

# Title: Non-paradoxical evolutionary stability of the recombination initiation landscape in *Saccharomycetes*

Authors: Isabel Lam<sup>1,2</sup> and Scott Keeney<sup>1,2,3,\*</sup>

## Affiliations:

<sup>1</sup> Louis V. Gerstner Jr. Graduate School of Biomedical Sciences, Memorial Sloan Kettering Cancer Center, New York, NY, 10065, USA.

<sup>2</sup> Molecular Biology Program, Memorial Sloan Kettering Cancer Center, New York, NY, 10065, USA.

<sup>3</sup> Howard Hughes Medical Institute, Memorial Sloan Kettering Cancer Center, New York, NY, 10065, USA.

\*Correspondence to: s-keeney@ski.mskcc.org.

**Abstract:** The nonrandom distribution of meiotic recombination shapes heredity and genetic diversification. A widely held view is that individual hotspots — favored sites of recombination initiation — are always ephemeral because they evolve rapidly toward extinction. An alternative view, often ignored or dismissed as implausible, predicts conservation of the positions of hotspots if they are chromosomal features under selective constraint, such as gene promoters. Here we empirically test opposite predictions of these theories by comparing genome-wide maps of meiotic recombination initiation from widely divergent species in the *Saccharomyces* clade. We find that the frequent overlap of hotspots with promoters is true of the species tested and, consequently, hotspot positions are well conserved. Remarkably, however, the relative strength of individual hotspots is also highly conserved, as are larger-scale features of the distribution of recombination initiation. This stability, not predicted by prior models, suggests that the particular shape of the yeast recombination landscape is adaptive, and helps in understanding evolutionary dynamics of recombination in other species.

**One Sentence Summary:** Hotspots for meiotic double-strand breaks decay rapidly in some species but are maintained over long evolutionary times in yeast, where they are targeted to functional genomic features.

**Main Text:** DNA double-strand breaks (DSBs) generated by Spo11 protein initiate meiotic recombination, which in turn alters genetic linkage and promotes pairing and accurate segregation of homologous chromosomes (*1*). DSBs are not distributed randomly, occurring often within narrow (typically < 200 bp wide in budding yeast) regions called hotspots (*2*). Theoretical work exploring evolutionary dynamics has led to a prevailing hypothesis, the “hotspot paradox”, that predicts rapid hotspot extinction (*3-8*). This view rests on biased gene conversion: the broken chromosome copies genetic information from its uncut homolog (**Fig. 1A**). As a result, hotspot alleles with different DSB activity deviate from a Mendelian segregation ratio, with colder (less active) alleles overrepresented in offspring. Such meiotic drive is observed in yeast (*9*) and humans (*10*) and predicts that mutations that reduce or

eliminate hotspot activity will be rapidly fixed in populations, while hotspot-activating mutations should be rapidly extinguished (3, 5, 11). The paradox is that hotspots for DSBs and recombination exist at all.

An alternative view predicts that hotspot positions can be preserved if Spo11 targets genomic features that are under selective constraint to maintain functions unrelated to their roles as hotspots (9, 12, 13). This hypothesis derives from the fact that most DSB hotspots in the budding yeast *S. cerevisiae* correspond to promoter-containing intergenic regions (13). However, theoretical studies have considered this implausible as a mechanism to preserve hotspots (3, 11). Instead, many studies take as their starting point the assumption that hotspot lifespan must always be short and that the fine-scale recombination initiation landscape will always be highly dynamic over long evolutionary times (4, 6-8, 14). This assumption is appropriate for humans and most other mammals (discussed further below), but has not been evaluated for other taxa.

To distinguish between these models, we asked whether the DSB landscape is conserved in yeast. A prior attempt to address this issue used population genetic data to deduce a recombination map in *S. paradoxus* and compared it to *S. cerevisiae* (15). Partial conservation was inferred, but the data had insufficient resolution and sensitivity to detect individual hotspots and were confined to one small chromosome that may not be representative of most of the yeast genome. We overcame these limitations by comparing high-resolution, whole-genome DSB maps between more widely diverged *Saccharomyces* species, as well as between *S. cerevisiae* strains from different lineages (the laboratory strain SK1 and wild-derived strains YPS128 and UWOPS03-461.4) (16) (**Fig. 1B, Table S1**). DSB maps were generated by deep-sequencing of DNA oligonucleotides (oligos) that remain covalently bound to Spo11 as a byproduct of DSB formation and processing (13, 17) (**Fig. 1A, Table S2**).

The *Saccharomyces sensu stricto* clade shared a last common ancestor ~20 million years ago (18). We chose species at varying evolutionary distances from *S. cerevisiae* (**Fig. 1B**) ranging from *S. paradoxus*, with protein coding sequence divergence from *S. cerevisiae* comparable to that between humans and mouse, to *S. kudriavzevii*, roughly as distant as mammals from birds (19). The *S. cerevisiae* strains chosen display 0.5–0.7% sequence divergence, comparable to the polymorphism density between humans and chimpanzees. Most differences are simple sequence polymorphisms, with few large-scale structural differences aside from one discussed below (20). The *Saccharomyces* karyotype is thus an evolutionarily stable, successful one.

These yeasts all underwent synchronous and efficient meiosis (**fig. S1A**), hence SK1 is not anomalous in this regard. As in *S. cerevisiae*, two major size classes of Spo11-oligo complexes were observed (**Fig. 1C, fig. S1B**), reflecting oligos of similar length distributions (**fig. S1C**). Each oligo is a tag recording precisely where Spo11 generated a DSB, and maps based on deep sequencing Spo11 oligos agree spatially and quantitatively with direct detection of DSBs by Southern blot (13, 21, 22). Biological replicate maps were highly reproducible (**Fig. 1D, fig. S2**) and most sequenced reads (>98%) mapped uniquely to their respective genomes (**Table S2**).

We first asked whether targeting of promoters is conserved. To aid analysis, we generated nucleosome occupancy maps by sequencing micrococcal nuclease-resistant DNA (MNase-seq) from meiotic cultures. In *S. cerevisiae*, DSBs form preferentially in promoter-associated nucleosome-depleted regions (NDRs) (13, 23), and promoter chromatin structure is

strongly conserved in other *Saccharomyces* species (24-26). Spo11 oligos were highly enriched in promoter NDRs in all species tested, whether examined at individual locations (**Fig. 1E, fig. S3A**), or averaged across annotated genes (**Fig. 1F, fig. S3B**). Many Spo11 oligos mapped to promoter-containing intergenic regions (IGRs) (i.e., those flanked by divergent or tandemly oriented transcription units), whereas few mapped to convergent IGRs (i.e., those lacking promoters) or within genes (**fig. S3D**). Clearly, Spo11 preference for promoters is a stable feature of the *Saccharomyces* DSB landscape.

Similar numbers of Spo11-oligo hotspots (~4000) were identified in all species (**Table S3**). When ranked by heat (Spo11-oligo count), hotspots formed a smooth continuum over a wide range in both *S. cerevisiae* and *S. kudriavzevii* (**Fig. 1H**), with nearly superimposable cumulative curves in all species (**fig. S3E**). Hence, the distribution of DSBs among hotspots is the same. Hotspots decisively tended toward low nucleosome occupancy (**Fig. 1G, fig. S3C**) consistent with open chromatin structure providing a window of opportunity for Spo11 to access DNA (27). The distribution of hotspot widths was also nearly identical, with wider hotspots tending to have more Spo11 oligos (**Fig. 1I, fig. S3F**). Conserved hotspot width agrees with previously established conservation of NDR width (24-26). Most important, the majority of hotspots overlapped the same promoter-containing IGRs in all species examined (**Fig. 1J, fig. S3G**). The low frequency of sex and outcrossing in yeasts could slow hotspot extinction compared to obligately outcrossed species (15, 28), but the yeasts examined here have had ample sexual generations to allow biased gene conversion to erode hotspots (estimated >200,000 outcrossed sexual generations since divergence of *S. cerevisiae* from *S. kudriavzevii*, comparable to the number of human sexual generations since divergence from chimpanzees; see Methods). Thus, as predicted (9), DSB hotspot locations can be preserved when the underlying chromosome architecture being targeted is conserved.

We next examined conservation of hotspot heat. The hotspot paradox predicts that hotspot activity will be in a state of flux: even if hotspot locations are conserved, they should vary widely in absolute and relative strength. Furthermore, the rapidity of hotspot extinction should scale with hotspot heat, because alleles that experience high DSB frequency provide more chances for loss for a given number of sexual generations (3-5, 11, 29). The selective constraint model is agnostic in this regard. If cis-acting sequence polymorphisms can quantitatively modulate DSB formation without ablating Spo11 targeting per se (experimentally shown to be a plausible proposition (e.g., 30)), then hotspot heats will be free to change and will do so rapidly (13). On the other hand, if DSB frequency (not just position) is tied to selectively constrained aspects of the targeted feature, or if DSB frequency is itself constrained, then hotspot heats will tend to be conserved.

To address this question, we summed Spo11 oligos mapping within 3426 promoter-containing IGRs that could be stringently and unambiguously matched between species based on conservation of the flanking coding sequences (**Fig. 2A, Table S4**). (This accounts for 81% of divergent and tandem IGRs and 83% of promoter-proximal DSB hotspots in *S. cerevisiae*, so most of the relevant Spo11-targeted genomic space is included. An IGR-centric approach is preferable to relying on more arbitrary hotspot definitions; see Methods.) Intra-IGR Spo11-oligo counts were highly similar between *S. cerevisiae* strains, with correlation coefficients (0.89–0.92) nearly as high as for comparisons between biological replicates (0.97–1.00) (**Fig. 2B,C, fig. S4A**, and data not shown). Thus, yeast intra-species variation in local DSB heat is low despite ~0.7–1% median sequence divergence within these IGRs. Strong positive correlations

were also found between species, with little change in correlation strength over large evolutionary distances (**Fig. 2B,C, fig. S4A**). Moreover, the hottest 1% of promoter IGRs in *S. cerevisiae* SK1 (based on Spo11-oligo count) were highly enriched among the hottest IGRs in other strains and species, with a median percentile ranking within the top 5% even in *S. kudriavzevii* (**Fig. 2D,E**). This was only modestly greater than the extent of conservation of the coldest IGRs (**Fig. 2D**), whereas theoretical modeling of the effects that biased gene conversion alone should have on hotspot longevity predicts that strong hotspots are much less likely to be shared between species than weak ones (5). We did find specific examples where strong hotspots in one species were substantially weaker in other species (**Fig. 2F**), so there is no absolute barrier to evolutionary changes in DSB activity. But the behavior of most IGRs leads us to conclude that the hottest hotspots present in the last *Saccharomyces* common ancestor have tended to retain their places at the top of the list of Spo11 targets, and that it has been rare for ancestrally cold promoter regions to acquire strong hotspot activity.

This high degree of hotspot conservation differs markedly from the situation in most mammals, where PRDM9 provides a different solution to the hotspot paradox (31). PRDM9 is a histone methyltransferase with an array of Zn-finger modules that rapidly evolve new DNA binding specificity. The DNA motifs it recognizes, which have no known intrinsic function, are lost quickly from genomes of humans and mice because of meiotic drive from biased gene conversion (29, 32, 33), but periodic appearance of new *PRDM9* alleles with different sequence specificity creates new hotspots suddenly and thus redraws the recombination initiation landscape (31). Accordingly, DSB hotspot heat between men sharing the same or similar *PRDM9* alleles (34) was less conserved than between *S. cerevisiae* strains despite much greater sequence identity (**Fig. 2C, fig. S4B**). This difference is consistent with PRDM9 motif erosion contributing to variation in hotspot strength between individuals (34).

Hotspots are only one level of non-randomness in the DSB landscape in that they reside within larger domains of intrinsically greater or lesser DSB potential (12, 13). Conservation has been noted for the distribution of crossover recombination over broad regions in several taxa (14), but DSB distributions have not been evaluated. We therefore asked if large-scale features of the DSB landscape are also conserved in yeast. Spo11-oligo maps demonstrated that DSB suppression near telomeres and centromeres (13, 27) is preserved across *Saccharomyces* evolution (**Fig. 3A,B, fig. S5A,B**). This behavior is not surprising, as recombination in these subchromosomal domains can interfere with genome integrity (subtelomeric regions are rife with repetitive DNA elements that can undergo nonallelic homologous recombination (35), and crossing over too close to centromeres can provoke meiotic segregation errors (36)). More remarkably, however, Spo11-oligo counts were highly correlated when compared in ~20-kb segments within larger blocks of synteny across the remaining interstitial portions of chromosomes (**Fig. 3C,D, fig. S5C-F, Table S5**). Thus, the domain structure of the DSB landscape is also evolutionarily stable. As in *S. cerevisiae*, Spo11-oligo counts correlated positively with G+C content of DNA in each species tested, with similar scale-dependent patterns of weaker correlation at short distances (~1 kb) and stronger correlation at large distances (**Fig. 3E**). This pattern is consistent with the hypothesis that large-scale DSB domains, like hotspots, reflect features of the underlying chromosome architecture that are selectively constrained (13). Furthermore, such large-scale domains presumably reflect operation of factors that work in cis but at a distance from DSB hotspots, i.e., too far to be frequently included in gene conversion tracts and thus not subject to loss through biased gene conversion. Such factors are not expected to evolve as rapidly as hotspots (3, 5, 14, 37).

Smaller chromosomes undergo more crossing over per kb than larger chromosomes (38). This behavior can be accounted for by an anticorrelation of DSB density with size (13), which is a patterning effect of a negative feedback circuit that inhibits DSB formation when homologous chromosomes successfully engage one another (39). We find that the negative correlation between Spo11-oligo density and chromosome length is conserved (**Fig. 4A, fig. S6**). Chromosome bisection and fusion experiments demonstrated that variation in crossover density is determined by chromosome length (38), but this has not been formally tested for DSBs. *S. mikatae* provides a natural experiment: reciprocal translocations between ancestral chromosomes VI, VII, and XVI gave rise to a longer chromosome VI and a shorter chromosome VII in *S. mikatae* than in other *Saccharomyces* species (40) (**Fig. 4B**). Spo11-oligo density within segments syntenic with the left and right arms of ancestral chromosome VI adopted the Spo11-oligo density predicted by their chromosome length context: density was higher when the segments resided on the short chromosome VI in *S. cerevisiae* but lower when on longer chromosomes in *S. mikatae* (**Fig. 4B**). Control syntenic segments on similar-length chromosomes exhibited matched Spo11-oligo densities (**Fig. 4C**). We conclude that whole-chromosome variation in DSB density is a direct consequence of chromosome size per se and is thus extrinsic to the DNA sequence.

That yeast hotspot positions are conserved fits the hypothesis that hotspots tend to be stable if Spo11 targets functional genomic elements that are evolutionarily constrained (9). *Drosophila* provides an interesting contrast to reinforce this conclusion: this genus lacks a PRDM9-like system but also does not preferentially target recombination to promoters or known functional elements (in fact, it appears to lack hotspots altogether); accordingly, the fine-scale recombination pattern appears to evolve rapidly (41). A corollary is that evolutionary stability may imply that DSB hotspots have a constrained function(s), even if that function is presently obscure. Interestingly, DSB hotspots are well conserved between the *Schizosaccharomyces* species *S. pombe* and *S. kambucha* (~0.5% sequence divergence) (42) despite mapping to sites without known function (i.e., in large intergenic regions but not gene promoters (43)).

Strong conservation in *Saccharomycetes* of relative DSB frequencies among hotspots, across subchromosomal domains, and even across whole chromosomes supports the hypothesis that this conservation traces back to the DSB landscape being shaped by selectively constrained chromosomal features that work combinatorially, hierarchically, and over multiple size scales (13). For example, transcription, replication, telomere and centromere function, sister chromatid cohesion, and chromosome compaction rely on and shape chromosome structures over scales ranging from tens to millions of base pairs. These structures in turn also mold the DSB landscape, so selective pressure to maintain them for gene expression, cell division, and other processes imposes a tendency to conserve the DSB landscape. However, the remarkable strength of conservation across millions of years of evolution leads us to speculate that the specific shape of the yeast DSB landscape may confer fitness benefits. It has long been understood that the recombination distribution is a heritable trait subject to selection (14, 44). We speculate that selective pressures may operate more directly on the DSB landscape genome-wide, perhaps related to accurate meiotic chromosome segregation and/or beneficial effects of disrupting or maintaining linkage groups at various size scales (44, 45).

Finally, we note that available evidence in plants, birds, and dogs — all of which appear to lack a PRDM9-like hotspot targeting mechanism — point to Spo11 acting preferentially at promoters, CpG islands, and/or other genomic elements that are under selective constraint to



maintain functions separate from being Spo11 targets (46-48). In finches, high-resolution recombination maps inferred from population genetic data reveal evolutionary stability of recombination hotspots, analogous to *Saccharomyces* spp. but wholly unlike PRDM9-reliant apes or mice (48). Thus, not only is it untrue that recombination initiation landscapes inevitably evolve rapidly, but conservation is likely to be a common pattern for many sexual species.

## References and Notes:

1. B. de Massy, Initiation of meiotic recombination: how and where? Conservation and specificities among eukaryotes. *Annu Rev Genet* **47**, 563 (2013).
2. L. Kauppi, A. J. Jeffreys, S. Keeney, Where the crossovers are: recombination distributions in mammals. *Nat Rev Genet* **5**, 413 (2004).
3. A. Boulton, R. S. Myers, R. J. Redfield, The hotspot conversion paradox and the evolution of meiotic recombination. *Proc Natl Acad Sci U S A* **94**, 8058 (1997).
4. P. Calabrese, A population genetics model with recombination hotspots that are heterogeneous across the population. *Proc Natl Acad Sci U S A* **104**, 4748 (2007).
5. G. Coop, S. R. Myers, Live hot, die young: transmission distortion in recombination hotspots. *PLoS Genet* **3**, e35 (2007).
6. U. Friberg, W. R. Rice, Cut thy neighbor: cyclic birth and death of recombination hotspots via genetic conflict. *Genetics* **179**, 2229 (2008).
7. F. Ubeda, J. F. Wilkins, The Red Queen theory of recombination hotspots. *J Evol Biol* **24**, 541 (2011).
8. W. P. Wahls, M. K. Davidson, DNA sequence-mediated, evolutionarily rapid redistribution of meiotic recombination hotspots. *Genetics* **189**, 685 (2011).
9. A. Nicolas, D. Treco, N. P. Schultes, J. W. Szostak, An initiation site for meiotic gene conversion in the yeast *Saccharomyces cerevisiae*. *Nature* **338**, 35 (1989).
10. A. J. Jeffreys, R. Neumann, Reciprocal crossover asymmetry and meiotic drive in a human recombination hot spot. *Nat Genet* **31**, 267 (2002).
11. M. Pineda-Krch, R. J. Redfield, Persistence and loss of meiotic recombination hotspots. *Genetics* **169**, 2319 (2005).
12. M. Lichten, B. de Massy, The impressionistic landscape of meiotic recombination. *Cell* **147**, 267 (2011).
13. J. Pan *et al.*, A hierarchical combination of factors shapes the genome-wide topography of yeast meiotic recombination initiation. *Cell* **144**, 719 (2011).
14. C. S. Smukowski, M. A. Noor, Recombination rate variation in closely related species. *Heredity* **107**, 496 (2011).
15. I. J. Tsai, A. Burt, V. Koufopanou, Conservation of recombination hotspots in yeast. *Proc Natl Acad Sci U S A* **107**, 7847 (2010).
16. G. Liti *et al.*, Population genomics of domestic and wild yeasts. *Nature* **458**, 337 (2009).
17. M. J. Neale, J. Pan, S. Keeney, Endonucleolytic processing of covalent protein-linked DNA double-strand breaks. *Nature* **436**, 1053 (2005).
18. T. Replansky, V. Koufopanou, D. Greig, G. Bell, *Saccharomyces sensu stricto* as a model system for evolution and ecology. *Trends in ecology & evolution* **23**, 494 (2008).
19. B. Dujon, Yeasts illustrate the molecular mechanisms of eukaryotic genome evolution. *Trends Genet* **22**, 375 (2006).

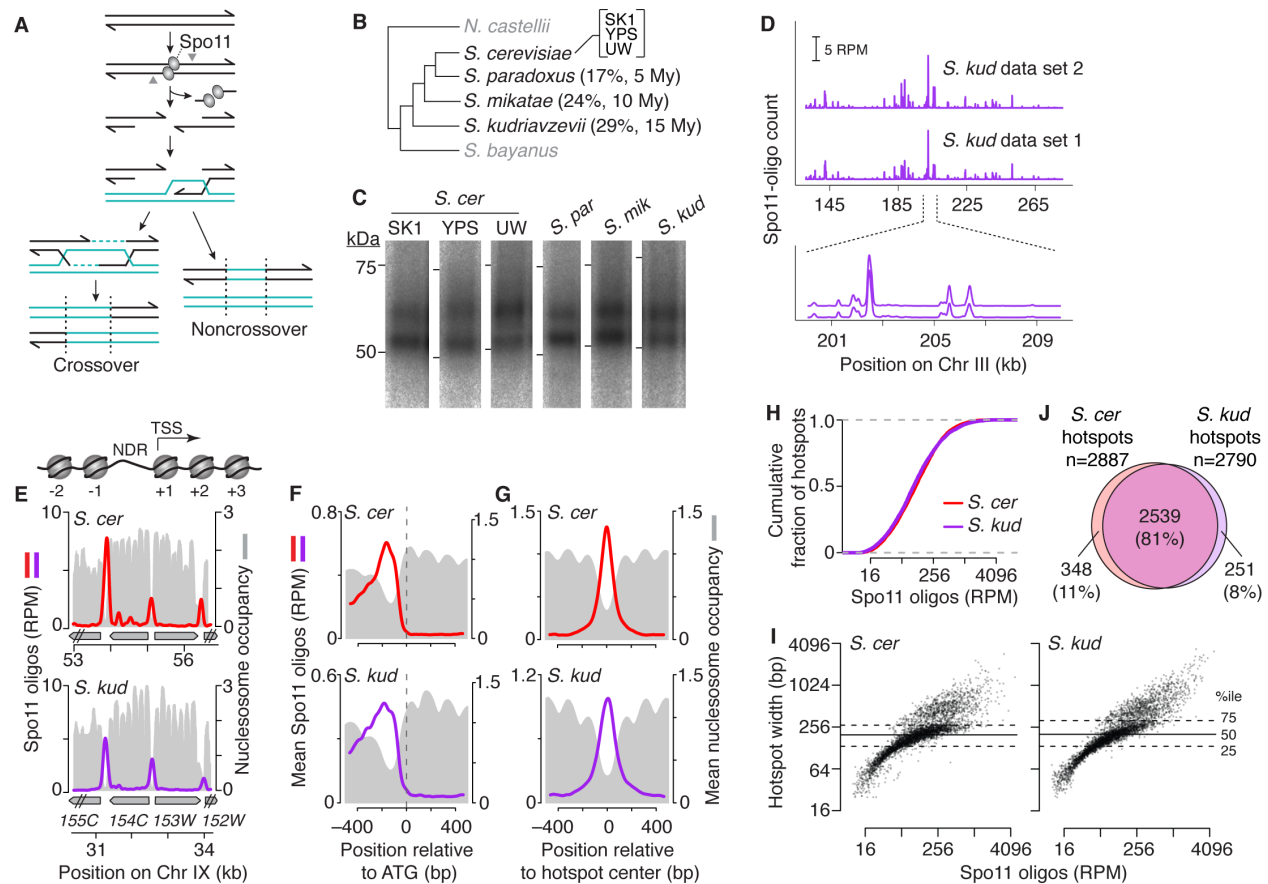
20. M. Kellis, N. Patterson, M. Endrizzi, B. Birren, E. S. Lander, Sequencing and comparison of yeast species to identify genes and regulatory elements. *Nature* **423**, 241 (2003).
21. M. Sasaki, S. E. Tischfield, M. van Overbeek, S. Keeney, Meiotic recombination initiation in and around retrotransposable elements in *Saccharomyces cerevisiae*. *PLoS Genet* **9**, e1003732 (2013).
22. X. Zhu, S. Keeney, High-resolution global analysis of the influences of Bas1 and Ino4 transcription factors on meiotic DNA break distributions in *Saccharomyces cerevisiae*. *Genetics*, (in press).
23. L. E. Berchowitz, S. E. Hanlon, J. D. Lieb, G. P. Copenhaver, A positive but complex association between meiotic double-strand break hotspots and open chromatin in *Saccharomyces cerevisiae*. *Genome Res* **19**, 2245 (2009).
24. I. Tirosh, N. Sigal, N. Barkai, Divergence of nucleosome positioning between two closely related yeast species: genetic basis and functional consequences. *Molecular systems biology* **6**, 365 (2010).
25. A. M. Tsankov, D. A. Thompson, A. Socha, A. Regev, O. J. Rando, The role of nucleosome positioning in the evolution of gene regulation. *PLoS Biol* **8**, e1000414 (2010).
26. K. Tsui *et al.*, Evolution of nucleosome occupancy: conservation of global properties and divergence of gene-specific patterns. *Mol Cell Biol* **31**, 4348 (2011).
27. M. Lichten, in *Recombination and Meiosis*, R. Egel, D. H. Lankenau, Eds. (Springer-Verlag, Heidelberg, 2008), vol. 3, pp. 165-193.
28. L. Goodstadt, C. P. Ponting, Is the control of recombination conserved among diverse eukaryotes? *Heredity* **106**, 710 (2011).
29. F. Cole *et al.*, Mouse tetrad analysis provides insights into recombination mechanisms and hotspot evolutionary dynamics. *Nat Genet* **46**, 1072 (2014).
30. M. A. White, M. Dominska, T. D. Petes, Transcription factors are required for the meiotic recombination hotspot at the *HIS4* locus in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* **90**, 6621 (1993).
31. F. Baudat, Y. Imai, B. de Massy, Meiotic recombination in mammals: localization and regulation. *Nat Rev Genet* **14**, 794 (2013).
32. C. L. Baker *et al.*, PRDM9 drives evolutionary erosion of hotspots in *Mus musculus* through haplotype-specific initiation of meiotic recombination. *PLoS Genet* **11**, e1004916 (2015).
33. S. Myers *et al.*, Drive against hotspot motifs in primates implicates the *PRDM9* gene in meiotic recombination. *Science* **327**, 876 (2010).
34. F. Pratto *et al.*, DNA recombination. Recombination initiation maps of individual human genomes. *Science* **346**, 1256442 (2014).
35. E. J. Louis, The chromosome ends of *Saccharomyces cerevisiae*. *Yeast* **11**, 1553 (1995).
36. B. Rockmill, K. Voelkel-Meiman, G. S. Roeder, Centromere-proximal crossovers are associated with precocious separation of sister chromatids during meiosis in *Saccharomyces cerevisiae*. *Genetics* **174**, 1745 (2006).
37. A. D. Peters, A combination of cis and trans control can solve the hotspot conversion paradox. *Genetics* **178**, 1579 (2008).
38. D. B. Kaback, V. Guacci, D. Barber, J. W. Mahon, Chromosome size-dependent control of meiotic recombination. *Science* **256**, 228 (1992).
39. D. Thacker, N. Mohibullah, X. Zhu, S. Keeney, Homologue engagement controls meiotic DNA break number and distribution. *Nature* **510**, 241 (2014).
40. G. Fischer, S. A. James, I. N. Roberts, S. G. Oliver, E. J. Louis, Chromosomal evolution in *Saccharomyces*. *Nature* **405**, 451 (2000).

41. C. S. Smukowski Heil, C. Ellison, M. Dubin, M. A. Noor, Recombining without hotspots: A comprehensive evolutionary portrait of recombination in two closely related species of *Drosophila*. bioRxiv doi: 10.1101/016972. (2015).
42. S. E. Zanders *et al.*, Genome rearrangements and pervasive meiotic drive cause hybrid infertility in fission yeast. *eLife* **3**, e02630 (2014).
43. K. R. Fowler, M. Sasaki, N. Milman, S. Keeney, G. R. Smith, Evolutionarily diverse determinants of meiotic DNA break and recombination landscapes across the genome. *Genome Res* **24**, 1650 (2014).
44. G. Coop, M. Przeworski, An evolutionary view of human recombination. *Nat Rev Genet* **8**, 23 (2007).
45. P. J. Keller, M. Knop, Evolution of mutational robustness in the yeast genome: a link to essential genes and meiotic recombination hotspots. *PLoS Genet* **5**, e1000533 (2009).
46. J. Berglund, J. Quilez, P. F. Arndt, M. T. Webster, Germline methylation patterns determine the distribution of recombination events in the dog genome. *Genome biology and evolution* **7**, 522 (2015).
47. K. Choi, I. R. Henderson, Meiotic recombination hotspots - a comparative view. *Plant J* **83**, 52 (2015).
48. S. Singhal *et al.*, Stable recombination hotspots in birds. (co submission).
49. G. Liti *et al.*, High quality de novo sequencing and assembly of the *Saccharomyces arboricolus* genome. *BMC genomics* **14**, 69 (2013).
50. D. R. Scannell *et al.*, The Awesome Power of Yeast Evolutionary Genetics: New Genome Sequences and Strain Resources for the *Saccharomyces sensu stricto* Genus. *G3* **1**, 11 (2011).
51. K. Kugou *et al.*, Rec8 guides canonical Spo11 distribution along yeast meiotic chromosomes. *Mol Biol Cell* **20**, 3064 (2009).
52. M. J. Neale, S. Keeney, End-labeling and analysis of Spo11-oligonucleotide complexes in *Saccharomyces cerevisiae*. *Methods Mol Biol* **557**, 183 (2009).
53. H. Murakami, S. Keeney, Temporospatial coordination of meiotic DNA replication and recombination via DDK recruitment to replisomes. *Cell* **158**, 861 (2014).
54. C. A. Brown, A. W. Murray, K. J. Verstrepen, Rapid expansion and functional divergence of subtelomeric gene families in yeasts. *Curr Biol* **20**, 895 (2010).
55. F. Baudat *et al.*, PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice. *Science* **327**, 836 (2010).
56. I. J. Tsai, D. Bensasson, A. Burt, V. Koufopanou, Population genomics of the wild yeast *Saccharomyces paradoxus*: Quantifying the life cycle. *Proc Natl Acad Sci U S A* **105**, 4957 (2008).
57. G. Liti, The fascinating and secret wild life of the budding yeast *S. cerevisiae*. *eLife* **4**, (2015).
58. L. J. Johnson *et al.*, Population genetics of the wild yeast *Saccharomyces paradoxus*. *Genetics* **166**, 43 (2004).
59. D. M. Ruderfer, S. C. Pratt, H. S. Seidel, L. Kruglyak, Population genomic analysis of outcrossing and recombination in yeast. *Nat Genet* **38**, 1077 (2006).
60. K. Brick, F. Smagulova, P. Khil, R. D. Camerini-Otero, G. V. Petukhova, Genetic recombination is directed away from functional genomic elements in mice. *Nature* **485**, 642 (2012).
61. P. P. Khil, F. Smagulova, K. M. Brick, R. D. Camerini-Otero, G. V. Petukhova, Sensitive mapping of recombination hotspots using sequencing-based detection of ssDNA. *Genome Res* **22**, 957 (2012).

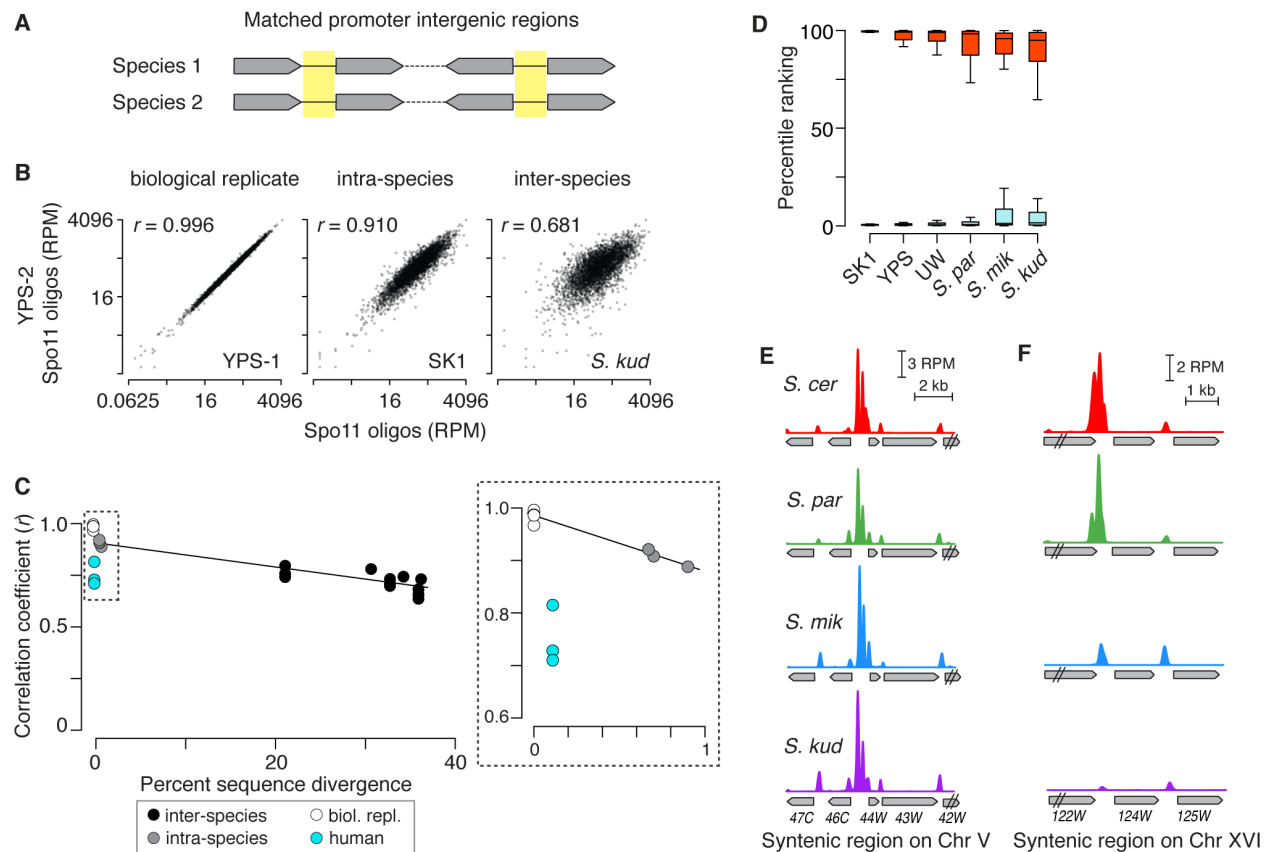
**Acknowledgments:** Submission of Spo11-oligo and MNase-seq data to the Gene Expression Omnibus (GEO) repository is pending. We are grateful to Duncan Greig, Chris Hittinger, Gianni Liti, Ed Louis, Carolin Müller, and Jeremy Roop for providing strains and advice



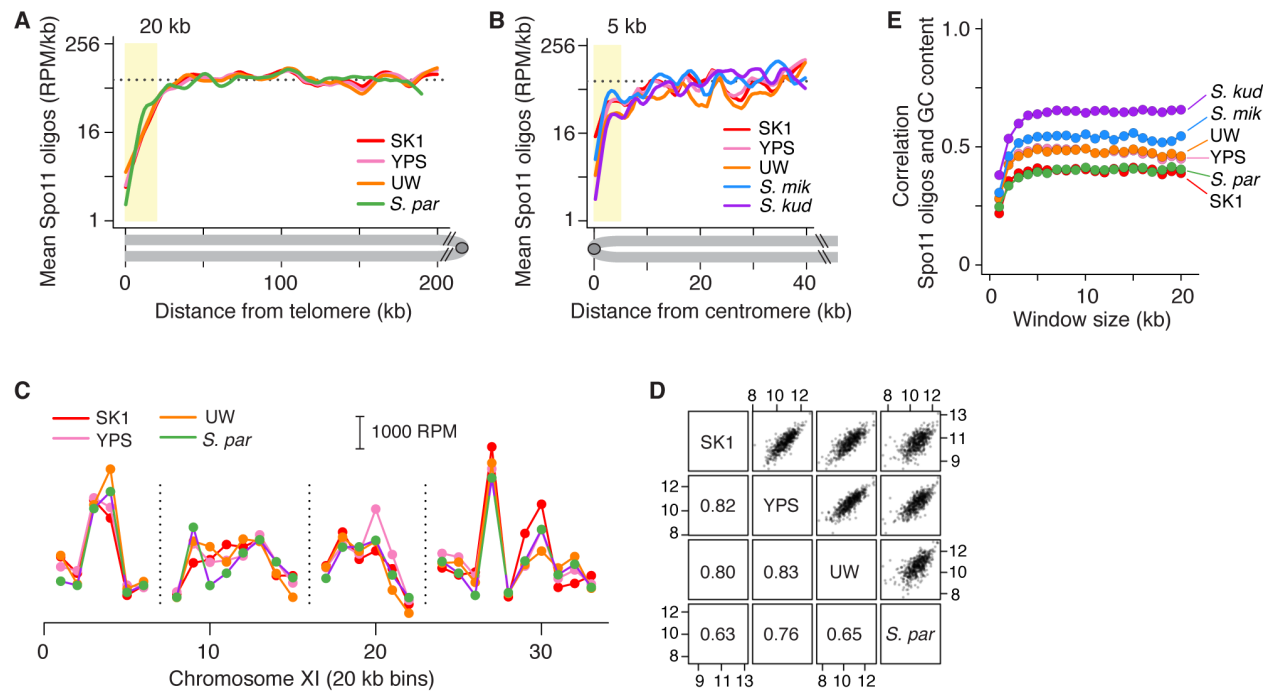
on culturing *Saccharomyces* species, Agnes Viale and the Integrated Genomics Operation (MSKCC) for sequencing, Nick Socci (Bioinformatics Core Facility, MSKCC) for mapping Spo11-oligo reads, Stewart Shuman for gifts of T4 RNA ligase, and Sarah Kim, Julian Lange, Neeman Mohibullah, Sam Tischfield, and Xuan Zhu for helpful discussions on the bioinformatic analyses. We thank Molly Przeworski for discussions and communicating data prior to publication, and for providing helpful comments on the manuscript. This work was supported by NIH grant R01 GM058673. IL was supported in part by NIH fellowship F31 GM097861. SK is an investigator of the Howard Hughes Medical Institute.



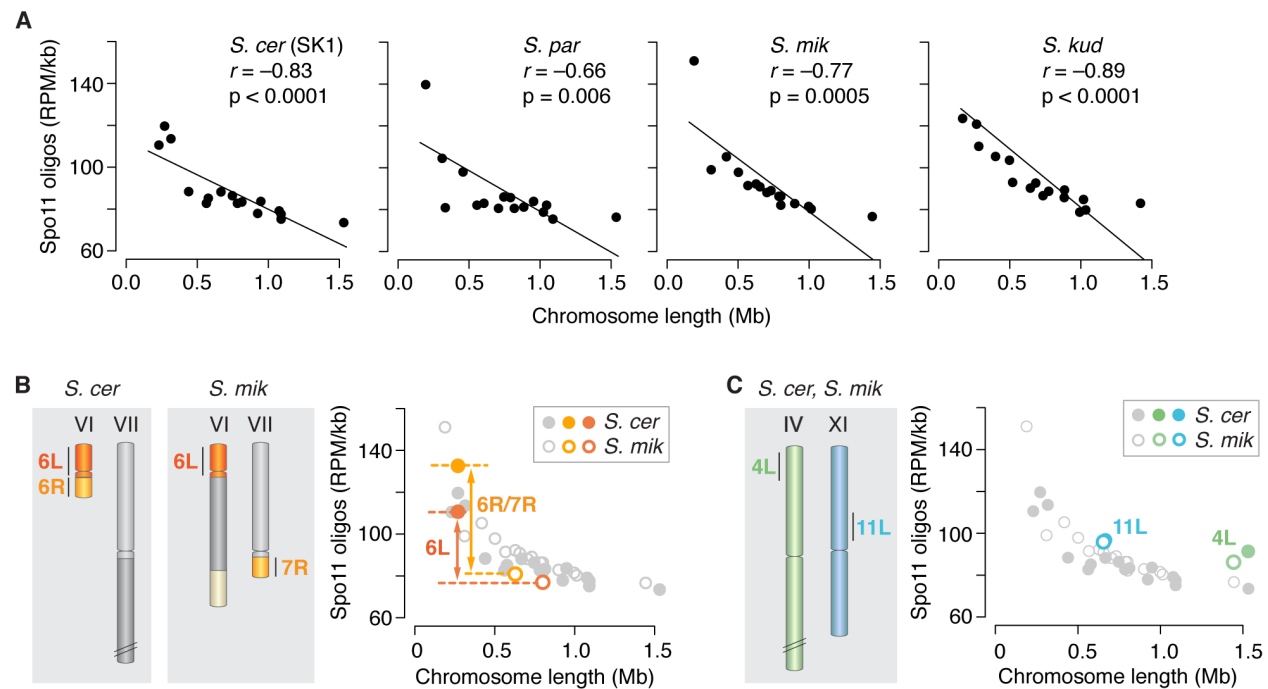
**Fig. 1.** Conserved targeting of DSBs to promoters. (A) Meiotic recombination. Spo11 generates covalent protein-linked DSBs; endonucleolytic cleavage (grey arrowheads) liberates Spo11 bound to short oligos exploited to map DSBs. DSB ends are resected and repaired to yield crossover or non-crossover products. The broken chromosome (black) copies information from the uncut allele (teal). (B) *Saccharomyces* phylogeny (schematic of outgroup (*N. castellii*)-rooted neighbor-joining tree using Rad51 protein alignment). Black, species/strains in this study; genic percent sequence divergence from *S. cerevisiae* (49) and estimated time since last common ancestor (18, 50) are shown. YPS, YPS128 strain; UW, UWOPS03-461.4 strain. (C) Conserved sizes of Spo11-oligo complexes. Autoradiographs of immunoprecipitated, radiolabelled complexes separated by SDS-PAGE. (D) Reproducibility of *S. kudriavzevii* Spo11-oligo maps. RPM, reads per million mapped. (E) Overlap of DSB hotspots with promoter NDRs is evolutionarily conserved. The cartoon depicts typical yeast promoter chromatin structure, with an NDR upstream of the transcription start site (TSS). Sample region (around *YIL154C*) compares Spo11 oligos with the nucleosome map (MNase-seq read depth relative to genome average). (F) Average Spo11-oligo and nucleosome profiles around start codons (*S. cer* SK1, *n*=5766 genes; *S. kud*, *n*=5578). (G) Average Spo11-oligo and nucleosome profiles at hotspots (*S. cer* SK1, *n*=4099; *S. kud*, *n*=3976). (H) Hotspot intensity varies over similar smooth continua in *S. cerevisiae* (SK1) and *S. kudriavzevii*. (I) Similar distributions of widths vs. Spo11-oligo counts in hotspots from *S. cerevisiae* (SK1) and *S. kudriavzevii*. (J) Conservation of hotspot position. Most promoter-containing IGRs hosting Spo11-oligo hotspots in *S. cerevisiae* (SK1) also had hotspots in *S. kudriavzevii* (see **fig. S3G** for more detail and other species). Spo11-oligo profiles were smoothed with 201-bp (panels D,E) or 75-bp (panels F,G) Hann window.



**Fig. 2.** Conservation of hotspot strength. (A) Promoter-containing IGRs were matched between species based on conservation of both flanking genes. (B) Comparison of Spo11-oligo counts within 3426 IGRs that were matched in all four species. Correlation coefficients for the log<sub>2</sub>-transformed data are shown (Pearson's  $r$ ). (C) Spo11-oligo counts in promoter IGRs remain highly correlated despite wide sequence divergence. Correlation coefficients (as in panel B) are plotted against the median sequence divergence within the IGRs, which is substantially greater than the coding sequence divergence (**Fig. 1B**). Black line highlights relation between the yeast comparisons; it is not a regression line. Human data (from ref. 34) are for three men with identical or similar *PRDM9* alleles (37,345 hotspots (see **fig. S4B**); each had ~0.1% sequence difference from the reference genome (34)). (D) The hottest hotspots have stayed hot, and the coldest have stayed cold. Percentile rankings in other strains and species are shown for the matched promoter IGRs with the most (red) and least (cyan) Spo11 oligos in SK1 (top and bottom 1%). Box plots are as in **fig. S3D**. (E,F) Examples of a strong Spo11-oligo hotspot from SK1 whose heat is conserved (E, *YEL046C*) and one whose heat is not (F, *YPR124W*).



**Fig. 3.** Conservation of large-scale features of the DSB landscape. (A and B) Telomere-proximal and pericentric DSB suppression. Lines are smoothed fit (Lowess) of Spo11-oligo densities in 500 bp bins averaged across 32 chromosome arms. Dashed line, genome average in SK1; yellow shading, DSB suppression zones. Genome assemblies are not complete enough to evaluate telomeric zones for *S. mikatae* or *S. kudriavzevii*. (C,D) Large-scale hot and cold interstitial domains are conserved. Interstitial segments (excluding chromosome ends and peri-centromeres) were defined as syntenic between *S. cerevisiae* and *S. paradoxus* if orthologous genes were in the same order in both species. Spo11-oligo counts were then summed in these syntenic segments divided into 20-kb bins (**Table S5**). A representative example is shown in panel C (vertical dashed lines denote breaks in synteny, many of which reflect unresolved annotation discrepancies) and genome-wide scatter plots and correlation coefficients are in panel D. See also **fig. S5**. (E) Correlation (Pearson's  $r$ ) between mean Spo11-oligo counts and GC content binned in windows of varying size.



**Fig. 4.** DSB density is influenced by chromosome length. (A) The negative correlation between chromosome length and DSB density is conserved. Each point is one chromosome. (B) A natural experiment demonstrating chromosome length-dependent DSB control. The schematic illustrates a pair of syntenic segments on chromosomes of different size in *S. cerevisiae* and *S. mikatae*. The plot shows that Spo11-oligo density is higher on these segments in *S. cerevisiae* (when on a short chromosome) than in *S. mikatae* (when on longer chromosomes). (C) Control syntenic regions that reside on similarly sized chromosomes have equivalent Spo11-oligo densities in both species.