo-C cryptic 2	aCCCGCAC	54
Oligo-C cryptic 3	tCCCGCAC	54
Oligo-C cryptic 4	CgCCGCAC	27
Oligo-C cryptic 5	CaCCGCAC	66
Oligo-C cryptic 6	CtCCGCAC	34
Oligo-C cryptic 7	CCgCGCAC	31
Oligo-C cryptic 8	CCaCGCAC	57
Oligo-C cryptic 9	CCtCGCAC	66
Oligo-C cryptic 10	CCCgGCAC	34

Oligo-C cryptic 11	CCCaGCAC	99
Oligo-C cryptic 12	CCCtGCAC	68
Oligo-C cryptic 13	CCCCaCAC	52
Oligo-C cryptic 14	CCCCtCAC	45
Oligo-C cryptic 15	CCCCcCAC	55
Oligo-C cryptic 16	CCCCGgAC	18
Oligo-C cryptic 17	CCCCGaAC	57
Oligo-C cryptic 18	CCCCGtAC	49
Oligo-C cryptic 19	CCCCGCgC	7
Oligo-C cryptic 20	CCCCGCtC	36
Oligo-C cryptic 21	CCCCGCcC	14
Oligo-C cryptic 22	CCCCGCAg	17
Oligo-C cryptic 23	CCCCGCAa	40
Oligo-C cryptic 24	CCCCGCAt	25
4095 (a)	**GGTCTAGAC**	**4**
4095 (a) cryptic 1	aGTCTAGAC	13
4095 (a) cryptic 2	tGTCTAGAC	11
4095 (a) cryptic 3	cGTCTAGAC	11
4095 (a) cryptic 4	GaTCTAGAC	20
4095 (a) cryptic 5	GtTCTAGAC	11
4095 (a) cryptic 6	GcTCTAGAC	32
4095 (a) cryptic 7	GGgCTAGAC	6
4095 (a) cryptic 8	GGaCTAGAC	21
4095 (a) cryptic 9	GGcCTAGAC	5
4095 (a) cryptic 10	GGTgTAGAC	28
4095 (a) cryptic 11	GGTaTAGAC	43
4095 (a) cryptic 12	GGTtTAGAC	59
4095 (a) cryptic 13	GGTCgAGAC	10
4095 (a) cryptic 14	GGTCaAGAC	47

4095 (a) cryptic 15	GGTCcAGAC	11
4095 (a) cryptic 16	GGTCTtGAC	53
4095 (a) cryptic 17	GGTCTcGAC	16
4095 (a) cryptic 18	GGTCTAaAC	53
4095 (a) cryptic 19	GGTCTAtAC	32
4095 (a) cryptic 20	GGTCTAcAC	12
4095 (a) cryptic 21	GGTCTAGgC	8
4095 (a) cryptic 22	GGTCTAGtC	21
4095 (a) cryptic 23	GGTCTAGcC	7
4095 (a) cryptic 24	GGTCTAGAg	11
4095 (a) cryptic 25	GGTCTAGAa	35
4095 (a) cryptic 26	GGTCTAGAt	22
4095 (b)	**GGTCTGGAC**	**5**
4095 (b) cryptic 1	aGTCTGGAC	12
4095 (b) cryptic 2	tGTCTGGAC	30
4095 (b) cryptic 3	cGTCTGGAC	17
4095 (b) cryptic 4	GaTCTGGAC	30
4095 (b) cryptic 5	GtTCTGGAC	31
4095 (b) cryptic 6	GcTCTGGAC	31
4095 (b) cryptic 7	GGgCTGGAC	11
4095 (b) cryptic 8	GGaCTGGAC	25
4095 (b) cryptic 9	GGcCTGGAC	8
4095 (b) cryptic 10	GGTgTGGAC	20
4095 (b) cryptic 11	GGTaTGGAC	53
4095 (b) cryptic 12	GGTtTGGAC	77
4095 (b) cryptic 13	GGTCgGGAC	7
4095 (b) cryptic 14	GGTCaGGAC	15
4095 (b) cryptic 15	GGTCcGGAC	6
4095 (b) cryptic 16	GGTCTtGAC	53

4095 (b) cryptic 17	GGTCTcGAC	16
4095 (b) cryptic 18	GGTCTGaAC	27
4095 (b) cryptic 19	GGTCTGtAC	19
4095 (b) cryptic 20	GGTCTGcAC	14
4095 (b) cryptic 21	GGTCTGGgC	6
4095 (b) cryptic 22	GGTCTGGtC	22
4095 (b) cryptic 23	GGTCTGGcC	9
4095 (b) cryptic 24	GGTCTGGAg	11
4095 (b) cryptic 25	GGTCTGGAa	36
4095 (b) cryptic 26	GGTCTGGAt	21
8-6	**TCGGCCGA**	**8**
8-6 cryptic 1	gCGGCCGA	38
8-6 cryptic 2	aCGGCCGA	41
8-6 cryptic 3	cCGGCCGA	16
8-6 cryptic 4	TgGGCCGA	72
8-6 cryptic 5	TaGGCCGA	50
8-6 cryptic 6	TtGGCCGA	172
8-6 cryptic 7	TCaGCCGA	159
8-6 cryptic 8	TCtGCCGA	180
8-6 cryptic 9	TCcGCCGA	79
8-6 cryptic 10	TCGaCCGA	115
8-6 cryptic 11	TCGtCCGA	112
8-6 cryptic 12	TCGcCCGA	54

The under-representation of recombinogenic (hotspot) DNA sequence motifs in the genome has been reported (WAHLS AND SMITH, 1994; STEINER *et al.* 2011). To adjust for compositional bias and to shed light upon hotspot dynamics, we compared the frequencies of known hotspot motifs to those of closely related, known/presumptively inactive motifs. Data are the frequencies of hotspot DNA sequence motifs (bold) and single base pair variant, "cryptic" motifs (substitution in

lower case) in the assembled sequence of the fission yeast genome (12,571,820 base pairs). There are 64,622 cryptic motifs, giving an average distribution of one cryptic motif every 194 base pairs along the genome. It is worth emphasizing that there are additional recombination-promoting DNA sequences whose hotspot motifs remain to be identified and even more are predicted statistically (STEINER *et al.* 2009), so the actual density of cryptic motifs is greater than that calculated here.