Note

Context Dependence of Meiotic Recombination Hotspots in Yeast: The Relationship Between Recombination Activity of a Reporter Construct and Base Composition

Thomas D. Petes[1] and Jason D. Merker

Department of Biology and Curriculum in Genetics and Molecular Biology, University of North Carolina, Chapel Hill, North Carolina 27599-3280

Manuscript received August 17, 2002
Accepted for publication September 3, 2002

ABSTRACT

Borde and colleagues reported that a reporter plasmid inserted at different genomic locations in *Saccharomyces cerevisiae* had different levels of meiotic recombination activity. We show that the level of recombination activity is very significantly correlated with the GC content of DNA sequences flanking the insertion.

In eukaryotes, chromosomal regions of high (hotspots) and low (coldspots) meiotic recombination activity have been identified (Lichten and Goldman 1995; Petes 2001). In *Saccharomyces cerevisiae*, the recombination activity of a chromosomal region is primarily a function of the frequency of local double-stranded DNA breaks (DSBs). DSBs generally occur between genes in nuclease-sensitive chromatin (Wu and Lichten 1994). Three different types of hotspots (not necessarily mutually exclusive) have been described. The activity of α-hotspots requires transcription factor binding, but not transcription *per se* (White *et al.* 1992, 1993). There are also hotspots associated with constitutively nucleosome-free regions of DNA (β; Kirkpatrick *et al.* 1999) and local regions of high G + C base composition (γ; Gerton *et al.* 2000). The mechanisms responsible for hotspot activity are not understood, although it has been suggested that all hotspots may be in regions with hypermodified nucleosomes (Petes 2001).

The relationship between naturally occurring hotspots in *S. cerevisiae* and chromosomal regions of high G + C base composition has been examined in several studies. Sharp and Lloyd (1993) found that chromosome III had two broad (~50 kb) regions of high G + C content and pointed out that three of the four known recombination hotspots were located in the GC-rich regions. Baudat and Nicolas (1997) mapped all of the DSB sites on chromosome III, demonstrating that most of the strong DSB sites were located in these two regions. Gerton *et al.* (2001) used microarrays to map genomic hotspots to single open reading frame resolution and found many hotspots were associated with local regions of high GC content. One interpretation of these results is that GC-rich intergenic regions are a preferred substrate for the recombination machinery.

Borde *et al.* (1999) monitored recombination events within reporter plasmids (Figure 1) introduced into 10 different locations on yeast chromosome III. Two different assays of recombination activity were performed. First, Borde and colleagues measured the frequency of DSBs located at two positions within the inserted plasmid, indicated as DSB-left and DSB-right in Figure 1. Second, they determined the frequency of heteroallelic recombination between the two plasmids (pMJ113 and pMJ115) placed at allelic positions. In general, these two measurements gave the same relative strength of recombination activity. Borde *et al.* (1999) found that recombination activity, as measured by either of these assays, varied by more than a factor of 10, depending on the position of the insertion.

To determine whether the recombination activity of the plasmid insertions in the study of Borde *et al.* (1999) was correlated with the G + C content of DNA sequences flanking the insertion, we used the "Composition" program of GCG to analyze the G + C content in various "windows" of flanking DNA for all 10 insertions, varying the size of the window from 500 bp (250 bp to each side of the insertion) to 100 kb (50 kb to each side of the insertion). For each window, the correlation between G + C content and the recombination activity (as measured by the sum of the DSBs at DSB-left and DSB-right) was determined using both the Pearson parametric test and the Spearman nonparametric test. For almost all windows, the *P* values of the correlation were highly

[1]*Corresponding author:* Department of Biology, University of North Carolina, Chapel Hill, NC 27599-3280.
E-mail: tompetes@email.unc.edu

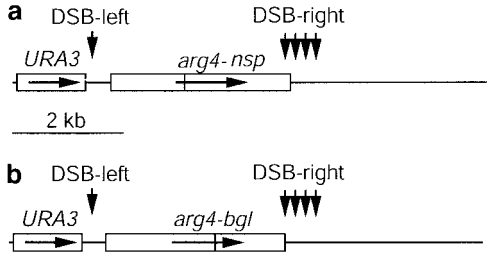

FIGURE 1.—Plasmids used to monitor recombination activity. The 8.5-kb plasmids (a) pMJ113 and (b) pMJ115 contain pBR322 sequences and the *URA3* and *ARG4* genes (BORDE *et al.* 1999). Plasmids pMJ113 and pMJ115 contain the mutant alleles *arg4-nsp* and *arg4-Bgl*, respectively. Genomic DNA fragments derived from yeast chromosome III were inserted into these plasmids, and the resulting plasmid derivatives were integrated into the chromosomes of haploid *a* and α-strains; these haploid strains were subsequently crossed to generate diploids. Meiotic recombination was measured in two ways (BORDE *et al.* 1999). In *RAD50* diploids, the frequency of Arg^+ spores, resulting from heteroallelic exchange between *arg4-nsp* and *arg4-Bgl*, was measured. In *rad50S* diploids, Southern analysis was used to measure the frequency of DSB formation at the positions indicated by arrows.

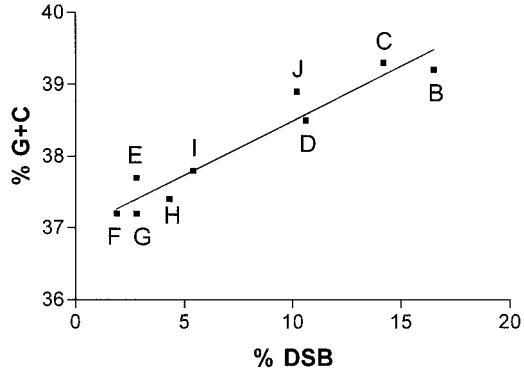

FIGURE 2.—The correlation between recombination activity in the reporter constructs (measured by the frequency of DSB formation) and the GC content in a 100-kb window. BORDE *et al.* (1999) measured recombination activity of the reporter plasmids inserted at 10 positions on chromosome III. One measurement of this activity was the sum of the DSB frequency for both DSB-left and DSB-right. In this graph, this parameter is shown as a function of the GC content of the flanking genomic sequences, measured 50 kb to each side of the insertion (100-kb window); since one of the plasmid insertions (at *CHA1*) was located <50 kb from the telomere, this insertion was not included in the analysis. The insert locations, the "% DSB" and the "% G + C" content for the nine other insertions, are: *HIS4* (B, 16.5% DSB, 39.2% GC); *LEU2* (C, 14.2% DSB, 39.3% GC); *YCL011c* (D, 10.6% DSB, 38.5% GC); *YCR004c* (E, 2.8% DSB, 37.7% GC); *RVS161* (F, 1.9% DSB, 37.2% GC); *YCR017c* (G, 2.8% DSB, 37.2% GC); *YCR026c* (H, 4.3% DSB, 37.4% GC); *RIM1* (I, 5.4% DSB, 37.8% GC); *MAT* (J, 10.2% DSB, 38.6% GC).

significant (Table 1). The correlation between DSB activity and the flanking G + C content (100-kb window) for nine insertions is shown in Figure 2. The positions of the insertions in relation to base composition for windows of 5 and 100 kb are shown in Figure 3.

TABLE 1

Correlation coefficients (*r* values) and the associated probabilities of correlations (*P* values) for the relationship of G + C content (measured in windows from 0.5 to 100 kb) and the frequency of meiotic recombination (assayed by double-strand break formation)

Window size (kb)	Pearson test		Spearman test	
	r values	*P* values	*r* values	*P* values
0.5	0.808	0.005	0.857	0.002
1	0.639	0.05	0.657	0.04
2	0.774	0.009	0.809	0.005
2.5	0.760	0.01	0.809	0.005
5	0.750	0.01	0.681	0.03
10	0.778	0.008	0.686	0.03
20	0.709	0.02	0.511	0.13
30	0.930	<0.0001	0.726	0.02
40	0.894	0.001	0.789	0.01
50	0.904	0.0008	0.765	0.02
60	0.888	0.001	0.702	0.04
75	0.934	0.0002	0.740	0.02
100	0.961	<0.0001	0.929	0.0003

For each window, the correlation between G + C content and the recombination activity (as measured by the sum of the DSBs at DSB-left and DSB-right) was determined using the SAS JMP software (version 4.0.4; SAS Institute, Cary, NC); we determined correlation coefficients (*r* values) and their associated probabilities (*P* values) using both the Pearson parametric test and the Spearman nonparametric test. Since one of the insertions (near the *CHA1* gene) is ~15 kb from the telomere, this insertion was not used to calculate *r* and *P* values for windows >30 kb.

In Table 1, the *P* values for correlations determined with the Pearson test (which assumes a linear correlation) fall into two groups. For windows between 0.5 and 20 kb, the *P* values are between 0.005 and 0.02; for windows between 30 and 100 kb, the *P* values are all substantially lower, between <0.0001 and 0.001. This analysis indicates that the recombination activity of inserted sequences is better predicted by the GC content of large chromosomal regions (≥30 kb) than by that of smaller regions (≤20 kb). The *P* values for the nonparametric Spearman test did not follow the same simple pattern. Since the correlation coefficients (*r* values) obtained with the Pearson test for the large chromosomal regions are larger than those obtained with the Spearman test for any window (except for the 100-kb window), it is likely that the Pearson test represents the better method of analyzing the data.

From our analysis, we suggest two conclusions. First, the context dependence of the recombination properties of plasmid insertions observed by BORDE *et al.* (1999) reflects, at least in part, the base composition of the DNA sequences flanking the insertion. These effects act at a distance, since the DSB sites used to assay recombination are located >1 kb from the junction with the chromosomal DNA. Second, our analysis supports the

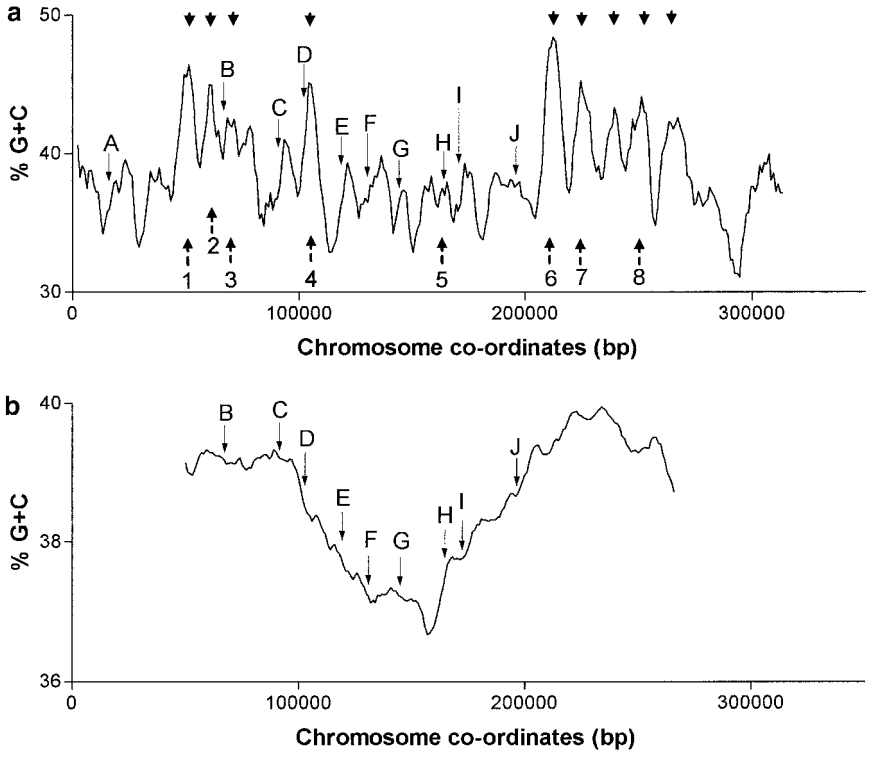

FIGURE 3.—Base composition of chromosome III assayed using moving windows of 5 and 100 kb. (a) 5-kb moving window: GC content was determined (using the "Window" Program of GCG) in a window of 5 kb moved in 1-kb intervals. The arrows with letters show the positions of the insertions (A represents the *CHA1* insertion; code for B–J in Figure 2). The short arrows at the top of the figure indicate nine peaks of G + C that are >3% above the average for chromosome III (38.7%). The numbered arrows represent hotspots for recombination as determined by GERTON *et al.* (2000). (b) 100-kb moving window: GC content was determined in windows of 100 kb moved in 1-kb intervals.

earlier conclusion (SHARP and LLOYD 1993; BAUDAT and NICOLAS 1997) that there are recombination-stimulating effects of high G + C that involve large (≥30 kb) chromosomal regions. On the basis of the colocalization of hotspots and peaks of local (≤5 kb) G + C content (GERTON *et al.* 2000; BIRDSELL 2002), there appear to be both local and regional effects of G + C content on meiotic recombination. Local effects could reflect a preference of the recombination machinery to associate with high-GC intergenic regions whereas regional effects may represent some global feature of chromosome structure that is affected by base composition (ZICKLER and KLECKNER 1999); such global features could be regions of recombination-promoting hyperacetylated nucleosomes or recombination-suppressing hypermethylated nucleosomes. The reporter plasmids were GC rich (46.2%) compared to yeast genomic DNA (39%) and therefore would be expected to contribute to both local and regional GC richness. Since this contribution is the same for each position of integration, the relative recombination activities of the insertions should be independent of this effect.

A correlation between high recombination activity and high GC content has also been observed for humans (EISENBARTH *et al.* 2001), *Caenorhabditis elegans*, and *Drosophila melanogaster* (MARAIS *et al.* 2001). Two different types of explanation have been given for this relationship. We suggested that regions of high GC stimulate recombination (GERTON *et al.* 2000; PETES 2001). Alternatively, high levels of recombination may create genomic regions with high GC content (EYRE-WALKER 1993; GALTIER *et al.* 2001; MARAIS *et al.* 2001), since high recombination rates result in elevated levels of heteroduplex formation and since mismatches in heteroduplexes are usually repaired with a bias toward GC over AT in mitotic mammalian and yeast cells (BROWN and JIRICNY 1989; BIRDSELL 2002) and in meiotic yeast cells (BIRDSELL 2002). The first model is more consistent with the observation that GC-rich sequences derived from *Escherichia coli* plasmids often have unusually strong hotspot activity when introduced into the yeast genome (STAPLETON and PETES 1991; WU and LICHTEN 1995), although (as discussed above) the hotspot activity of these sequences is influenced by the GC content of the chromosomal DNA sequences that flank the insertion.

Although our results and those of others support the conclusion that GC content affects hotspot activity, local and regional GC content is only one factor. In *S. cerevisiae*, other factors associated with hotspot activity include: (1) "open" chromatin, (2) a location between divergently transcribed genes, and (3) a requirement for local transcription factor binding (LICHTEN and GOLDMAN 1995; PETES 2001); hotspots are found infrequently in centromeric or telomeric regions (BAUDAT and NICOLAS 1997; GERTON *et al.* 2000). Although these associations have some predictive value in determining which DNA sequences will be recombination hotspots, the mechanisms responsible for these associations are not understood.

We thank P. Mieczkowski, J. Stone, and M. Lichten for useful discussions and J. Stone, M. A. Amamoo, D. Moore, and N. Degtyareva

for help with data analysis. The research was supported by National Institutes of Health grant GM-24110.

LITERATURE CITED

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

Birdsell, J. A., 2002 Integrating genomics, bioinformatics, and classical genetics to study the effects of recombination on genome evolution. Mol. Biol. Evol. **19:** 1181–1197.

Borde, V., T.-C. Wu and M. Lichten, 1999 Use of a recombination reporter insert to define meiotic recombination domains on chromosome III of *Saccharomyces cerevisiae*. Mol. Cell. Biol. **19:** 4832–4842.

Brown, T. C., and J. Jiricny, 1989 Repair of base-base mismatches in simian and human cells. Genome **31:** 705–711.

Eisenbarth, I., A. M. Striebel, E. Moschgath, W. Vogel and G. Assum, 2001 Long-range sequence composition mirrors linkage disequilibrium pattern in a 1.13 Mb region of human chromosome 22. Hum. Mol. Genet. **10:** 2833–2839.

Eyre-Walker, A., 1993 Recombination and mammalian genome evolution. Proc. R. Soc. Lond. Ser. B Biol. Sci. **252:** 237–243.

Galtier, N., G. Piganeau, D. Mouchiroud and L. Duret, 2001 GC-content evolution in mammalian genomes: the biased gene conversion hypothesis. Genetics **159:** 907–911.

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

Kirkpatrick, D. T., Y.-H. Wang, M. Dominska, J. D. Griffith and T. D. Petes, 1999 Control of meiotic recombination and gene expression in yeast by a simple repetitive DNA sequence that excludes nucleosomes. Mol. Cell. Biol. **19:** 7661–7671.

Lichten, M., and A. S. H. Goldman, 1995 Meiotic recombination hotspots. Annu. Rev. Genet. **29:** 423–444.

Marais, G., D. Mouchiroud and L. Duret, 2001 Does recombination improve selection on codon usage? Lessons from nematode and fly complete genomes. Proc. Natl. Acad. Sci. USA **98:** 5688–5692.

Petes, T. D., 2001 Meiotic recombination hot spots and cold spots. Nat. Rev. Genet. **2:** 360–369.

Sharp, P. M., and A. T. Lloyd, 1993 Regional base composition variation along yeast chromosome III: evolution of chromosome primary structure. Nucleic Acids Res. **21:** 179–183.

Stapleton, A., and T. D. Petes, 1991 The Tn3 β-lactamase gene acts as a hotspot for meiotic recombination in yeast. Genetics **127:** 39–51.

White, M. A., P. Detloff, M. Strand and T. D. Petes, 1992 A promoter deletion reduces the rate of mitotic, but not meiotic, recombination at the *HIS4* locus in yeast. Curr. Genet. **21:** 109–116.

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.

Wu, T.-C., and M. Lichten, 1994 Meiosis-induced double-strand break sites determined by yeast chromatin structure. Science **263:** 515–518.

Wu, T.-C., and M. Lichten, 1995 Factors that affect the location and frequency of meiosis-induced double-strand breaks in *Saccharomyces cerevisiae*. Genetics **140:** 55–66.

Zickler, D., and N. Kleckner, 1999 Meiotic chromosomes: integrating structure and function. Annu. Rev. Genet. **33:** 603–754.

Communicating editor: A. Nicolas