

Global mapping of meiotic recombination hotspots and coldspots in the yeast *Saccharomyces cerevisiae*

Jennifer L. Gerton^{*†}, Joseph DeRisi^{†*}, Robert Shroff[§], Michael Lichten[§], Patrick O. Brown[‡], and Thomas D. Petes^{*}

^{*}Department of Biology and Curriculum in Genetics and Molecular Biology, University of North Carolina, Chapel Hill, NC 27599-3280; [†]Department of Biochemistry, Stanford University School of Medicine, Howard Hughes Medical Institute, Stanford, CA 94305-5428; [‡]Department of Biochemistry and Biophysics, University of California, 513 Parnassus Avenue, San Francisco, CA 94143-0448; and [§]Laboratory of Biochemistry, National Cancer Institute, National Institutes of Health, Building 37, Bethesda, MD 20892

This contribution is part of the special series of Inaugural Articles by members of the National Academy of Sciences elected on April 27, 1999.

Contributed by Thomas D. Petes, August 10, 2000

In the yeast *Saccharomyces cerevisiae*, meiotic recombination is initiated by double-strand DNA breaks (DSBs). Meiotic DSBs occur at relatively high frequencies in some genomic regions (hotspots) and relatively low frequencies in others (coldspots). We used DNA microarrays to estimate variation in the level of nearby meiotic DSBs for all 6,200 yeast genes. Hotspots were nonrandomly associated with regions of high G + C base composition and certain transcriptional profiles. Coldspots were nonrandomly associated with the centromeres and telomeres.

Hotspots are genomic regions with unusually high levels of recombination (1). From studies in yeast, several generalizations concerning hotspots can be made. First, most hotspots are intergenic rather than intragenic (2). Second, genetically defined hotspots are associated with local double-strand DNA breaks (DSBs) (1). Third, DSBs usually occur in DNase I-sensitive regions (3). Fourth, activity of the *HIS4* hotspot in *Saccharomyces cerevisiae* requires the binding of transcription factors in the hotspot region (4–6), but does not require high levels of meiosis-specific transcription (7). Hotspots that require transcription factor binding are called α hotspots (8). Fifth, certain DNA sequences result in hotspots (β hotspots) that do not require the binding of known transcription factors (9).

Meiotic DSB formation requires Spo11p, a topoisomerase II-related protein that is transiently covalently attached to the 5' ends of the DNA fragments (10, 11). In wild-type yeast cells, the Spo11p is removed to allow subsequent steps in recombination. In strains with the *rad50S* mutation, however, Spo11p stays covalently attached to the broken DNA ends (11).

Coldspots in yeast have received less attention than hotspots. Lambie and Roeder (12) showed that the centromere of chromosome III reduced crossing-over and gene conversion of nearby markers, and Baudat and Nicolas (13) noted a lack of DSB sites near the centromere. Several researchers have found a relative lack of DSB sites in *rad50S* strains near the telomeres (13, 14).

Although observations concerning individual hotspots and coldspots have given clues as to the mechanism of recombination initiation, our ability to predict hotspots and coldspots from DNA sequence information is very limited. A complementary approach is to map hotspots and coldspots globally and to determine whether they share common DNA sequences and/or structural elements. Such mapping studies have been performed to map DSB sites to single-gene resolution on chromosome III (13) and to the resolution of a pulsed-field gel on chromosomes I, III, and VI (14, 15). By using DNA samples enriched for meiosis-specific DSBs as hybridization probes on DNA microarrays, we extended these analyses to measure the global distribution of DSBs at single-gene resolution.

Materials and Methods

Yeast Strains. FX4, FX6, and QFY105 are diploid *rad50S* strains that have been described (4). These strains (used in the microar-

ray analysis) are isogenic except for changes introduced by transformation. The strain MJL2688 (*MAT α /MAT α ura3/ura3 lys2/lys2 ho::LYS2/ho::LYS2 arg4-nsp,bgl/arg4-nsp,bgl SPO11-3XHA-His6::KanMX4/SPO11-3XHA-His6::KanMX4 rad50-K181::URA3/rad50-K181::URA3*) is isogenic with SK1 (11).

Purification of DSB-Enriched DNA. Cells were sporulated for 24 h (16). Genomic DNA was isolated by using a nonproteolyzing protocol (11). The DNA was sheared to a size of 2–3 kb. A portion of this DNA was used as a hybridization probe (total genomic DNA). From the remaining sample, we isolated DNA fragments covalently attached to proteins by using a glass fiber filter (11). This DNA was the DSB-enriched hybridization probe.

Microarray Analysis. Both genomic and DSB-enriched DNA samples were used as templates for a two-step random PCR amplification. About 25 ng of each DNA sample was used for two 8-min extensions with 2.7 mM Round A primer [5'-GTTTC-CCAGTCACGATCNNNNNNNNN (Integrated DNA Technologies, Coralville, IA), N being a mixture of all four nucleotides with 60% A + T and 40% G + C] at 37°C with 267 units/ml Sequenase 2.0 (Amersham Pharmacia). DNA was denatured at 94°C for 1.5 min, and cooled to 10°C, and Sequenase 2.0 was added between extensions. The resulting products were used as templates for 15 rounds of PCR using *Taq* polymerase (Perkin-Elmer) and 10 mM Round B primer (5'-GTTTCCAGTCACGATC). This DNA then was labeled by 20 additional rounds of PCR including fluorescently labeled Cy3-dUTP or Cy5-dUTP (Amersham Pharmacia). These DNA samples were mixed and used as probes for DNA microarrays (17). Microarray images were acquired by using a GenePix 4000A scanner (Axon Instruments, Foster City, CA). For image analysis, GENEPIX PRO 3.0 software was used (Axon Instruments).

Analysis of Hotspots Using Nylon Filters. DSB-enriched DNA was isolated from the SK1-derivative MJL2688 by chromatin immunoprecipitation (18) with the following alterations: (i) formaldehyde was omitted, (ii) the lysis buffer contained 1 M NaCl, and (iii) sonication resulted in DNA fragments of 1 kb. Hemagglutinin (HA)-tagged Spo11p was precipitated with anti-HA antibody 12CA5 (Roche Molecular Biochemicals) and Immobilized Protein G (Pierce). Immunoprecipitates were washed five times with lysis buffer with 1 M NaCl, three times with lysis buffer with 0.5 M NaCl, followed by two washes as described (18). DSB-enriched DNA and total genomic DNA (25 ng) were labeled by the incorporation of ³³P-dCTP by using the High Prime DNA Labeling system (Roche Molecular Biochemicals).

Genes were amplified individually from strain S288C, mostly using primers purchased from Research Genetics, Huntsville, AL. About 2 ng of each PCR product was spotted on Zeta-Probe

Abbreviations: DSB, double-strand DNA break; P1, probe 1; P2, probe 2.

GT membranes (Bio-Rad) (19). Hybridization methods were as described (20), except that 5× Denhardt’s solution replaced skim milk. Filters were analyzed on a BAS2000 PhosphorImager system after being exposed for 40–70 h and quantitated by using IMAGE GAUGE 3.3 software (Fuji Medical Systems, Stamford, CT). For each filter, the average background was subtracted from the measured intensity of all members of the array. The data were normalized by using the ratio of the median signal intensities of the array members measured from each filter. The normalized data for each gene were used to calculate the ratio of (Spo11p-enriched DNA)/(genomic DNA).

Data Analysis. Most of the data analysis was done by using GCG programs. The microarray data generated in our study has the website address: <http://derisilab.ucsf.edu/hotspots/>. Statistical analyses were done with the INSTAT 1.12 program for Macintosh.

Results and Discussion

Experimental Rationale. The mapping of hotspots and coldspots is based on the assumptions that: (i) Spo11p-catalyzed meiosis-specific DSBs are the initiating lesions for most recombination events in yeast, and (ii) DNA fragments from a *rad50S* strain (induced to undergo meiosis) that are covalently attached to proteins will be enriched for the Spo11p-associated DNA sequences adjacent to hotspots and deficient for the DNA sequences adjacent to coldspots.

Microarray Mapping of Hotspots and Coldspots. Two closely related *rad50S* yeast strains were used for the microarray analysis, FX4 and FX6 (4). FX4 has a small insertion upstream of *HIS4* that elevates the already high level of recombination at this locus (5), and FX6 has the same insertion upstream of *ARG4*.

We prepared two samples of meiotic DNA, one containing total genomic DNA (probe 1, P1) and one enriched for DNA fragments covalently bound to proteins, including Spo11p (probe 2, P2) (11); we will refer to the P2 sample as DSB-enriched. Southern analysis showed that P2 was enriched at least 100-fold for DNA fragments resulting from DSB formation (Fig. 1). P1 and P2 were labeled with different fluorescently tagged

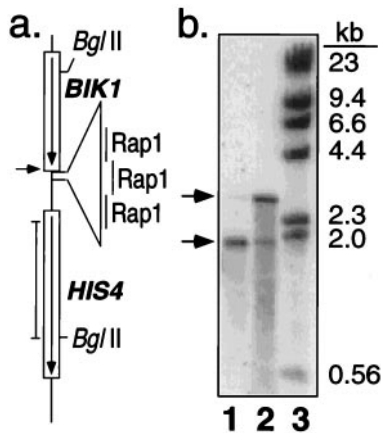


Fig. 1. Characterization of DSB-enriched probe. (a) Map of the *HIS4* recombination hotspot in FX4. An insertion of telomeric DNA with three Rap1 binding sites results in high levels of meiosis-specific DSBs at the position indicated by the arrow (4). The position of the probe (pDN42; 4) used in *b* is shown by the vertical line adjacent to the map. (b) Southern analysis of FX4 meiotic DNA samples used in the microarray experiments. Lanes 1 and 2 contain *Bgl*II-treated DSB-enriched DNA (P2) isolated by using glass fiber filters (11), and *Bgl*II-treated total genomic DNA (P1), respectively. Lane 3 contains *Hind*III-treated λ DNA. Arrows indicate the position of the intact 2.9-kb *HIS4-BIK1 Bgl*II fragment and the 1.9-kb fragment resulting from the meiosis-specific DSB.

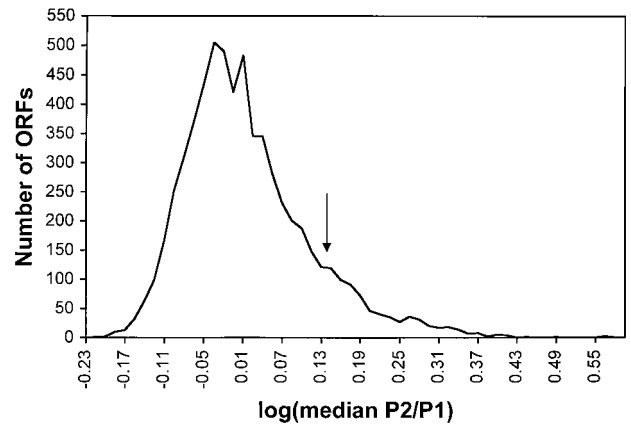


Fig. 2. Hotspot activity of yeast ORFs. To estimate DSB formation adjacent to each ORF, we measured the ratio of hybridization to a DSB-enriched probe (P2) to a total genomic probe (P1) in seven microarrays. The log of the median hybridization ratio for each of the 6,200 ORFs is graphed; ratios were grouped into bins of 0.015 log units. An arrow marks the separation between the ORFs with hybridization ratios ranked in the top eighth from those in the bottom seven-eighths.

nucleotides (Cy3-dUTP and Cy5-dUTP), mixed, and hybridized to DNA microarrays containing the 6,200 ORFs of the genome (17). Most DSBs occur within the intergenic regions (3, 13); because the probe DNA is derived from DNA fragments 2–3 kb in size and the intergenic regions in yeast average 500 bp (21), P2 DNA should contain portions of ORFs that border the intergenic regions.

Dual-wavelength images of microarrays were obtained by scanning at a resolution of 10 microns/pixel. The relative hybridization signal at each spot was determined by calculating the ratio of mean pixel intensity within a spot for both wavelengths, each wavelength measuring the hybridization of one of the two probes. The hybridization signals were normalized to a value of approximately 1 over all 6,200 ORFs. Thus, a ratio of hybridization of 1 represents an ORF with an average level of nearby DSBs. Seven ORF-containing microarrays were analyzed, four with FX4 and three with FX6. In four of these experiments (two with FX4 and two with FX6 DNA), P1 was labeled with the Cy5-dUTP (red fluorophore) and P2 was labeled with the Cy3-dUTP (green fluorophore); in the other three experiments, the labeling scheme was reversed. This procedure controls for the possibility of DNA sequences that preferentially label with one type of fluorophore.

Fig. 2 is a graph of the frequency distribution of all genomic ORFs with various P2/P1 hybridization ratios (log of the median P2/P1 ratio). The frequency distribution is asymmetric, with approximately one-eighth of the ORFs located in a shoulder of high P2/P1 ratios. We classified an ORF as “hot” if it ranked in the top 12.5% in at least five of the seven microarrays and “cold” if it ranked in the bottom 12.5% in at least five of seven experiments. A binomial expansion calculation predicts that these criteria will result in only three false-positive hot or cold ORFs. Thus, for our analysis, we operationally define hot and cold ORFs as those that reproducibly have hybridization ratios in the top and bottom eighths, respectively. This classification scheme is conservative, although somewhat arbitrary, and emphasizes the hottest and coldest ORFs in the genome.

By these criteria, we detected 303 hot ORFs, clustered into 177 hotspots (Table 1). Hot ORFs were considered to represent a single hotspot if they were adjacent or if they were separated by a single nonhot ORF. We found 49 cold ORFs, clustered into 40 coldspots (Table 2). The positions of the hotspots and coldspots along the chromosomes are shown in Fig. 3.

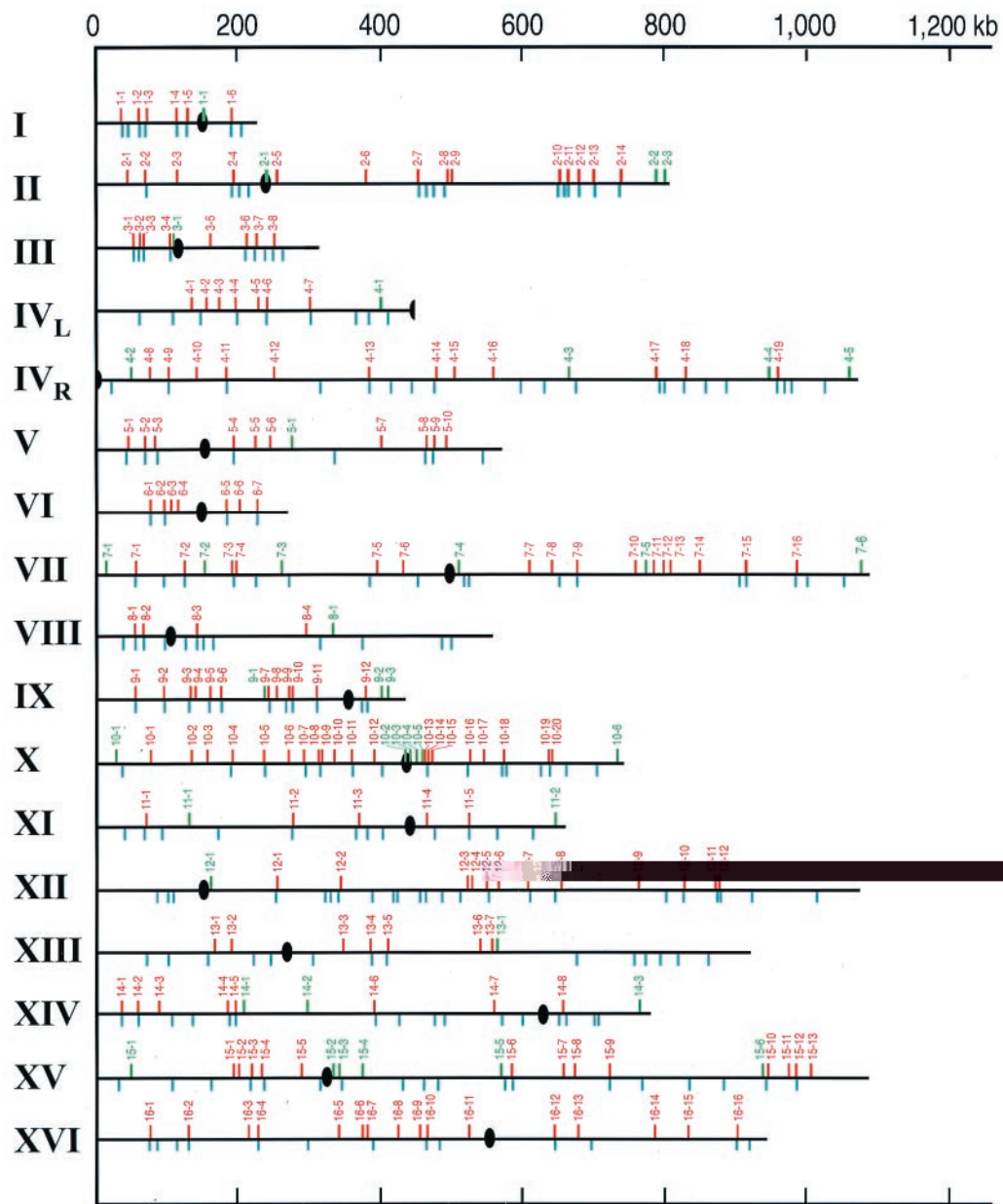


Fig. 3. Chromosomal map positions of the hotspots (red) and coldspots (green). Ovals indicate centromeres. Blue lines represent peaks of G + C base composition that are 3% higher than the average for the chromosome.

The relative strengths of the hot ORFs were determined by the following procedure: (i) within each microarray data set, each ORF was assigned a ranking based on the P2/P1 hybridization ratio, (ii) the median ranking of each ORF within the seven data sets was determined, and (iii) the 303 hot ORFs were reordered by their median ranking and renumbered with the highest ranking assigned a value of 1. These values are shown in Table 1 (first number in parentheses); because some hotspots share the same ranking, the values extend from 1 to 134. An analogous method was used to rank the coldspots (Table 2). In Tables 1 and 2, we also show the median ratio of P2/P1 hybridization (second number in parentheses). We chose to use the median ranking, instead of the median ratio of hybridization, to compare hotspots and coldspots because the rank positions in different microarrays varied less than the ratios of hybridization.

The highest ranked ORFs (YGR176W and YGR177C) had

ratios of hybridization (P2/P1) of about five. The weakest hot ORFs had ratios of hybridization about 1.4-fold above the average. As described above, the strains FX4 and FX6 were isogenic except for small recombination-stimulating insertions located between *BIK1* and *HIS4* in FX4 and *ARG4* and *DED81* in FX6. *ARG4* (YHR018C) in FX4 had a hotspot ranking of 134, and a hybridization ratio of 1.4 (Table 1); in FX6, the *ARG4* ranking was 2, with a hybridization ratio of 5.5. No significant difference was observed for the *HIS4/BIK1* hotspot in FX4 compared with FX6 (rankings of 4 and 2, and hybridization ratios of 3.2 and 4.2, respectively), possibly because of the very high level of hotspot activity in the wild-type (FX6) strain. We also examined two microarrays (one using FX4- and one using FX6-derived probes) that had both intergenic regions and ORFs. The results obtained were consistent with those seen with the ORF-only arrays. For example, 76% of the top 25 ranked

Table 1. Recombination activity of hotspot ORFs

HS# [†]	ORFs (rank) [*]	HS# [†]	ORFs (rank) [*]	HS# [†]	ORFs (rank) [*]	HS# [†]	ORFs (rank) [*]
1-1*	YAL061W(16/2.3)	4-17	YDR382W(82/1.6)	<u>2-1*</u>	-154C(20/2.0)	12-7*	-232W(50/2.0)
	-62W(57/1.8)	4-18*	YDR405W(101/1.5)		-155C(12/2.2)		-234W(101/1.5)
1-2*	YAL043C(120/1.5)	4-19*	YDR476C(100/1.7)	9-2*	YIL136W(118/1.5)		-236C(112/1.5)
	-44C(66/1.9)	5-1*	YEL058W(51/1.8)	9-3*	YIL121W(85/1.8)	12-8*	YLR255C(120/1.4)
1-3*	YAL036C(76/1.9)		-59C(81/1.8)		-123W(105/1.6)		-257W(109/1.7)
	-37W(122/1.6)		-59W(60/1.9)	9-4	YIL117C(111/1.5)	12-9	YLR316C(102/1.6)
	-38W(3/3.0)!!	5-2*	YEL044W(69/1.9)		-118W(107/1.5)	12-10*	YLR348C(48/2.0)
	-39C(48/1.9)		-45C(66/2.1)	9-5*	YIL108W(125/1.6)		-349C(99/1.6)
1-4*	YAL020C(71/1.7)		-46C(56/2.1)	<u>2-6*</u>	YIL099W(5/2.6)		-350W(30/2.0)
1-5*	YAL012W(71/2.0)	5-3	YEL037C(132/1.5)		-100W(40/2.1)	12-11*	YLR374C(90/1.6)
	-13W(48/2.0)	5-4*	YER019C-A(55/1.8)		-101C(104/1.8)	12-12*	!!YLR378C(76/1.8)
1-6*	YAR042W(102/1.6)	5-5	YER034W(131/1.4)	9-7*	YIL065C(61/1.9)		-379W(43/2.0)
2-1	YBL093C(79/1.9)		-35W(70/1.8)	9-8	YIL053W(120/1.7)	13-1	!!YML053C(85/1.6)
2-2*	YBL082C(115/1.5)	5-6	YER046W(105/1.6)		-54W(68/1.8)	13-2	YML042W(22/2.3)
2-3	YBL055C(123/1.6)	5-7	YER119C-A(133/1.4)		-55C(88/1.6)		-43C(63/1.8)
2-4*	YBL015W(89/1.7)	5-8*	YER145C(71/1.8)		-56W(66/1.7)	13-3	YMR038C(70/1.9)
	-16W(19/2.1)		-146W(120/1.5)		-57C(103/1.6)	13-4*	YMR055C(126/1.5)
2-5	YBR009C(124/1.5)		-147C(78/1.8)	9-9*	YIL046W(15/2.4)		-56C(10/2.4)
2-6	YBR072W(101/1.8)		-149C(105/1.5)	9-10*	YIL040W(118/1.5)	13-5*	YMR070W(42/2.0)
2-7*	YBR106W(32/1.1)	<u>5-9*</u>	YER152C(9/2.2)		-41W(56/2.0)	13-6	YMR135C(55/2.0)
	-107C(127/1.6)		-153C(55/2.0)		-42C(30/2.0)	13-7	YMR146C(72/1.8)
2-8*	YBR127C(30/2.2)		-154W(120/1.5)	9-11*	YIL023C(87/1.7)	14-1*	YNL322C(65/1.9)
	-128C(130/1.5)	5-10	YER159C(40/2.1) ³	9-12*	YIR017C(111/1.5)	14-2*	YNL305C(39/1.9)
2-9	YBR132C(109/1.6)	<u>6-1*</u>	YFL030W(4/2.6)	10-1	YJL189W(107/1.6)	14-3	YNL289W(117/1.5)
2-10*	YBR214W(20/2.1)		-31W(61/1.8)	10-2	YJL151C(113/1.5)	14-4*	YNL244C(49/2.0)
2-11*	YBR219C(109/1.7)		-32W(96/1.6)		-152W(60/1.9)		-245C(42/1.9)
	-220C(109/1.6)	6-2*	YFL021W(26/2.1)		-153C(92/1.8)	14-5*	YNL240C(63/1.8)
2-12*	YBR230C(32/2.2)		-22C(117/1.5)	10-3	YJL134W(92/1.7)		-241C(16/2.2)
	-231C(74/2.0)	6-3	YFL016C(78/1.6)	10-4*	YJL116C(121/1.5)	14-6*	YNL125C(119/1.7)
2-13*	YBR240C(16/2.3)	6-4	YFL010C(14/2.2)		-117W(79/1.8)	14-7	YNL036W(45/2.3)
	-241C(12/2.5)	6-5*	YFR017C(101/1.6)		-119C(105/1.6)	14-8*	YNR015W(106/1.5)
2-14	YBR267W(78/1.9)		-18C(112/1.5)	10-5*	-120W(121/1.6)	15-1	-17W(81/1.8)
	-268W(84/1.6)	6-6	YFR023W(103/1.6)		YJL100W(28/2.6)	15-2	YOL073C(121/1.4)
3-1*	YCL040W(104/1.8)		-24C(78/1.6)	10-6	!!-101C(93/1.8)	15-2	YOL067C(99/1.6)
3-2*	YCL033C(66/1.9)		-25C(25/2.1)	<u>10-6*</u>	YJL088W(90/1.9)	15-3*	YOL058W(130/1.4)
	-34W(79/1.9)		-26C(96/1.5)		YJL077C(97/1.8)		-59W(79/1.9) ³
3-3*	YCL029C(129/1.5)!!	6-7*	YFR033C(56/1.8)		-79C(6/2.3)	15-4*	YOL052C(29/2.3)
	-30C(2/4.2)		-34C(50/2.0)	10-8*	YJL066C(124/1.5)		-53W(67/1.7)
3-4*	YCL009C(95/1.7)		-35C(58/1.9)	10-9*	YJL063C(51/2.0)	15-5	YOL020W(72/1.9)
	-10C(121/1.5)		-36W(97/1.5)	10-10	YJL052W(46/1.9)	15-6*	YOR138C(56/2.1)
3-5	YCR025C(77/1.7)		-37C(19/2.3)		-53W(99/1.6)		-139C(54/1.9)
3-6*	YCR046C(7/2.4)!!	7-1*	YGL235W(82/1.7)	10-11*	YJL043W(56/1.9)	15-7	-140W(90/1.9)
	-47C(59/1.8)!!		-236C(87/1.7)	10-12	YJL026W(40/2.0)		YOR173W(32/2.1)
	-48W(41/2.0)	7-2*	YGL198W(13/2.3)		-27C(56/1.9)	15-8	YOR179C(69/1.8)
	-49C(20/2.0)		-199C(34/1.9)	10-13*	YJR016C(35/2.1)	15-9*	YOR203W(113/1.5)
	-50C(78/2.0)		-200C(42/2.1)		-17C(116/1.6)	15-10*	YOR331C(25/2.0)
	-51W(17/2.1)	7-3	YGL165C(94/1.6)	10-14*	YJR021C(132/1.5)		-332W(24/2.0)
	-52W(126/1.5)		-166W(75/1.7)	<u>10-15</u>	YJR024C(5/2.8)		-333C(118/1.5)
3-7*	YCR061W(28/2.0)	7-4*	YGL162W(30/2.2)		-25C(8/2.6)!!	15-11	YOR344C(94/1.7)
	-63W(128/1.5)	7-5	YGL056C(117/1.5)	10-16	YJR049C(118/1.5)	15-12*	YOR348(23/2.1)
3-8*	YCR077C(116/1.5)	7-6	YGL035C(102/1.6)	10-17	YJR057W(59/2.0)	15-13	YOR374W(97/1.6)
	-79W(85/1.8)		!!-37C(51/1.8)		-58C(32/2.1)!!		-376W(124/1.5)
4-1	YDL181W(125/1.5)	7-7	YGR059W(34/2.2)	10-18*	!!YJR072(125/1.5)	16-1*	!!YPL250C(53/1.9)
	-182W(87/1.8)		-60W(104/1.5)		!!-73C(36/2.1)		-252C(99/1.6)
4-2	YDL169C(83/1.6)	7-8	YGR078C(36/1.9)		-74W(16/2.4)	16-2*	YPL222W(79/1.6)
	-170W(129/1.6)		!!-79W(81/1.8)		-75W(116/1.6)		-223C(118/1.7)
4-3	YDL155W(109/1.6)	7-9*	YGR096W(18/2.1)		-77C(46/2.0)	16-3	YPL177C(92/1.6)
4-4*	YDL144C(47/1.9)	7-10	YGR136W(110/1.5)		-79W(116/1.5)	16-4*	YPL170W(108/1.5)
4-5	YDL131W(123/1.6)	7-11	YGR147C(74/1.8)	10-19*	YJR112W(97/106)		-171C(97/1.5)
4-6*	YDL124W(99/1.6)	7-12	YGR155W(43/2.1)	10-20*	YJR115W(21/2.2)	16-5	YPL111W(61/2.1)
4-7*	YDL086W(38/2.1)		-156W(11/2.2)		-116W(67/1.7)	16-6	YPL092W(66/1.9)
4-8	YDR037W(110/1.6)	7-13	YGR159C(64/1.7)	11-1*	YKL198C(80/1.7)	16-7	YPL088(62/1.8)
4-9*	!!YDR047W(65/1.9)		-161C(97/1.6)	11-2*	YKL087C(95/1.7)	16-8	YPL067C(90/1.8)
4-10	YDR073W(37/2.1)	7-14	YGR175C(12/2.3)	11-3	YKL035W(122/1.5)		-068C(109/1.6)
4-11*	YDR094W(132/1.5)!!		!!-176W(1/5.6)		-37W(117/1.5)	16-9	YPL055C(94/1.6)
4-12	YDR124W(27/2.3)		-177C(1/5.0)	11-4	YKR012C(84/1.7)	16-10*	YPL048W(102/1.8)
	-125C(90/1.8)	7-15*	YGR209(113/1.4)	11-5*	YKR048C(11/2.2)		-49C(58/1.8)
4-13*	!!YDR187C(43/1.9)		-210C(73/1.7)	12-1*	YLR055C(44/2.1)	16-11	YPL015C(121/1.6)
	-188W(21/2.4)	7-16*	YGR248W(114/1.5)		-56W(16/2.4)	16-12*	YPR036W(97/1.8)
4-14	YDR230W(78/1.8)	8-1*	YHL025W(64/1.8)	12-2*	YLR100W(111/1.5)		-37C(75/1.8)
	-231C(78/1.8)	8-2*	YHL020C(48/1.9)		-101C(124/1.7)	16-13	YPR063C(96/1.6)
	-232W(96/1.8)	8-3*	YHR018C(134/1.4)	12-3	YLR185W(65/1.9)	16-14	YPR123C(103/1.6)
	-233C(33/2.2)		!!-19C(113/1.5)!!	12-4	YLR188W(86/1.6)		-124W(127/1.5)
	-234W(88/1.7)	8-4	YHR096C(78/1.8)	12-5	YLR202C(75/1.7)	16-15	YPR151C(116/1.5)
4-15	YDR246W(91/1.7)	<u>2-1*</u>	YIL152W(120/1.5)	12-6	YLR213C(67/1.7)	16-16*	YPR183W(99/1.6)
4-16	YDR275W(118/1.5)		-153W(8/2.2)	12-7*	YLR231C(71/1.9)		

[†]Hotspots (HS) represent one or more ORFs ranked in the top 12.5% in five or more of the seven microarrays. Hotspots are numbered consecutively from the left end of the chromosome. Boldface indicates that a DSB has been mapped by Southern analysis; an asterisk indicates that the center of the hotspot is located within 2.5 kb of a peak of %(G+C) base composition that is at least 3% above the average %(G+C) for the chromosome. Underlining indicates that the hotspot has at least one ORF ranked among the 12 "hot"(G

Table 2. Recombination activity of coldspot ORFs

CS# [†]	ORFs (rank) [‡]	CS# [†]	ORFs (rank) [‡]	CS# [†]	ORFs (rank) [‡]	CS# [†]	ORFs (rank) [‡]
<u>1-1</u>	YAR002W(4/0.71)	7-2	YGL186C(9/0.72)	<u>10-2</u>	YJL001W(2/0.65)	14-1	YNL233W(19/0.76)
2-1	YBR001C(9/0.68)	<u>7-3</u>	YGL132W(1/0.6)		-R001W(12/0.75)		-235C(17/0.71)
2-2	YBR294W(5/0.71)	7-4	YGR008C(12/0.73)	<u>10-3</u>	YJR004C(2/0.65)		-237W(8/0.67)
2-3	YBR298C(11/0.72)		-009C(14/0.72)	<u>10-4</u>	YJR032W(12/0.68)		-238W(12/0.71)
3-1	YCL004W(18/0.75)	7-5	YGR143W(13/0.66)	<u>10-5</u>	YJR091C(4/0.64)	14-2	YNL183C(15/0.71)
4-1	YDL029W(17/0.75)	7-6	YGR289C(10/0.74)	10-6	YJR158C(6/0.72)	14-3	YNR070W(5/0.66)
4-2	YDR028C(5/0.65)	8-1	YHR111W(17/0.72)	11-1	YKL169C(12/0.69)	15-1	YOL147C(14/0.74)
4-3	YDR324C(9/0.71)	9-1	YIL067C(17/0.76)		-170W(12/0.64)	15-2	YOR003W(12/0.69)
4-4	YDR467C(12/0.65)	<u>9-2</u>	YIR023W(2/0.68)	11-2	YKR101W(7/0.62)	15-3	YOR008C(20/0.72)
	-468C(16/0.72)	9-3	YIR028W(9/0.68)		-102W(6/0.71)	15-4	YOR022C(5/0.64)
4-5	YDR573C(16/0.74)		-029W(9/0.69)	12-1	YLR005W(12/0.75)	15-5	YOR130C(14/0.65)
5-1	YER060W(12/0.73)	10-1	YJL214W(8/0.70)	13-1	YMR154C(15/0.65)	15-6	YOR329(6/0.75)
<u>7-1</u>	YGL257C(3/0.70)						

[†]Coldspots (CS) represent one or more ORFs ranked in the bottom 12.5% in five or more of the seven microarrays. Boldface indicates that Southern analysis was performed; no DSBs were detected for any coldspots examined. In addition, none of the coldspots was located within 2.5 kb of a GC peak. Underlining indicates that the coldspot has at least one ORF ranked among the seven "coldest" ORFs in the genome.

[‡]Coldspots were ranked in the same manner described in Table 1 with one exception. After assignment of the median ranking for each coldspot ORF, the rankings were reordered with the coldest ORF ranked 1. As in Table 1, the number after the ranking is the median ratio of hybridization (DBS-enriched probe/genomic probe).

hotspots from the ORF-only arrays were next to one of the intergenic regions ranked in the top 5% in the two ORF + intergenic microarrays.

Several arguments suggest that the microarray analysis reliably identifies hotspots and coldspots. First, all of the regions previously shown to have strong hotspot activity by genetic or physical criteria are represented among the 177 microarray-identified hotspots including: *HIS4*, *ARG4*, *DED81*, *CYS3*, *HIS2*, *ARE1/IMG1*, and *CDC19* (1). Second, most of the hotspots identified by the microarray analysis that we examined by Southern analysis had one or more nearby strong meiosis-specific DSBs; none of the seven coldspots examined had detectable DSBs (data not shown). The hotspots with mapped DSBs are shown in boldface in Table 1. Twenty of the 22 mapped DSB sites were located within intergenic regions.

For chromosome III, we compared the microarray mapping to other mapping procedures. Baudat and Nicolas (13) did a detailed mapping of DSB sites on chromosome III. In Fig. 4, we compare the "heat" of chromosome III ORFs (showed as log of the median P2/P1 hybridization ratios) as determined in our study (Fig. 4a) with the data from the Baudat and Nicolas analysis (Fig. 4b). Although there are quantitative differences in the patterns, both methods identify similar hot and cold regions on chromosome III. Some of the differences are likely to reflect strain-specific differences in hotspot activity. For example, the strain used in our study (related to S288c) has a very high level of hotspot activity at the *HIS4* locus (5) relative to the SK1 derivative examined by Baudat and Nicolas.

We also used a different method to examine chromosome III hotspots. PCR-amplified DNAs representing ORFs from chromosome III were spotted onto a nylon filter. Probe DNA was isolated from a *rad50S* SK1 derivative strain containing an epitope-tagged version of Spo11p. An antibody directed against the epitope was used to purify the Spo11p-associated DNA and that DNA sample was used as a hybridization probe. The resulting profile (Fig. 4c) of hybridization mimics that observed by Baudat and Nicolas. We also used the nylon filter array with an FX6-derived probe (isolated with the glass fiber filter method) (Fig. 4d). To obtain sufficient quantities of the hybridization probe for the microarray analysis and for the analysis shown in Fig. 4d, the DNA samples were amplified by PCR. We repeated the analysis using nylon filters with hybridization probes that had not been amplified. The patterns were similar (data not shown).

Distribution of Hotspots and Coldspots. All of the chromosomes have at least one hotspot (Fig. 3). There was a significant correlation between chromosome size and the number of hotspots [parametric (Pearson) linear correlation $P = 0.008$]; there was also a significant correlation ($P = 0.03$) between chromosome size and the number of coldspots.

Large chromosomes in yeast have fewer crossovers per kb than small chromosomes (22). Although larger chromosomes have more hotspots than smaller chromosomes, the density of hotspots on the smaller chromosomes is significantly greater than on the larger chromosomes ($P = 0.03$). In addition, the hotspots on the smaller chromosomes are significantly "hotter" than those on the larger chromosomes (linear regression analysis, $P = 0.05$). The simplest interpretation of these data are that both hotspot strength and hotspot density contribute to the high levels of crossovers observed for smaller chromosomes.

The average spacing between hotspots was 54 kb (median value of 36 kb). The average distance between hotspots for intervals that include the centromere was 117 kb (median value of 94 kb). The average interhotspot distance excluding the intervals containing the centromeres was 48 kb, with a median distance of 32 kb. The interhotspot distances for the centromere-containing intervals exceed those for the intervals that do not contain the centromere ($P < 0.0001$, Mann-Whitney nonparametric comparison), consistent with other studies indicating suppression of meiotic recombination near the centromere (12, 13).

The average distance from the telomere to the proximal hotspot was 103 kb. If the 177 hotspots were placed randomly along the chromosome, we would expect 14 located within 30 kb of the telomeres. None were observed, a significant departure from this expectation ($P < 0.001$ by χ^2 analysis). These results support previous observations indicating a lack of DSBs near the telomeres of *rad50S* yeast strains (13, 14). The distribution of coldspots is strikingly different from that observed for hotspots (Fig. 3). If the 40 coldspots were randomly distributed, we expect three within 30 kb of the telomeres and two within 30 kb of the centromeres. We observed 10 telomere-associated coldspots and eight centromere-associated coldspots, a significant departure from the random distribution ($P = 0.003$ by χ^2 analysis).

Correlation Between Peaks of G + C Base Composition and Hotspots. Sharp and Lloyd (23) noted that chromosome III had broad peaks of high G + C base composition located in the middle of the left and right arms. They pointed out that three of the four known hotspots were located within these regions. Similar

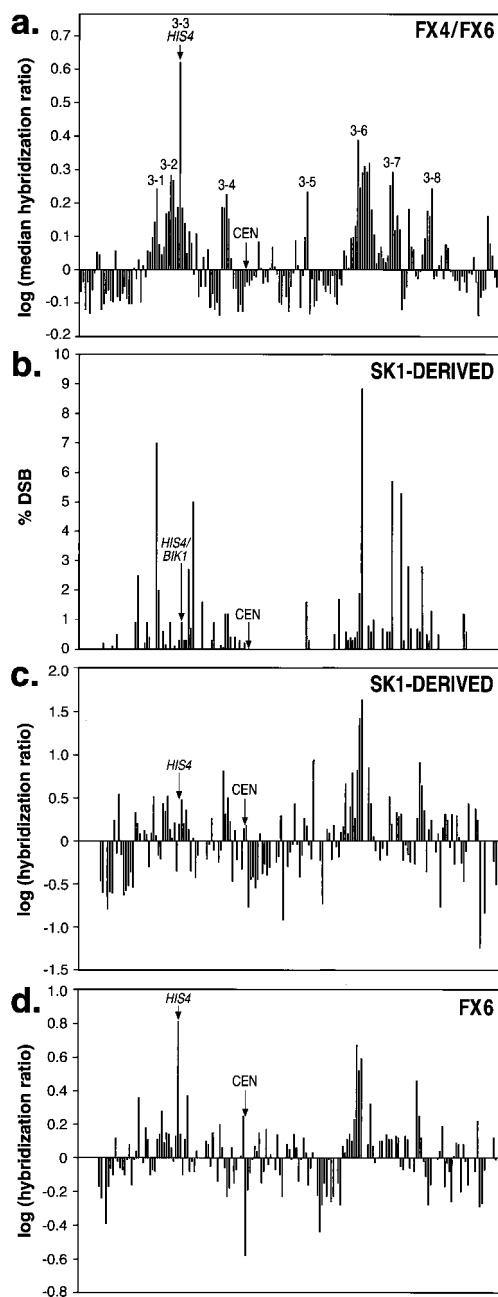


Fig. 4. Comparison of different methods of mapping hotspots and coldspots on chromosome III. (a) Summary of hybridization ratios using microarrays for all chromosome III ORFs. These data were based on yeast strains FX4 and FX6, except for the *HIS4* locus for which only the FX6 data were used. Each ORF was assigned the median hybridization ratio of the seven microarrays and we graphed the log function of this value. The 182 ORFs represented in the analysis (left to right on the x axis) are (Inclusive Stanford Genome Database numbering system): YCL076W-YCL073C, YCL069W-YCL016C, YCL014W-YCL001W, YCR001W-YCR020C, YCR020C-A, YCR021C-YCR024C, YCR024C-A, YCR025C-YCR029C, YCR029C-A, YCR030C-YCR077C, YCR079W-YCR097WA, and YCR097WB-YCR107W. The hotspot labels correspond to those listed in Table 1. (b) Summary of data of Baudat and Nicolas (13). These researchers monitored DSBs in an SK1 derivative by Southern analysis. (c) Analysis of hotspots in an SK1 derivative using DNA from immunoprecipitated Spo11p/hemagglutinin-DNA complexes as a probe of nylon filters containing chromosome III ORFs. The y axis shows the log of the average normalized ratio of hybridization (DSB-enriched probe/total genomic DNA probe). Gaps on the x axis reflect ORFs that were not present on the nylon filters. (d) Analysis of hotspots in FX6 using macroarrays. In this experiment, the probe was prepared by the method used in a.

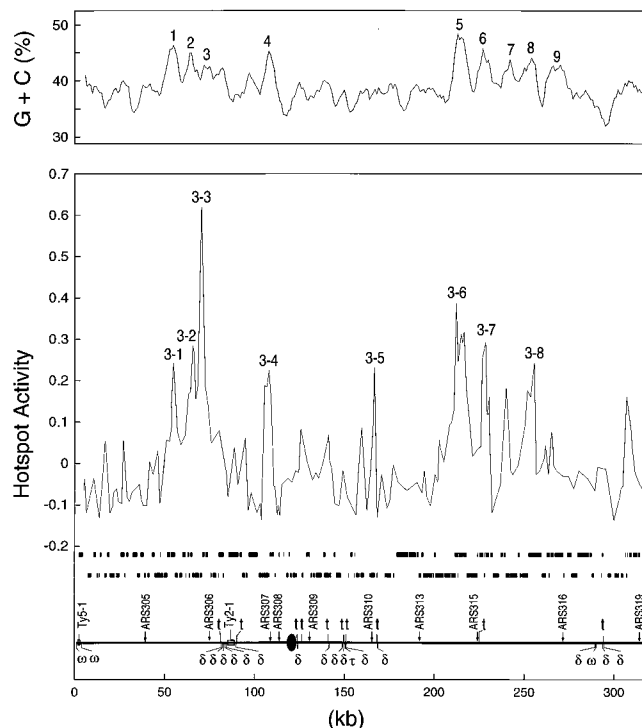


Fig. 5. Correlation of hotspots with peaks of high G + C content on chromosome III. (Upper) A scan of base composition at 1-kb intervals (5-kb sliding window). The marked peaks have a G + C base composition $\geq 3\%$ above the average for chromosome III (38.7%). In addition, we show the log values of the median hybridization ratio (DSB-enriched probe/total genomic DNA). (Lower) The ORFs (two lines of black rectangles) and the positions of a number of structural chromosomal elements are indicated.

patterns of variation in base composition have been detected in most yeast chromosomes by researchers using sliding windows of either 5 kb (19) or 30 kb (21). To determine whether hotspots or coldspots were correlated with peaks of G + C, we monitored base composition along all of the yeast chromosomes in 5- and 30-kb sliding windows, moving the windows in steps of 1 kb. As described below, there was a striking correlation between hotspots and peaks of high G + C obtained with the 5-kb window; the correlation was much less striking with the 30-kb window (data not shown). We found no obvious correlation between hotspot locations and peaks of gene density analyzed with a 30-kb window (24).

When scanned with a 5-kb window, chromosome III had eight G + C peaks in which the peak G + C content was at least 3% above the average (38.7%) for the chromosome. The midpoints of seven of the eight hotspots on chromosome III were within 2.5 kb of one of the eight peaks. This correlation is depicted in Fig. 5. In Fig. 3, the vertical blue lines drawn below each chromosome show peaks of G + C that are at least 3% above the average for the chromosome.

Several statistical tests were done to assess the significance of this correlation. A total of 221 peaks of G + C were observed. If the 177 hotspots were distributed at random with respect to the peaks, the expected number within 2.5 kb of a peak is 18. We observed 99 peak-associated hotspots, a deviation from this expectation ($P < 0.001$ by χ^2 analysis). We also calculated the G + C content in a 5-kb window centered at the midpoint of each hotspot. In 162 of 177 hotspots, the G + C content exceeded the average G + C content of the chromosome ($P < 0.001$ by χ^2 analysis). We also found an association between the ranking of the hotspots (Table 1) and their G + C content by linear

regression analysis ($P < 0.002$). In addition, the average rankings for the peak-associated hotspots and the nonpeak-associated hotspots were different ($P < 0.05$).

Of 22 mapped DSBs, 14 (13 intergenic and one intragenic) were associated with high G + C peaks (Table 1). The observed G + C content of these intergenic regions was 38.7%. The average G + C content of yeast genomic intergenic regions is about 34% (21). The average G + C content for the DNA sequences flanking the intergenic regions in a 5-kb window was 43.3%, higher than the 40.2% G + C observed for the average yeast ORF. Thus, the high G + C content associated with many of the hotspots reflects higher than average G + C content in both the intergenic region and flanking ORFs.

Relationship Between Hotspots and Known Repetitive Chromosomal Elements. We looked for associations between hotspots and repetitive DNA elements. One repetitive sequence is the ARS (autonomously replicating sequence) element; a subset of ARS elements are chromosomal replication origins. Active origins have been mapped only on chromosomes III (25) and VI (26, 27). There are 15 hotspots located on these two chromosomes that have a combined total length of 585 kb. A total of 75/585 or 0.128 of the sequences on these two chromosomes are within 2.5 kb of a hotspot. There are 17 active origins on the two chromosomes. If these are located randomly with respect to the hotspots, we expect to find 0.128×17 or two origins within 2.5 kb of a hotspot; we find one. Thus, there is no significant association between the location of replication origins and hotspots.

By similar arguments, we failed to find any association between hotspots and other repetitive DNA elements including the tRNA genes, small nuclear RNA genes, Ty elements (Ty1–5), and the long terminal repeat elements derived from the Ty elements (σ , δ , and τ). Our calculations were based on the numbers of repetitive DNA elements reported by Kim *et al.* (28) and the positions indicated in the Stanford Genome Database. In addition, no associations between repetitive elements and coldspots were found, except for the centromere/telomere associations described above.

Relationship Between Hotspots and Transcription. α Hotspots require the binding of transcription factors (8). Consequently, we looked for associations between hotspots and various parameters of gene expression. We mapped 20 DSBs within intergenic regions of the hotspots (Table 1). Intergenic regions can be located between the 5' ends of two genes (class 1), between the 3' ends of two genes (class 2), or between the 5' end of one gene and the 3' end of the other (class 3). The expected ratio of these classes is 1:1:2 if the transcriptional orientation of the flanking genes is random; this ratio is approximately that observed if all intergenic regions are analyzed (1,342, class 1; 1,427, class 2; 2,621, class 3). In contrast, of the 20 DSB-associated intergenic regions, we found 13 class 1, six class 3, and only one class 2. One explanation of this departure from expectation ($P = 0.0001$ by Fisher exact test) is that it reflects the requirement of transcription factor binding for hotspot activity. Alternatively, the negative superhelical stress resulting from divergent transcription may promote hotspot activity.

It is likely that some transcription factors are more proficient at stimulating recombination than others. For example, activity of the *HIS4* hotspot requires the binding of Bas1p, Bas2p, and Rap1p, but not the binding of Gcn4p (6). If other hotspots behave similarly, one might find conserved DNA sequences reflecting conserved transcription factor binding sites at hotspots. We compared putative transcription factor binding sites from our 20 mapped DSB sites with 31 random intergenic regions. There was a significant enrichment for Pho4p ($P = 0.05$) and Gcr1p ($P = 0.02$) binding sites. Because we examined the intergenic regions for the binding sites of 35 different factors,

however, the enrichment for Pho4p and/or Gcr1p binding sites at hotspots might represent a false-positive. Using the MEME program of GCG, we also looked for other conserved DNA sequence motifs specific to the 20 intergenic DSB sites. None were found. Blumenthal-Perry *et al.* (29) reported a motif associated with some recombination hotspots. We found no striking correlation between this motif and the hotspots identified in our study, suggesting that this motif might be associated with a subset of hotspots.

Several interesting correlations were observed between hotspot ORFs and gene functions. A total of 3,725 of the yeast ORFs have been assigned one or more functions (Munich Information Center for Protein Sequences database), and 181 of the hotspot ORFs are within this group. There was a very significant ($P < 0.0001$) over-representation of the hotspot ORFs in the metabolism functional class (81 of the 181 classified ORFs), particularly involving amino acid biosynthetic genes. Of the 25 hotspots with the highest rankings, 19 had one or more ORFs classified in the metabolism functional class. In addition, there was an over-representation ($P = 0.03$) of hotspot ORFs classified in the functional category of ionic homeostasis. There was an over-representation of coldspot ORFs in the categories of transport facilitation and intracellular transport ($P = 0.007$). One interpretation of these correlations is that certain categories of genes are associated with a particular chromatin structure that is favorable (hotspots) or unfavorable (coldspots) for initiating meiotic recombination.

Using a database that examined mRNA levels for most of the yeast genes in vegetative cells (30), we found that hotspot ORFs were expressed at higher levels than average ORFs. The average ORF has a median transcript level of 0.8/cell (30). Of the 250 hotspot ORFs in the database, 155 were expressed at higher levels than this value and 95 were expressed at lower values. This effect, at least in part, might be related to the correlation between hotspots and genes that function in metabolic pathways, because such genes are likely to be abundantly expressed in vegetative cells.

Microarrays have been used to examine transcription profiles for many conditions of cell growth including sporulation (31). There is an association between hotspot ORFs and genes whose expression is repressed in meiosis. This association is most significant at times 5 h and 11.5 h after induction of sporulation ($P = 0.001$ and $P = 0.003$, respectively, by contingency χ^2 test). This association is not a useful predictor of hotspot activity, however, because most hotspot ORFs show no meiosis-specific reduction in gene expression and most ORFs that show such a reduction are not hot ORFs.

Caveats Concerning Our Hotspot and Coldspot Mapping. Our mapping of hotspots and coldspots has several caveats. First, the method used to prepare P2 DNA enriches for any protein covalently bound to meiotic DNA (for example, other topoisomerases), in addition to Spo11p. It is unclear whether DSBs made by other proteins contribute to meiotic recombination. A related issue is that the difference in the level of hybridization between a hotspot ORF and a nonhotspot ORF in the microarray experiments (often less than a factor of 2) is smaller than expected by direct measurement of DSBs using standard Southern analysis (13). Although this difference may be an artifact of the PCR amplification required to prepare the probe, an alternative possibility is that the microarray analysis detects nonlocalized DSBs more sensitively than standard Southern analysis. Another caveat is that we restricted our analysis to DNA samples prepared from *rad50S* strains. Although the level of DSBs measured in *rad50S* strains often correlates well with the frequency of recombination measured by genetic techniques (3), a recent study has shown that DNA sequences near the telomeres of two chromosomes have lower levels of DSBs in *rad50S* strains

than in wild-type strains (V. Borde, A. Goldman, and M.L., unpublished data). Finally, our analysis was restricted to one specific genetic background. As discussed above, the same locus (for example, *HIS4* on chromosome III) can have very different levels of recombination in different genetic backgrounds.

Summary and Conclusions. No single distinctive feature is shared among all of the hotspots. The strongest correlation is the association with peaks of high G + C. One interpretation of this association is that DSBs are dictated, at least partly, by chromosome structure. Cohesins are bound at AT-rich regions located at approximately 15-kb intervals along chromosome III (19); some subset of these sites may be used to structure the DNA loops (averaging 20 kb) (32) of meiotic yeast chromosomes, and it has been suggested that DSB formation may be related to DNA loop structure (33). Because the average distance between hotspots measured in our study (54 kb) exceeds the average size of the DNA loops, all DNA loops do not represent hotspots. An alternative possibility is that the GC-rich regions bind protein(s) that result in a chromatin structure that is open to the recombination machinery. A related possibility is that the GC-rich regions are open to the recombination machinery by default, access to other chromosomal regions being blocked by cohesins. The association between hotspots and GC-rich regions is consistent with the finding that GC-rich *Escherichia coli* plasmid sequences often act as hotspots in yeast (20, 34).

One interpretation of the correlations of hotspots with genes of specific functions and specific expression patterns is that most hotspots require the binding of transcription factors to open the chromatin to permit access for the recombination machinery. We suggest that certain transcription factors are preferred by the

recombination machinery. The repression of transcription observed for genes located within some of the hotspots could reflect competitive interactions between the transcription and recombination machinery for access to the intergenic region.

We found, as had others previously, that many of the coldspots are nonrandomly associated with telomeres and centromeres. Borde *et al.* (35) showed that the lack of DSBs in cold regions does not reflect an absence of nuclease hypersensitive sites, but is likely to reflect some aspect of higher-order chromosome structure.

In conclusion, the position of hotspots may be regulated at two levels: the approximate position regulated by some feature of chromosome structure related to GC-richness, and a more precise position regulated by chromatin openness controlled by transcription factors. Although we (and others) have begun to define operational rules that may help in predicting hotspots, more mechanistic studies will be critical in refining these rules. In addition, studies in other organisms (such as *Schizosaccharomyces pombe*) eventually will reveal to what extent these rules are organism-specific.

We thank M. Dominska, L. Stefanovich, P. Greenwell, N. Perabo, and K. Werner for technical assistance and S. Whitfield for photographic assistance. We thank S. Keeney for hemagglutinin-tagged Spo11p and G. S. Roeder, C. Newlon, and S. Keeney for useful suggestions. We thank J. Sekelsky, Y. Blat, and Jing Zhu for assistance in data analysis. The research was supported by National Institutes of Health Grants GM24110 (T.D.P.) and HG00983 (P.O.B.) and by the Howard Hughes Medical Institute. P.O.B. is an Associate Investigator of the Howard Hughes Medical Institute. J.L.G. was supported by a training grant to the Lineberger Cancer Center (T32 CA09156) and by American Cancer Society Grant 5-39833; J.D.R. was supported by a Sandler grant.

- Lichten, M. & Goldman, A. S. (1995) *Annu. Rev. Genet.* **29**, 423–444.
- Petes, T. D., Malone, R. E. & Symington, L. S. (1991) in *The Molecular and Cellular Biology of the Yeast Saccharomyces*, eds. Broach, J., Jones, E. W. & Pringle, J. R. (Cold Spring Harbor Lab. Press, Plainview, NY), pp. 407–521.
- Wu, T. C. & Lichten, M. (1994) *Science* **263**, 515–518.
- Fan, Q.-Q., Xu, F. & Petes, T. D. (1995) *Mol. Cell. Biol.* **15**, 1679–1688.
- White, M. A., Wierdl, M., Detloff, P. & Petes, T. D. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 9755–9759.
- White, M. A., Dominska, M. & Petes, T. D. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 6621–6625.
- White, M. A., Detloff, P., Strand, M. & Petes, T. D. (1992) *Curr. Genet.* **21**, 109–116.
- Kirkpatrick, D. T., Fan, Q.-Q. & Petes, T. D. (1999) *Genetics* **152**, 101–115.
- Kirkpatrick, D. T., Wang, Y.-H., Dominska, M., Griffith, J. D. & Petes, T. D. (1999) *Mol. Cell. Biol.* **19**, 7661–7671.
- Bergerat, A., de Massy, B., Gadelle, D., Varoutas, P. C., Nicolas A. & Forterre, P. (1997) *Nature (London)* **386**, 414–417.
- Keeney, S., Giroux, C. N. & Kleckner, N. (1997) *Cell* **88**, 375–384.
- Lambie, E. J. & Roeder, G. S. (1988) *Cell* **52**, 863–873.
- Baudat, F. & Nicolas, A. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 5213–5218.
- Klein, S., Zenvirth, D., Dror, V., Barton, A. B., Kaback, D. B. & Simchen, G. (1996) *Chromosoma* **105**, 276–284.
- Zenvirth, D., Arbel, T., Sherman, A., Goldway, M., Klein, S. & Simchen, G. (1992) *EMBO J.* **11**, 3441–3447.
- Goyon, C. & Lichten, M. (1993) *Mol. Cell. Biol.* **13**, 373–382.
- DeRisi, J. L., Iyer, V. R. & Brown, P. O. (1997) *Science* **278**, 680–686.
- Strahl-Bolsinger, S., Hecht, A., Luo, K. & Grunstein, M. (1997) *Genes Dev.* **11**, 83–93.
- Blat, Y. & Kleckner, N. (1999) *Cell* **98**, 249–259.
- Wu, T. C. & Lichten, M. (1995) *Genetics* **140**, 55–66.
- Dujon, B. (1996) *Trends Genet.* **12**, 263–270.
- Kaback, D. B., Steensma, H. Y. & de Jonge, P. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 3694–3698.
- Sharp, P. M. & Lloyd, A. T. (1993) *Nucleic Acids Res.* **21**, 179–183.
- Goffeau, A., Aert, R., Agostini-Carbone, L., Ahmed, A., Aigle, M., Alberghina, L., Albermann, K., Albers, M., Aldea, M., Alexandraki, D. *et al.* (1997) *Nature (London)* **387**, 3–105.
- Newlon, C. S., Collins, I., Dershowitz, A., Deshpande, A. M., Greenfeder, S. A., Ong, L. Y. & Theis, J. F. (1993) *Cold Spring Harbor Symp. Quant. Biol.* **58**, 415–423.
- Yamashita, M., Hori, Y., Shinomiya, T., Obuse, C., Tsurimoto, T., Yoshikawa, H. & Shirahige, K. (1997) *Genes Cells* **2**, 655–665.
- Friedman, K. L., Brewer, B. J. & Fangman, W. L. (1997) *Genes Cells* **2**, 667–678.
- Kim, J. M., Vanguri, S., Boeke, J. D., Gabriel, A. & Voytas, D. F. (1998) *Genome Res.* **8**, 464–478.
- Blumenthal-Perry, A., Zenvirth, D., Klein, S., Onn, I. & Simchen, G. (2000) *EMBO J.*, in press.
- Holstege, F. C., Jennings, E. G., Wyrick, J. J., Lee, T. I., Hengartner, C. J., Green, M. R., Golub, T. R., Lander, E. S. & Young, R. A. (1998) *Cell* **95**, 717–728.
- Chu, S., DeRisi, J., Eisen, M., Mulholland, J., Botstein, D., Brown, P. O. & Herskowitz, I. (1998) *Science* **282**, 699–705.
- Moens, P. B. & Pearlman, R. E. (1988) *BioEssays* **9**, 151–153.
- Zickler, D. & Kleckner, N. (1999) *Annu. Rev. Genet.* **33**, 603–754.
- Stapleton, A. & Petes, T. D. (1991) *Genetics* **127**, 39–51.
- Borde, V., Wu, T. C. & Lichten, M. (1999) *Mol. Cell. Biol.* **19**, 4832–4842.