GC-Biased Gene Conversion in Yeast Is Specifically Associated with Crossovers: Molecular Mechanisms and Evolutionary Significance

Yann Lesecque,[1] Dominique Mouchiroud,[1] and Laurent Duret*,[1]

[1]Laboratoire de Biométrie et Biologie Evolutive, UMR CNRS 5558, Université de Lyon, Université Lyon 1, Villeurbanne, France

***Corresponding author:** E-mail: laurent.duret@univ-lyon1.fr.

Associate editor: Kenneth Wolfe

Abstract

GC-biased gene conversion (gBGC) is a process associated with recombination that favors the transmission of GC alleles over AT alleles during meiosis. gBGC plays a major role in genome evolution in many eukaryotes. However, the molecular mechanisms of gBGC are still unknown. Different steps of the recombination process could potentially cause gBGC: the formation of double-strand breaks (DSBs), the invasion of the homologous or sister chromatid, and the repair of mismatches in heteroduplexes. To investigate these models, we analyzed a genome-wide data set of crossovers (COs) and noncrossovers (NCOs) in *Saccharomyces cerevisiae*. We demonstrate that the overtransmission of GC alleles is specific to COs and that it occurs among conversion tracts in which all alleles are converted from the same donor haplotype. Thus, gBGC results from a process that leads to long-patch repair. We show that gBGC is associated with longer tracts and that it is driven by the nature (GC or AT) of the alleles located at the extremities of the tract. These observations invalidate the hypotheses that gBGC is due to the base excision repair machinery or to a bias in DSB formation and suggest that in *S. cerevisiae*, gBGC is caused by the mismatch repair (MMR) system. We propose that the presence of nicks on both DNA strands during CO resolution could be the cause of the bias in MMR activity. Our observations are consistent with the hypothesis that gBGC is a nonadaptive consequence of a selective pressure to limit the mutation rate in mitotic cells.

Key words: **recombination, biased gene conversion, crossover, meiotic drive.**

Introduction

In many eukaryotes, recombination is required for the proper segregation of chromosomes during meiosis. This process involves the programmed formation of double-strand breaks (DSBs), which are subsequently repaired by using homologous sequences as a template. It is generally accepted that the profound *raison d'être* of meiosis is to enhance the efficacy of natural selection by allowing the formation of new combinations of alleles via this process of recombination. Thus, asexual taxa (which cannot create new haplotypes by recombination) are expected to be evolutionary dead ends, because of their reduced potential for adaptation (for review, see Coop and Przeworski [2007]).

Recently, many studies have shown that besides its fundamental impact on selection efficacy, recombination also strongly contributes to genome evolution via the nonadaptive process of biased gene conversion (BGC) (for review, see Duret and Galtier [2009] and Webster and Hurst [2012]). Gene conversion is a process intrinsically associated with recombination that results in the nonreciprocal transfer of genetic information between the two recombining sequences. This process is said to be biased if one of the two alleles has a higher probability to be the donor than its homolog. BGC tends to raise the frequency of the donor allele in the pool of gametes and therefore leads to increase its probability of fixation in the population. It is a nonadaptive process, because the spread of one allele through BGC is independent of its effect on fitness. However, its impact on the dynamics of allele frequency within populations is very similar to that of directional selection (Nagylaki 1983). Different lines of evidence indicate that in many eukaryotes, BGC tends to favor the transmission of GC alleles in AT/GC heterozygotes (for review, see Duret and Galtier [2009] and Webster and Hurst [2012]). In mammals, it has been shown that gBGC (i.e., GC-favoring BGC) is the main determinant of the evolution of genomic base composition (Meunier and Duret 2004; Duret and Arndt 2008; Katzman et al. 2011; Auton et al. 2012), and there is indirect evidence that this process is widespread in eukaryotes (Capra and Pollard 2011; Escobar et al. 2011; Pessia et al. 2012). Moreover, it has been shown that gBGC can interfere with natural selection and lead to the fixation of deleterious alleles (Galtier and Duret 2007; Berglund et al. 2009; Galtier et al. 2009; Glémin 2010, 2011; Ratnakumar et al. 2010; Necşulea et al. 2011). However, despite its major impact on genome evolution, the molecular mechanisms leading to gBGC are still unknown.

Much of our knowledge of the molecular mechanisms of meiotic recombination in eukaryotes has come from the study of yeasts (for review, see de Massy [2003]). Recombination is initiated by the formation of DSBs followed by 5′- to 3′-end resection (Smith and Nicolas 1998; Krogh and Symington 2004). DSBs are then repaired, using homologous sequences as a template, either from the sister chromatid or, more frequently, from the nonsister chromatid (the

Article

homolog). Recombination events between homologs can lead to the exchange of flanking regions (i.e., crossovers [COs]) or not (i.e., noncrossover [NCO] recombination events). The two types of events result from different recombination pathways (fig. 1). In budding yeast (*Saccharomyces cerevisiae*), NCOs result principally from the synthesis-dependant strand annealing pathway, and secondarily from double Holliday junction (dHj) dissolution, whereas COs result from dHj resolution (class I COs) and from the Mus81 pathway (class II COs) (McMahill et al. 2007; Martini et al. 2011). In all cases, the repair of DSBs by the homolog involves the formation of heteroduplex DNA, with one DNA strand coming from the broken chromosome and the other from the intact template. When homologs are not identical, mismatches are formed in this heteroduplex, and their repair leads to the conversion of one allele by the other. The segment of the chromosome affected by a conversion event is called the conversion tract. Mancera et al. (2008) recently published a high-resolution recombination map that allowed a very detailed genome-wide analysis of conversion tracts in *S. cerevisiae*. The median length of conversion tracts is 2 kb for COs and 1.8 kb for NCOs. They found that the majority of conversion tracts (89% for COs and 97% for NCOs) are "simple," that is, with one single-donor haplotype along the whole tract (Mancera et al. 2008). They notably demonstrated that conversion events overlapping AT/GC heterozygous sites lead to a significant overtransmission of the GC allele (1.3% greater than expected under the null hypothesis of Mendelian transmission), thus providing the first direct evidence of gBGC in a eukaryote (Mancera et al. 2008).

Several hypotheses can be proposed concerning the molecular mechanisms responsible for gBGC. First, the analysis of gene conversion tracts, in yeasts or in mammals, indicates that in most cases, gene conversion occurs from the intact chromosome toward the broken one (Nicolas et al. 1989; Mancera et al. 2008; Webb et al. 2008). Thus, if in an AT/GC heterozygote, DSBs occur more frequently on the AT-richer haplotype, this could lead to the overtransmission of the GC allele. This model is hereafter referred to as the "initiation bias" hypothesis. An alternative model is that gBGC could result from the activity of the mismatch repair (MMR) machinery. MMR plays a major role during recombination, not only for the repair of mismatches in heteroduplex DNA but also for the choice of the DNA template to be used to repair the DSB. Indeed, during the process of invasion of the homologous chromosome by the single-stranded 3′-overhang, MMR is able to sense the mismatches present in the heteroduplex and to reject the invading strand if the level of sequence divergence is too high (Hunter et al. 1996; Chen and Jinks-Robertson 1999). This activity is crucial to avoid recombination between nonallelic loci (ectopic recombination) (Surtees et al. 2004). Current models suggest that in the cases where MMR prevents the invasion of the homolog, DSBs get subsequently repaired by using the sister chromatid (Martini et al. 2011), which leads to Mendelian transmission, without any conversion (fig. 1). In theory, it is possible that the decision to reject the invading strand or to repair the mismatch depends on the nature of the allele present on the invading strand: If strands carrying AT alleles were less prone to be rejected than those carrying GC alleles, then the

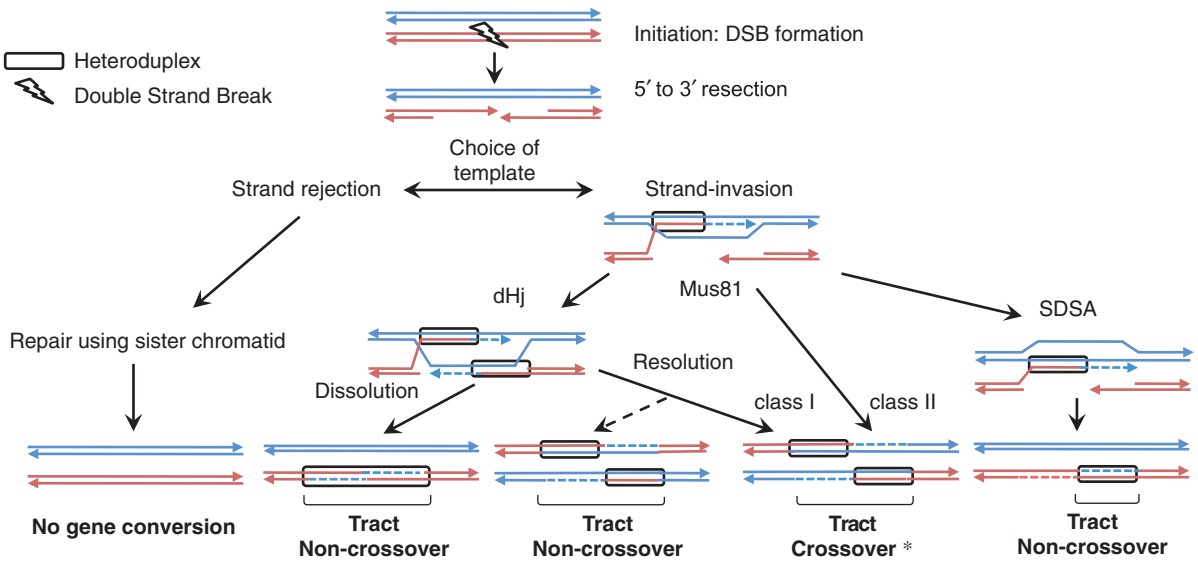

Fig. 1. Canonical model of meiotic recombination in *Saccharomyces cerevisiae*. For simplicity, only two homologous dsDNA molecules are represented (one red and one blue). Meiotic recombination is initiated by the formation of a DSB (here represented by a flash on the red haplotype), followed by 5′- to 3′-end resection. The DSB is subsequently repaired using as a template either the sister chromatid (not shown here; left part) or the homolog (here represented in blue; right part). There exist several DSB repair pathways, which, when the homolog is used as a template, can lead to COs or NCOs. Current models indicate that NCOs result principally from the synthesis-dependant strand annealing (SDSA) pathway and secondarily from double Holliday junction (dHj) dissolution, whereas COs result from dHj resolution (class I) and from the Mus81 pathway (class II) (Martini et al. 2011). The resolution of dHj into NCOs is represented by a dashed harrow, to indicate that this is a minor pathway. Dashed lines (blue and red) represent newly synthesized DNA and boxes show heteroduplex associations during the whole process. The * symbol next to the CO product refers to figure 4.

former would have more opportunities to get converted, which would lead to a conversion bias in favor of GC alleles. An alternative hypothesis is that MMR could cause gBGC via its activity in the repair of mismatches in heteroduplex DNA. The directionality of the repair by MMR depends on the presence of nicks flanking the mismatch (Jiricny 2006) and is not known to be biased toward the GC allele. It is, however, possible that a weak bias, such as the one causing gBGC in *S. cerevisiae*, might have remained unnoticed. Finally, we and others proposed that gBGC could be caused by the base excision repair (BER) machinery (Memisoglu and Samson 2000). Indeed, although MMR is the prominent repair system active during recombination (Evans and Alani 2000; Hoffmann and Borts 2004; Surtees et al. 2004; Jiricny 2006), there is evidence that other systems contribute to the repair of mismatches in heteroduplex DNA (Coïc et al. 2000). Given that BER is intrinsically biased toward GC, it is a priori expected that if this repair machinery is active on heteroduplex DNA during meiotic recombination, then it should induce gBGC (Brown and Jiricny 1989; Galtier et al. 2001; Birdsell 2002; Marais 2003). One clear difference between BER and MMR is the length of the region affected by the repair: Although MMR involves DNA resynthesis over hundreds of base pairs (i.e., about the size of conversion tracts) (Holmes and Clark 1990; Thomas et al. 1991), BER leads only to short-patch repair (1–13 bp) (Memisoglu and Samson 2000). Given the length of gene conversion tracts (~2 kb on average), if some single-nucleotide polymorphism (SNP) conversion events are driven by BER, then the conversion of these SNPs should occur independently of the conversion of neighboring SNPs. Thus, although MMR is expected to produce predominantly simple conversion tracts, BER—if active during recombination—is expected to lead frequently to complex conversion tracts (i.e., tracts involving conversion events from both parental haplotypes). Hence, if BER is responsible for the conversion bias, then gBGC should be much stronger among complex conversion tracts compared with simple conversion tracts.

To try to distinguish between the different processes possibly responsible for gBGC (initiation bias, MMR, or BER), we decided to analyze the high-resolution recombination data published by Mancera et al. (2008). We demonstrate that in *S. cerevisiae*, gBGC is associated with long-patch DNA repair and is specific of CO events. We further show that gBGC is associated with longer conversion tracts and that the conversion bias depends on the nature of mismatches at the boundaries of the tract. These observations are not consistent with the initiation bias and BER models and suggest that gBGC is caused by MMR.

Results

To analyze gene conversion tracts in yeast, we used the high-resolution recombination data published by Mancera et al. (2008). These data were obtained by genotyping tetrads resulting from 46 meioses, in a diploid hybrid of two wild-type *S. cerevisiae* strains (S96 and YJM789). Several other similar data sets have been published (Winzeler et al. 1998; Chen et al. 2008; Qi et al. 2009). However, the Mancera data set is currently the only one to provide exhaustive genotyping data (i.e., almost all the sites that differ between the two strains have been genotyped) for such a large number of meioses. The median distance between two consecutive markers is 78 bp. We analyzed all recombination events associated with detectable conversion tracts (2,884 COs and 2,090 NCOs). On average, conversion tracts overlap nine SNPs. Each of these SNP sites was genotyped in the two resulting spores. Thus, in total, 89,538 SNP sites involved in a conversion event have been genotyped. To test whether gene conversion shows a bias in favor of GC or AT allele, we focused on the subset of sites that correspond to AT/GC heterozygotes in the parental hybrid (87% of the total set of SNPs involved in conversion events). For this set of sites, we counted the proportion of GC alleles in the offspring (x). The existence of a conversion bias was tested by comparing x to the Mendelian expectation (50%), with a one-sample proportion test (see Materials and Methods). The intensity of the conversion bias in favor of GC alleles was measured by the coefficient $b = 2x - 1$. (NB: We chose this expression because it is equivalent to the definition of the selection coefficient of a semidominant mutation, see Nagylaki [1983].) In agreement with previous results (Mancera et al. 2008), we observed a significant conversion bias toward GC alleles ($b = 0.013$, $P < 10^{-3}$; table 1; NB: The properties of the conversion tracts that we studied are summarized in supplementary table S1, Supplementary Material online).

Table 1. Conversion Bias Toward GC Bases for AT/GC SNPs Involved in a Recombination Event.

Conversion Tract Type	Number of Genotyped SNP Sites with AT/GC Polymorphism	Conversion Bias Toward GC Bases (b)[a]	P[a]
All	77,901	0.013	<0.001
Simple	64,898	0.014	<0.001
Complex	13,003	0.008	0.36 (NS)

NOTE.—NS, nonsignificant.
[a]One-sample proportion test.

Transmission Biases in Simple and Complex Conversion Tracts

If BER is the unique cause of gBGC, it is expected that the conversion bias in favor of GC alleles should be much stronger among complex conversion tracts than among simple tracts. To test this prediction, we measured the conversion bias in favor of GC alleles separately for SNPs located in simple and complex conversion tracts. Interestingly, we observed that the conversion bias is not reduced among simple conversion tracts compared with complex ones (table 1). On the contrary, b tends to be higher for SNPs located in simple conversion tracts (although the difference is not significant; two-sample proportion test). It should be noted that complex conversion tracts tend to be longer than simple ones (because, by definition, complex tracts must contain at least two SNPs, whereas simple tracts may contain just one SNP).

To test whether this ascertainment bias might have affected our conclusions, we repeated the analysis on SNP sites located in tracts overlapping at least five SNPs. The results remained unchanged (supplementary table S2, Supplementary Material online). Note that a large majority (83%) of SNPs involved in a recombination event are located in simple conversion tracts. Hence, quantitatively, the conversion bias in favor of GC alleles is essentially due to recombination events associated with simple conversion tracts. This observation is, therefore, not consistent with the predictions of the BER model.

Conversion Biases Operate on Multiple Adjacent SNPs

In the above analysis, as in Mancera et al. (2008), the statistical significance of the conversion bias in favor of GC alleles was assessed under the assumption that each SNP conversion was an independent event. However, given that the observed conversion bias is essentially associated with simple conversion tracts, this assumption is clearly incorrect: All SNPs in a simple tract are converted together from the same donor haplotype. This nonindependence might lead to overestimate the statistical significance of conversion biases. To avoid this potential artifact, we reanalyzed conversion biases at the scale of the conversion event (i.e., a set of SNPs involved in a common conversion tract), focusing exclusively on simple conversion tracts ($N = 4{,}428$ recombination events). For each tract, we measured the difference in GC content between the two haplotypes involved in the conversion event (ΔGC, supplementary fig. S1, Supplementary Material online). We selected all cases where one of the two haplotypes had a higher GC content than the other (i.e., $\Delta GC \neq 0$, $N = 3{,}676$ recombination events). These conversion tracts were said to have a "AT/GC-richer" polymorphism. Among the 7,352 corresponding haplotypes in the pool of spores, we observed a clear and statistically significant conversion bias in favor of the GC-richer haplotype (fig. 2; $b = 0.030$, $P = 0.01$), which confirms the existence of gBGC. Note that this conclusion remains when using a more stringent threshold to categorize AT/GC-richer haplotypes (supplementary text S2 and fig. S2, Supplementary Material online). In the rest of this article, to avoid any statistical artifact due to the nonindependence of SNPs located in a same tract, we analyzed conversion biases at the scale of conversion tracts (and not individual SNPs), excluding complex conversion tracts.

gBGC Is CO Specific

Among CO recombination events, we observed a strong conversion bias in favor of the GC-richer haplotype (fig. 2; $b = 0.057$, $P = 3.6 \times 10^{-4}$). Interestingly, NCOs did not exhibit any conversion bias. This difference between COs and NCOs conversion biases was significant (fig. 2; $P = 0.014$). This indicates that gBGC observed in the whole data set is essentially driven by COs. This observation is not consistent with the initiation bias model, which predicts that gBGC should affect both COs and NCOs (see Discussion).

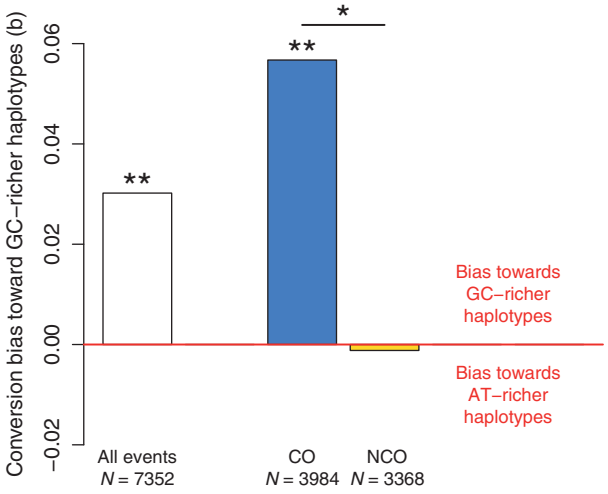

FIG. 2. Conversion bias toward GC-richer haplotypes. The conversion bias toward GC-richer haplotypes (b) was computed for simple conversion tracts, taken all together (white bar) or separating tracts associated with COs (blue bar) and NCOs (yellow bar). "N" is the number of genotyped haplotypes in each category. The red horizontal line indicates the Mendelian expectation ($b = 0$). Significant conversion biases are indicated by "*" for a P value ≤ 0.05 and "**" for a P value ≤ 0.01 (one-sample proportion test). The "*" between "CO" and "NCO" bars denotes the fact that the conversion bias toward GC-richer haplotypes is significantly different between CO and NCO events (two-sample proportion test).

gBGC Is Driven by Mismatches Located at the Extremities of Conversion Tracts and Is Associated with Longer Tracts

The previous observations are inconsistent with the BER and initiation bias models. We, therefore, investigated further the hypothesis of a mismatch repair bias driven by MMR. The fact that gBGC is observed in simple conversion tracts is compatible with a role of MMR in gBGC. However, this hypothesis raises the question of how the MMR machinery would be able to distinguish AT-richer versus GC-richer haplotypes. It seems a priori unlikely that the MMR machinery could sense the global difference in GC content along the region, typically 2-kb long, affected by the conversion. Given that the directionality of the repair by MMR depends on the presence of flanking nicks (Jiricny 2006), we hypothesized that the bias could depend specifically on the nature of the mismatches found at the boundaries of the conversion tract, that is, those that are closest to the nicks flanking the heteroduplexes (fig. 4).

To test this prediction, we classified conversion tracts according to the nature of the first and the last SNPs of the tract. When one particular strain had a G or a C for first and last SNPs in the region corresponding to the conversion tract, whereas the other strain had a A or T at those positions, the first haplotype was called "GC_f" (which stands for GC-flanked haplotype) and the second "AT_f" (AT-flanked haplotype) (supplementary fig. S1, Supplementary Material online). These conversion tracts were said to have a "GC_f/AT_f polymorphism." Similarly, conversion tracts with a GC/AT SNP at

one end and an AT/TA or GC/CG SNP at the other end were classified as "one-side GC_f/AT_f polymorphisms." All other cases were excluded: When haplotypes are flanked by G or C at one extremity and A or T at the other one, it is impossible to define a conversion bias in this fashion because the two parental haplotypes are indistinguishable in term of GC/AT flanking SNPs.

Among CO-associated simple conversion events with GC_f/AT_f polymorphism ($N = 1{,}104$ events, i.e., 38% of the set of CO-associated simple conversion tracts), we observed a strong conversion bias toward the GC_f haplotype (fig. 3). For CO-associated simple conversion events with one-side GC_f/AT_f polymorphism, the conversion bias toward the GC_f haplotype was slightly weaker and only marginally significant

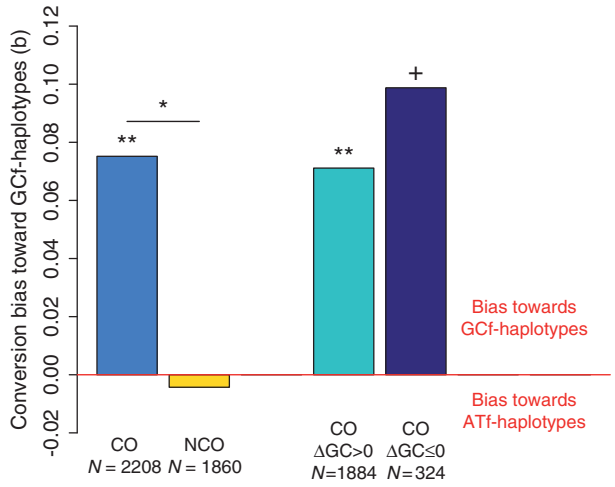

FIG. 3. Conversion bias toward GCf haplotypes. The conversion bias toward GCf haplotypes (b) was computed for simple conversion tracts, associated with COs (sky blue bar), NCOs (yellow bar), COs with $\Delta GC > 0$ (light blue bar), and COs with $\Delta GC \leq 0$ (dark blue bar). ΔGC is positive ($\Delta GC > 0$) when the GCf haplotype is globally richer in G + C than the ATf haplotype, it is negative or null otherwise ($\Delta GC \leq 0$). The red horizontal line indicates the Mendelian expectation ($b = 0$). Significant conversion biases are indicated by "**" for two-tailed one-sample proportion test with a P value < 0.01 and "+" for one-tailed one-sample proportion test with alternative hypothesis "$b > 0$" and P value < 0.05. The "*" between "CO" and "NCO" bars denotes the fact that the conversion bias toward GCf haplotypes is significantly different between CO and NCO events (two-sample proportion test, P value < 0.05).

($b = 0.06$, $P = 0.07$, one-sample proportion test), probably because of limited sample size ($N = 435$ events). Note that for NCO recombination events, the conversion of GC_f/AT_f haplotypes was unbiased (fig. 3), Thus, as noticed previously, the conversion bias appears to be CO specific. Interestingly, we noticed that for CO-associated simple conversion tracts, the length of tracts varies according to the direction of conversion: The median tract length (computed as the distance between the two most distal SNPs within the tract, for all conversion tracts overlapping at least two SNPs) is 1,322 bp for GC_f conversion tracts, compared with 1,146 bp for other conversion tracts (Wilcoxon test, $P = 0.0017$). This difference is not observed for NCO recombination events (median tract length: 1,046 bp for GC_f conversion tracts, compared with 1,044 bp for other conversion tracts).

The fact that we observed a conversion bias toward the GC_f haplotype is consistent with the hypothesis that the bias depends on the nature of mismatches located at the extremities of the conversion tracts. However, given the way they are defined, GC_f haplotypes also tend to be GC rich. Thus, the observed conversion bias toward the GC_f haplotype might in fact be driven by a conversion bias toward the GC-richer haplotype (i.e., it might depend on the GC richness of the whole haplotype and not specifically on the SNPs located at the extremities). To test this hypothesis, we considered the subset of CO-associated simple conversion tracts with GC_f/AT_f polymorphism for which the GC_f haplotype is not richer in GC than the AT_f haplotype ($\Delta GC \leq 0$ in fig. 3). If the bias toward GC_f haplotypes was driven by the bias toward GC-richer haplotypes, one would expect the GC_f-conversion bias to be negative for these 162 events. In contradiction with this prediction, we observed a strong and positive bias in favor of GC_f haplotypes ($b = 0.099$, fig. 3). This indicates that the conversion bias toward GC_f haplotypes exists regardless of the difference in GC content between homologous haplotypes and that this conversion bias is predominant over the conversion bias toward the GC-richer haplotype. And indeed, when we categorized conversion tracts into AT/GC-richer haplotypes based on internal SNPs (i.e., ignoring the two SNPs at the extremities of the tract), then the conversion bias in favor of the GC-richer haplotype becomes much weaker and nonsignificant (table 2, supplementary table S3, Supplementary Material online). Thus, gBGC in yeast is essentially driven by a conversion bias in favor of GC_f haplotypes. Given that in 85% of the cases, GC_f haplotypes are GC richer

Table 2. Conversion Bias Toward GC-Richer Haplotypes among AT_f/GC_f Polymorphism, Considering All SNPs, or Only Flanking or Internal SNPs.

SNPs Considered to Classify Haplotypes as AT- or GC Richer	Number of Genotyped Haplotypes with AT/GC-Richer Polymorphism[a]	Conversion Bias Toward GC-Richer Haplotypes (b)[b]	P[b]
All	1,114	0.070	0.02
Flanking SNPs only	1,246	0.101	<0.001
Flanking SNPs excluded	1,072	0.034	0.28 (NS)

NOTE.—NS, nonsignificant.

[a]Haplotypes were categorized in AT- or GC richer according to their difference in GC content, considering all SNPs in the tract or only the two flanking SNPs or only the SNPs that are not the two flanking SNPs.

[b]One-sample proportion test.

than AT_f haplotypes, this GC_f bias leads to an overall bias in favor of GC-richer haplotypes.

Discussion

Our analyses confirm that in yeast, when a GC/AT heterozygote site is involved in the conversion tract of a recombination event, the GC allele has a higher probability to be transmitted than the AT allele (Mancera et al. 2008). We show that this pattern of non-Mendelian segregation is specific of CO recombination events. Furthermore, we found that gBGC is essentially associated with simple conversion tracts (i.e., where all SNPs within the tract are converted from the same donor haplotype) and that the conversion bias depends on the nature of mismatches located at the extremities of the conversion tract. Thus, it appears that the decision to repair distal mismatches in one direction or the other affects all other mismatches in the heteroduplex, independently of their base composition. This phenomenon of "conversion sweep" (by analogy to selective sweeps) therefore tends to decrease the strength of gBGC. Indeed, the bias observed at the scale of conversion events in favor of GC-flanked haplotypes ($b = 0.075$, fig. 3) is much stronger than gBGC observed among the whole set of SNPs ($b = 0.013$, table 1). The departure from Mendelian expectation observed in the whole set of SNPs (50.6% instead of 50%, $b = 0.013$) might seem relatively weak. However, similar to natural selection, the impact of gBGC on the probability of allele fixation depends on the effective population size (N_e) and becomes strong when $N_e b \gg 1$ (Nagylaki 1983). Given that 1% of the yeast genome is affected by gene conversion during each meiosis (Mancera et al. 2008), the genome-wide gBGC coefficient is $b = 1.3 \times 10^{-4}$. Thus, in an obligate outcrossing species, such a gBGC drive would have a very strong impact, even for relatively small effective population sizes ($N_e \geq 10^5$). Yeast show a very low level of sexual reproduction and outcrossing, which reduces the population genetic effect of gBGC (Tsai et al. 2010). Nonetheless, there is evidence that gBGC affects the long-term evolution of yeast genomes (Birdsell 2002; Lynch et al. 2010; Tsai et al. 2010; Cutter and Moses 2011; Harrison and Charlesworth 2011).

Invalidation of the BER Hypothesis

To better understand the proximal causes of gBGC and the selective pressure that might operate on this genetic system, it is essential to identify the molecular mechanisms responsible for this conversion bias. In mammals, experiments in somatic cells demonstrated that the repair of DNA mismatches is strongly GC biased (Brown and Jiricny 1988, 1989; Bill et al. 1998). This GC bias results, at least in part, from the activity of the BER pathway, which involves DNA glycosylases that specifically excise thymines (and/or uracils) in DNA mismatches. Given that BER is intrinsically GC biased, it has been previously proposed that this repair mechanism, if active during meiosis, could be the cause of gBGC (Brown and Jiricny 1989; Galtier et al. 2001; Birdsell 2002; Marais 2003). BER leads to short patch repair and should therefore be frequently associated to complex conversion tracts. In the Mancera data set, the majority (≥89%) of conversion tracts are simple, as expected given the prominent role of MMR during recombination. However, a minor contribution of BER to the repair of mismatches in heteroduplex DNA cannot be a priori excluded. Calculations show that if a fraction of SNP conversion events result from the action of BER, then such cases must be at least 10 times more frequent among complex conversion tracts compared with simple conversion tracts (for details, see supplementary text S1, Supplementary Material online). Hence, if BER is the unique cause of gBGC, it is expected that the conversion bias in favor of GC alleles should be much stronger among complex conversion tracts than among simple tracts. However, in contradiction with this prediction, our analyses show that the largest source of gBGC corresponds to recombination events associated with simple conversion tracts. We, therefore, conclude that in *S. cerevisiae*, gBGC occurs in conversion events associated with a long-patch repair machinery and that the contribution of BER to the gBGC process, if any, is at most very minor.

Invalidation of the Initiation Bias Hypothesis

An alternative hypothesis is that gBGC could be the consequence of a bias in the initiation of recombination. It has been shown that the rate of DSB formation at a given locus may vary strongly between different haplotypes (Webb et al. 2008), and there is clear evidence that this initiation bias leads to a strong conversion bias in favor of the haplotype that is less prone to initiate recombination (Myers et al. 2010). Thus, if DSBs tend to occur more frequently on the AT-richest haplotype, this initiation bias might lead to gBGC. The analysis of DSB maps in *S. cerevisiae* did not reveal any clear association with AT-rich motifs (Murakami and Nicolas 2009; Pan et al. 2011), but a weak sequence preference, sufficient to cause the observed gBGC, cannot be a priori excluded. However, this initiation bias hypothesis is not consistent with our observation that gBGC is exclusively associated with CO recombination events. In yeast, CO hotspots and NCO hotspots generally coincide: Some recombination hotspots with biased CO/NCO ratios have been detected, but they represent only a tiny fraction (1.4%) of the regions involved in recombination events (Mancera et al. 2008). This indicates that generally, the same initiating regions can lead to both COs and NCOs. Hence, if the distribution of DSBs was the cause of gBGC, one would expect the same conversion bias in CO and NCO recombination events. The fact that gBGC is CO specific is therefore a strong argument indicating that the conversion bias is the consequence of a process that is posterior to the formation of DSBs. Note that in humans, the location of recombination hotspots is determined by a DNA-binding protein (PRDM9), which recognizes a specific sequence motif (Baudat et al. 2010). As predicted by the initiation bias model, the 13-bp genomic sequence motif targeted by PRDM9 has been subject to a rapid accumulation of substitutions in the human lineage (Myers et al. 2010). However, given that this motif is GC rich, this initiation bias tends to favor the fixation of G:C to A:T mutations. Hence, this

initiation bias cannot account for the gBGC process observed in the human genome.

MMR Model 1: Strand Rejection

Given that the BER and initiation bias models are rejected, an alternative hypothesis is that gBGC could be due to MMR. MMR plays a major role during recombination as a sensor of sequence homology during the process of strand invasion (Hunter et al. 1996; Chen and Jinks-Robertson 1999; Surtees et al. 2004). As mentioned in the introduction, it is in principle possible that the decision to reject the invading strand depends on the nature of mismatches present in the heteroduplex DNA. It is also possible that even in cases where the invading strand is not rejected, the extent of the heteroduplex is influenced by the presence of SNPs: When an SNP is encountered during the process of strand invasion, then either it is included in the heteroduplex (resulting in an additional mismatch) or the process of strand invasion is interrupted. Let us suppose that when the SNP that is encountered corresponds to an AT allele on the single-stranded 3′-overhang (and a GC allele on the intact homolog), the probability of interruption is lower than in the opposite configuration. Under this assumption, one expects an excess of cases where the mismatches at the extremities of heteroduplex DNA correspond to an AT on the broken chromatid and to a GC allele on the intact homolog. Thus, given that gene conversion occurs from the intact homolog toward the broken chromatid, this model predicts an excess of GC-flanked conversion tracts. Moreover, this model also predicts that GC-flanked conversion tracts should, on average, be longer than other conversion tracts. Both predictions, therefore, fit with our observations. However, one difficulty with this model is to explain why gBGC is CO specific. Yeast mutants lacking MSH2 show an increase both in the number of COs and NCOs (Martini et al. 2011), which suggests that MMR affects strand invasion for both categories of recombination events. Thus, if gBGC was due to the sensing of mismatches by MMR during the process of strand invasion, one would a priori expect to detect gBGC both in COs and NCOs.

MMR Model 2: Biased Mismatch Repair

An alternative (but non exclusive) hypothesis is that gBGC could result from the repair activity of MMR. MMR is composed of two main protein classes (Evans and Alani 2000; Jiricny 2006): MSH (MutS Homologs) proteins act as heterodimers to recognize mismatches along the sequence and recruit MLH (MutL Homologs) heterodimer proteins to form complexes. These complexes then migrate, in both directions from the mismatch, up to encountering a nick, where they will recruit an exonuclease. The degradation of the nick-containing strand is then followed by DNA resynthesis. It has been shown, both in vivo and in vitro, that the efficiency of mismatch repair by MMR depends on the nature of mismatches (Bishop et al. 1989; Mazurek et al. 2009; Martini et al. 2011). However, to our knowledge, it has not been investigated whether the "direction" of repair by MMR is affected by the nature of mismatches. In principle, it is only when nicks are present on both strands that there is a possibility of choice in the direction of repair. In the context of heteroduplex DNA formed during recombination, nicks are always present on the strand coming from the broken chromosome, but nicks can also be formed on the other strand during the resolution of recombination intermediates (Martini et al. 2011). The choice of one strand or the other will lead either to the conversion of the broken haplotype toward the nonbroken haplotype or to the restoration of Mendelian segregation. We propose a model, wherein the direction of the repair by MMR (toward conversion or restoration) depends on the nature of the mismatched bases that are close to the nicks flanking the heteroduplex (fig. 4). According to that model, when nicks are present on both strands, MMR would preferentially initiate DNA degradation from the nick closest to a mismatched A or T base. Thus, when the strand coming from the broken chromosome carries the GC allele, MMR would more frequently lead to the restoration of this haplotype, when compared with when it carries the AT allele. Hence, AT alleles would be converted more frequently than GC alleles, which would lead to an overall conversion bias in favor of GC alleles (fig. 4; supplementary text S3, Supplementary Material online). Note that the extent of the detected conversion tract (and hence the nature of the SNPs at its boundaries) depends on whether mismatch repair is directed toward conversion or restoration. As shown in figure 4, if AT alleles are less frequently restored than GC alleles, then GC-flanked conversion tracts are expected to be on average larger than other conversion tracts (supplementary text S3, Supplementary Material online, for details). Thus, this model would explain both the fact that gBGC is directed by the nature of the alleles located at the extremities of the conversion tract and the fact that GC-flanked conversion tracts are larger than other conversion tracts.

Again, one difficulty with this hypothesis is to explain why gBGC is CO specific. Current models indicate that COs result from dHj resolution (class I COs) and from the Mus81 pathway (class II COs) (Martini et al. 2011) (fig. 1). The class I CO pathway requires several meiosis-specific homologs of the MMR system (Hunter and Borts 1997; Argueso et al. 2004): The MLH1–MLH3 complex is involved in dHj resolution, whereas the MSH4–MSH5 complex is required in earlier steps of this pathway (Zakharyevich et al. 2012). However, MSH4 and MSH5 lack mismatch recognition domain and activity (Ross-Macdonald and Roeder 1994; Hollingsworth et al. 1995) and hence cannot be directly responsible for the biased mismatch repair. In fact, both in meiotic and mitotic cells, the recognition of base-base mismatches relies on the MSH2–MSH6 complex (for review, see Jiricny [2006]). In MSH2 mutants, meiotic recombination proceeds normally, but mismatches in heteroduplex DNA are left unrepaired, both for COs and NCOs (Martini et al. 2011). Given that both COs and NCOs rely on the same machinery for mismatch recognition, then how can gBGC be CO specific? One possible explanation is that the resolution of COs requires the formation of nicks of both DNA strands, in close vicinity (the average distance between Hj is approximately 260 bp [Cromie

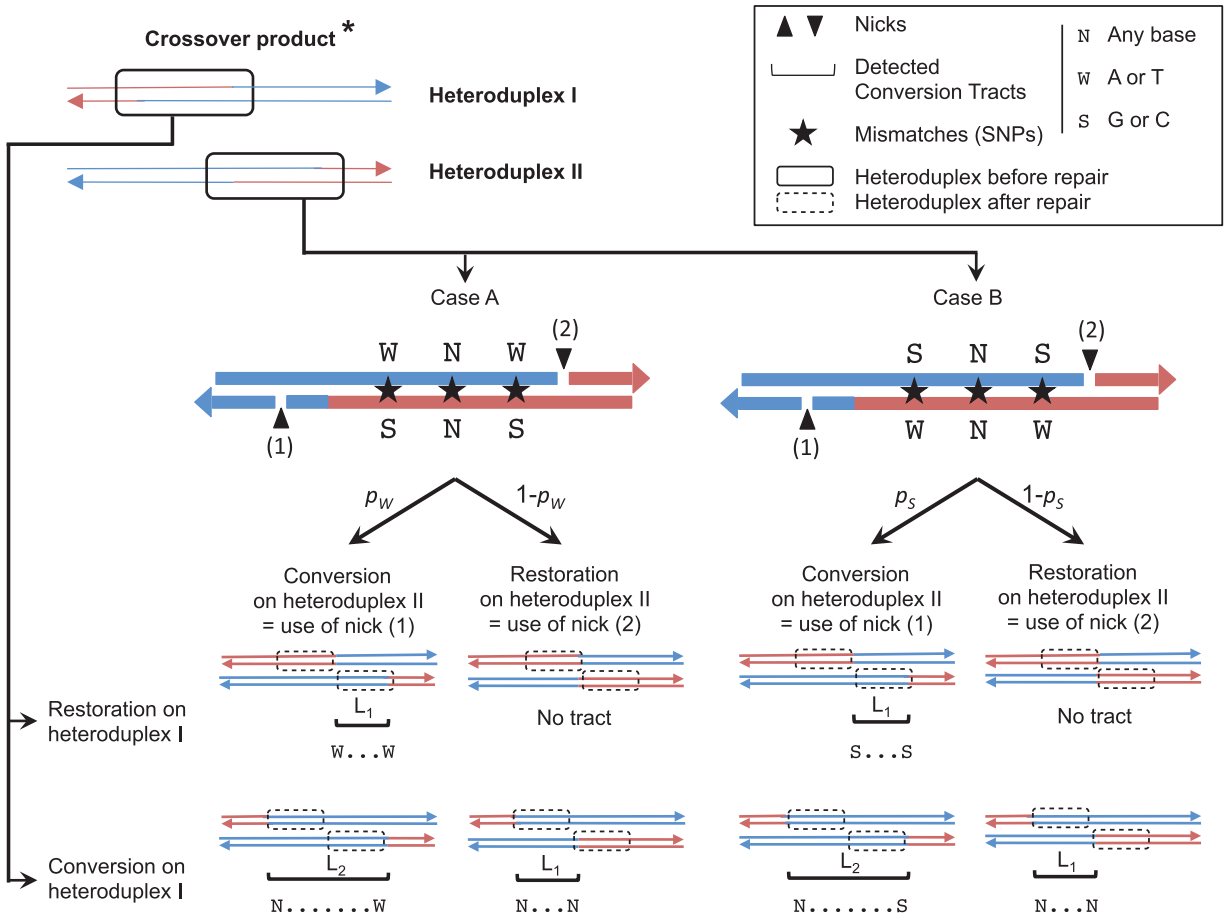

Fig. 4. Model of gBGC driven by GC-biased MMR repair. According to our model, gBGC results form a bias in the repair of mismatches by MMR, when nicks are present on both strands of the heteroduplex. This configuration potentially occurs during CO pathways (indicated by an * in fig. 1). COs involve the formation of two heteroduplexes. In the example shown, heteroduplex II consists of one GC-flanked haplotype (S = G or C) and one AT-flanked haplotype (W = A or T; *N* represents any mismatched base within the heteroduplex). MMR repair from nick (1) leads to the conversion of the red haplotype (the one that encountered the DSB initiating recombination), whereas the use of nick (2) leads to restoration. According to our model, the probability to use nick (2), instead of nick (1), is higher when the red strand carries the GC-flanked haplotype (case A) compared with when it carries the AT-flanked haplotype (case B). Thus, the probability of conversion is higher in case A than in case B (i.e., $P_S > P_W$). The detected conversion tract depends on the repair of both heteroduplexes: If both are restored, no tract can be detected. If only one heteroduplex is converted, the size of the tract (L_1) is expected to be smaller than if both are converted (L_2, with $L_1 < L_2$). In the case where the heteroduplex I is converted, the nature of the donor allele detected at the 5′-end of the tract (represented by an *N*) is independent of the haplotypes present in heteroduplex II. Given that $P_S > P_W$, this model predicts that among detected tracts with GC_f/AT_f polymorphism, there should be a transmission bias in favor of GC-flanked haplotype. Moreover, the model predicts that GC-flanked conversion tracts should on average be longer than AT-flanked ones (see details in supplementary text S3, Supplementary Material online). For simplicity, failures of mismatch repair (leading to postmeiotic segregation) are not considered here.

et al. 2006]). Thus, the presence of nicks on both DNA strands provides an opportunity for a bias in the direction of repair by MMR according to the nature of mismatches. Molecular pathways leading to NCOs also involve nicks on both strands (fig. 1). However, if these nicks are not in close proximity, or not present at the same time in NCO intermediates, then there would be no possible choice in the direction of repair. Thus, the fact that gBGC is CO specific could be due to differences in the spatiotemporal configuration of nicks in CO and NCO recombination intermediates.

Conclusion

In conclusion, our observations reject the BER and initiation bias models and are consistent with the hypotheses that gBGC is caused by MMR (via its role in strand invasion, in mismatch repair, or both). At this stage, the models of MMR-induced gBGC presented here remain speculative, and more data will be needed to test them. The hypothesis that gBGC is due to the repair activity of MMR makes several predictions that could be tested experimentally. First, this model predicts that the repair of AT:GC mismatches by MMR should be biased toward GC when nicks are present on both DNA strands. Second, this model implies that MSH2 should be active, not only during the early steps of recombination but also at the final step of CO pathway(s), during the resolution of joint molecules. Furthermore, it would be interesting to test whether gBGC is associated with both class I and class II CO pathways. In their analyses of recombination in yeast, Mancera et al. (2008) included five meioses from a mutant

of the class I CO pathway. Unfortunately, the limited number of recombination events detected was not sufficient to test whether gBGC occurs or not in this mutant (data not shown).

Given that the components of the recombination machinery are conserved across eukaryotes (Kolodner and Marsischky 1999), it seems likely that the same processes may be responsible for gBGC in other eukaryotes. One should note, however, that the relative contribution of the different CO recombination pathways differs among taxa. For example, fission yeast appears to rely exclusively on the class II pathway (Cromie et al. 2006), whereas most COs in mice result from the class I pathway (Holloway et al. 2008). If gBGC is specific of one of the two CO pathways, then one may expect differences in gBGC intensity among taxa.

One important issue is to understand the primary cause of the evolution of gBGC. In all taxa where some evidence of BGC has been reported, the conversion bias tends to favor GC alleles over AT alleles (Capra and Pollard 2011; Escobar et al. 2011; Pessia et al. 2012). This probably results from the fact that in most taxa, the pattern of mutation is biased toward AT (Lynch 2010), and hence any selective pressure to reduce the mutation rate is expected to favor the evolution of GC-biased mismatch repair. It should be noted that meiosis represents only a small fraction of the life cycle of eukaryotes. For example, in humans, germline cells are on average subject to 33 (in females) to approximately 200 (in males) mitotic cell divisions before meiosis (Chang et al. 1994). In nature, budding yeasts divide by meiosis only once every 1,000 generations (Tsai et al. 2010). Hence, most mutations occur in mitotic cells, where MMR plays a major role in the repair of DNA replication errors (Jiricny 2006). Thus, if the GC bias of MMR results from a selective pressure to reduce the mutation rate, then the strongest selective pressure should come from mutations that occur during mitosis. We therefore propose that the evolution of GC-biased MMR is driven by a selective pressure to reduce the rate of mutation in mitotic cells (including somatic cells, in the case of multicellular eukaryotes) and that gBGC simply results from the activity of this repair system during meiosis. Thus, under this hypothesis, gBGC would be a nonadaptive (and possibly maladaptive) indirect consequence of a selective pressure to limit the mutation rate in mitotic cells.

Materials and Methods

Data

We used recombination data, obtained by genotyping meiosis products of wild-type strains of *S. cerevisiae*, that were produced by Mancera et al. (2008). The list of conversion events associated with COs and NCOs, with details about parental and transmitted alleles, was kindly provided by Richard Bourgon. We filtered SNPs for which the base found in the spore was not called with enough confidence (labeled "NA" in the data). This led to a final list of 89,538 genotyped SNPs involved in conversion events, corresponding to 2,884 COs and 2,090 NCOs.

Measure of Gene Conversion Biases

We measured gene conversion biases at different scales: individual SNPs or haplotypes (see main text). In all cases, we considered sets of sites that are heterozygous in the parental hybrid and that were involved in conversion events (for the sake of generality, the two alleles are hereafter denoted Z and Y). For this set of sites, we counted the proportion of the allele Z in the offspring (x). We tested the existence of a conversion bias in favor of the allele Z by comparing x to the Mendelian expectation (50%), with a one-sample proportion test (see later). The intensity of the conversion bias in favor of the allele Z was measured by the coefficient $b = 2x - 1$.

Statistical Testing

Two types of tests were used on the proportion x as defined earlier. We used normal approximate two-tailed Z test with continuity correction to compare x to the Mendelian expectation of 50%. This is referred as "one-sample proportion test" in the text and legends. Additionally, we used normal approximate Z test with continuity correction to compare two different observed x proportions. This is referred as "two-sample proportion test" in the text and legends. Two-sample proportion tests are all two-tailed except when specified differently.

Supplementary Material

Supplementary texts S1–S3, figure S1 and S2, and tables S1–S5 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

The authors thank Richard Bourgon for providing the data set and for fruitful discussions about the observed transmission biases. They thank Bernard de Massy, Erika Kvikstad, Nicolas Lartillot, and Gaël Yvert for their very useful comments on an early version of the manuscript. This work was supported by the Centre National de la Recherche Scientifique.

References

Argueso JL, Wanat J, Gemici Z, Alani E. 2004. Competing crossover pathways act during meiosis in *Saccharomyces cerevisiae*. *Genetics* 168:1805–1816.

Auton A, Fledel-Alon A, Pfeifer S, et al. (23 co-authors). 2012. A fine-scale chimpanzee genetic map from population sequencing. *Science* 336: 193–198.

Baudat F, Buard J, Grey C, Fledel-Alon A, Ober C, Przeworski M, Coop G, de Massy B. 2010. PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice. *Science* 327:836–840.

Berglund J, Pollard KS, Webster MT. 2009. Hotspots of biased nucleotide substitutions in human genes. *PLoS Biol.* 7:45–62.

Bill CA, Duran WA, Miselis NR, Nickoloff JA. 1998. Efficient repair of all types of single-base mismatches in recombination intermediates in Chinese hamster ovary cells: competition between long-patch and G-T glycosylase-mediated repair of G-T mismatches. *Genetics* 149: 1935–1943.

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

Bishop DK, Andersen J, Kolodner RD. 1989. Specificity of mismatch repair following transformation of *Saccharomyces cerevisiae* with heteroduplex plasmid DNA. *Proc Natl Acad Sci U S A.* 86:3713–3717.

Brown TC, Jiricny J. 1988. Different base/base mispairs are corrected with different efficiencies and specificities in monkey kidney cells. *Cell* 54: 705–711.
Brown TC, Jiricny J. 1989. Repair of base-base mismatches in simian and human cells. *Genome* 31:578–583.
Capra JA, Pollard KS. 2011. Substitution patterns are GC-biased in divergent sequences across the metazoans. *Genome Biol Evol.* 3: 516–527.
Chang B, Shimmin L, Shyue S-K, Hewett-Emmett D, Li W-H. 1994. Weak male driven molecular evolution in rodents. *Proc Natl Acad Sci U S A.* 91:827–831.
Chen SY, Tsubouchi T, Rockmill B, Sandler JS, Richards DR, Vader G, Hochwagen A, Roeder GS, Fung JC. 2008. Global analysis of the meiotic crossover landscape. *Dev Cell.* 15:401–415.
Chen W, Jinks-Robertson S. 1999. The role of the mismatch repair machinery in regulating mitotic and meiotic recombination between diverged sequences in yeast. *Genetics* 151:1299–1313.
Coïc E, Gluck L, Fabre F. 2000. Evidence for short-patch mismatch repair in *Saccharomyces cerevisiae*. *EMBO J.* 19:3408–3417.
Coop G, Przeworski M. 2007. An evolutionary view of human recombination. *Nat Rev Genet.* 8:23–34.
Cromie GA, Hyppa RW, Taylor AF, Zakharyevich K, Hunter N, Smith GR. 2006. Single Holliday junctions are intermediates of meiotic recombination. *Cell* 127:1167–1178.
Cutter AD, Moses AM. 2011. Polymorphism, divergence, and the role of recombination in *Saccharomyces cerevisiae* genome evolution. *Mol Biol Evol.* 28:1745–1754.
de Massy B. 2003. Distribution of meiotic recombination sites. *Trends Genet.* 19:514–522.
Duret L, Arndt PF. 2008. The impact of recombination on nucleotide substitutions in the human genome. *PLoS Genet.* 4:1–19.
Duret L, Galtier N. 2009. Biased gene conversion and the evolution of mammalian genomic landscapes. *Annu Rev Genomics Hum Genet.* 10:285–311.
Escobar JS, Glémin S, Galtier N. 2011. GC-biased gene conversion impacts ribosomal DNA evolution in vertebrates, angiosperms, and other eukaryotes. *Mol Biol Evol.* 28:2561–2575.
Evans E, Alani E. 2000. Roles for mismatch repair factors in regulating genetic recombination. *Mol Cell Biol.* 20:7839–7844.
Galtier N, Duret L. 2007. Adaptation or biased gene conversion? Extending the null hypothesis of molecular evolution. *Trends Genet.* 23:273–277.
Galtier N, Duret L, Glémin S, Ranwez V. 2009. GC-biased gene conversion promotes the fixation of deleterious amino acid changes in primates. *Trends Genet.* 25:1–5.
Galtier N, Piganeau G, Mouchiroud D. 2001. GC-content evolution in mammalian genomes: the biased gene conversion hypothesis. *Genetics* 159:907–911.
Glémin S. 2010. Surprising fitness consequences of GC-biased gene conversion: I. Mutation load and inbreeding depression. *Genetics* 185: 939–959.
Glémin S. 2011. Surprising fitness consequences of GC-biased gene conversion. II. Heterosis. *Genetics* 187:217–227.
Harrison RJ, Charlesworth B. 2011. Biased gene conversion affects patterns of codon usage and amino acid usage in the *Saccharomyces* sensu stricto group of yeasts. *Mol Biol Evol.* 28:117–129.
Hoffmann ER, Borts RH. 2004. Meiotic recombination intermediates and mismatch repair proteins. *Cytogenet Genome Res.* 107: 232–248.
Hollingsworth N, Ponte L, Halsey C. 1995. MSH5, a novel MutS homolog, facilitates meiotic reciprocal recombination between homologs in *Saccharomyces cerevisiae* but not mismatch repair. *Genes Dev.* 9: 1728–1739.
Holloway JK, Booth J, Edelmann W, McGowan CH, Cohen PE. 2008. MUS81 generates a subset of MLH1-MLH3-independent crossovers in mammalian meiosis. *PLoS Genet.* 4:e1000186.
Holmes J, Clark S. 1990. Strand-specific mismatch correction in nuclear extracts of human and *Drosophila melanogaster* cell lines mismatch. *Proc Natl Acad Sci U S A.* 87:5837–5841.
Hunter N, Borts RH. 1997. MLH1 is unique among mismatch repair proteins in its ability to promote crossing-over during meiosis. *Genes Dev.* 11:1573–1582.
Hunter N, Chambers SR, Louis EJ, Borts RH. 1996. The mismatch repair system contributes to meiotic sterility in an interspecific yeast hybrid. *EMBO J.* 15:1726–1733.
Jiricny J. 2006. The multifaceted mismatch-repair system. *Nat Rev Mol Cell Biol.* 7:335–346.
Katzman S, Capra JA, Haussler D, Pollard KS. 2011. Ongoing GC-biased evolution is widespread in the human genome and enriched near recombination hot spots. *Genome Biol Evol.* 3:614–626.
Kolodner RD, Marsischky GT. 1999. Eukaryotic DNA mismatch repair. *Curr Opin Genet Dev.* 9:89–96.
Krogh BO, Symington LS. 2004. Recombination proteins in yeast. *Annu Rev Genet.* 38:233–271.
Lynch DB, Logue ME, Butler G, Wolfe KH. 2010. Chromosomal G + C content evolution in yeasts: systematic interspecies differences, and GC-poor troughs at centromeres. *Genome Biol Evol.* 2: 572–583.
Lynch M. 2010. Rate, molecular spectrum, and consequences of human mutation. *Proc Natl Acad Sci U S A.* 107:961–968.
Mancera E, Bourgon R, Brozzi A, Huber W, Steinmetz LM. 2008. High-resolution mapping of meiotic crossovers and non-crossovers in yeast. *Nature* 454:479–485.
Marais G. 2003. Biased gene conversion: implications for genome and sex evolution. *Trends Genet.* 19:330–338.
Martini E, Borde V, Legendre M, Audic S, Regnault B, Soubigou G, Dujon B, Llorente B. 2011. Genome-wide analysis of heteroduplex DNA in mismatch repair-deficient yeast cells reveals novel properties of meiotic recombination pathways. *PLoS Genet.* 7:1–18.
Mazurek A, Johnson CN, Germann MW, Fishel R. 2009. Sequence context effect for hMSH2-hMSH6 mismatch-dependent activation. *Proc Natl Acad Sci U S A.* 106:1–6.
McMahill MS, Sham CW, Bishop DK. 2007. Synthesis-dependent strand annealing in meiosis. *PLoS Biol.* 5:2589–2601.
Memisoglu A, Samson L. 2000. Base excision repair in yeast and mammals. *Mutat Res.* 451:39–51.
Meunier J, Duret L. 2004. Recombination drives the evolution of GC-content in the human genome. *Mol Biol Evol.* 21:984–990.
Murakami H, Nicolas A. 2009. Locally, meiotic double-strand breaks targeted by Gal4BD-Spo11 occur at discrete sites with a sequence preference. *Mol Cell Biol.* 29:3500–3516.
Myers S, Bowden R, Tumian A, Bontrop RE, Freeman C, MacFie TS, McVean G, Donnelly P. 2010. Drive against hotspot motifs in primates implicates the *PRDM9* gene in meiotic recombination. *Science* 327:876–879.
Nagylaki T. 1983. Evolution of a finite population under gene conversion. *Proc Natl Acad Sci U S A.* 80:6278–6281.
Necşulea A, Popa A, Cooper DN, Stenson PD, Mouchiroud D, Gautier C, Duret L. 2011. Meiotic recombination favors the spreading of deleterious mutations in human populations. *Hum Mutat.* 32: 198–206.
Nicolas A, Treco D, Schultes NP, Szostak JW. 1989. An initiation site for meiotic gene conversion in the yeast *Saccharomyces cerevisiae*. *Nature* 338:35–39.
Pan J, Sasaki M, Kniewel R, et al. (12 co-authors). 2011. A hierarchical combination of factors shapes the genome-wide topography of yeast meiotic recombination initiation. *Cell* 144:719–731.
Pessia E, Popa A, Mousset S, Rezvoy C, Duret L, Marais GAB. 2012. Evidence for widespread GC-biased gene conversion in eukaryotes. *Genome Biol Evol.* 4:675–682.
Qi J, Wijeratne A, Tomsho L, Hu Y, Schuster S, Ma H. 2009. Characterization of meiotic crossovers and gene conversion by whole-genome sequencing in *Saccharomyces cerevisiae*. *BMC Genomics* 10:475.
Ratnakumar A, Mousset S, Glémin S, Berglund J, Galtier N, Duret L, Webster MT. 2010. Detecting positive selection within genomes: the problem of biased gene conversion. *Philos Trans R Soc Lond B Biol Sci.* 365:2571–2580.

Ross-Macdonald P, Roeder GS. 1994. Mutation of a meiosis-specific MutS homolog decreases crossing over but not mismatch correction. *Cell* 79:1069–1080.

Smith KN, Nicolas A. 1998. Recombination at work for meiosis. *Curr Opin Genet Dev.* 8:200–211.

Surtees JA, Argueso JL, Alani E. 2004. Mismatch repair proteins: key regulators of genetic recombination. *Cytogenet Genome Res.* 107: 146–159.

Thomas DC, Roberts JD, Kunkel TA. 1991. Heteroduplex repair in extracts of human HeLa cells. *J Biol Chem.* 266:3744–3751.

Tsai IJ, Burt A, Koufopanou V. 2010. Conservation of recombination hotspots in yeast. *Proc Natl Acad Sci U S A.* 107:7847–7852.

Webb AJ, Berg IL, Jeffreys A. 2008. Sperm cross-over activity in regions of the human genome showing extreme breakdown of marker association. *Proc Natl Acad Sci U S A.* 105:10471–10476.

Webster MT, Hurst LD. 2012. Direct and indirect consequences of meiotic recombination: implications for genome evolution. *Trends Genet.* 28:101–109.

Winzeler EA, Richards DR, Conway AR, et al. (11 co-authors). 1998. Direct allelic variation scanning of the yeast genome. *Science* 281: 1194–1197.

Zakharyevich K, Tang S, Ma Y, Hunter N. 2012. Delineation of joint molecule resolution pathways in meiosis identifies a crossover-specific resolvase. *Cell* 149:334–347.