High-resolution genome-wide analysis of irradiated (UV and gamma rays) diploid yeast cells reveals a high frequency of genomic loss of heterozygosity (LOH) events.


*Department of Molecular Genetics and Microbiology and §Department of Pharmacology and Cancer Biology, Duke University Medical Center, Durham, NC 27710

**Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599

§§ contributed equally to the study
Running head: High-resolution genomic mapping of LOH

Key words/phrases: mitotic recombination, loss of heterozygosity, DNA repair, *Saccharomyces cerevisiae*, single-nucleotide polymorphism (SNP) microarrays.

Corresponding author:

Thomas D. Petes

Department of Molecular Genetics and Microbiology

Box 3054, Duke University Medical Center

Durham, NC 27710

Telephone: (919) 684-4986

Fax: (919) 684-6033

Email: tom.petes@duke.edu
ABSTRACT

In diploid eukaryotes, repair of double-stranded DNA breaks (DSBs) by homologous recombination often leads to loss of heterozygosity (LOH). Most previous studies of mitotic recombination in S. cerevisiae have focused on a single chromosome or a single region of one chromosome at which LOH events can be selected. In this study, we used two techniques (single-nucleotide polymorphism [SNP] microarrays and high-throughput DNA sequencing [HTS]) to examine genome-wide LOH in a diploid yeast strain at a resolution averaging one kb. We examined both selected LOH events on chromosome V and unselected events throughout the genome in untreated cells, and cells treated with either γ radiation or ultraviolet radiation (UV). Our analysis shows: 1) spontaneous and damage-induced mitotic gene conversion tracts are more than three times larger than meiotic conversion tracts, and conversion tracts associated with crossovers are usually longer and more complex than those unassociated with crossovers, 2) most of the crossovers and conversions reflect the repair of two sister chromatids broken at the same position, and 3) both UV and γ radiation efficiently induce LOH at doses of radiation that cause no significant loss of viability. Using HTS, we also detected new mutations induced by γ-rays and UV. To our knowledge, our study represents the first high-resolution genome-wide analysis of DNA damage-induced LOH events performed in any eukaryote.
INTRODUCTION

All organisms experience DNA damage from both exogenous and endogenous sources. Endogenous DNA damage includes spontaneous deamination of nucleotides, depurination/depyrimidination, oxidative damage, and double-stranded DNA breaks (DSBs) (FRIEDBERG et al. 2006). DSBs are likely to be particularly deleterious, since unrepaired DSBs can lead to chromosome rearrangements or chromosome loss. Although the sources of endogenous DSBs have not been completely determined, some DSBs appear to reflect nuclease processing of secondary DNA structures (such as DNA “hairpins”) or head-on collisions between the replication and transcription machineries (AGUILERA 2002). Below, in addition to examining spontaneous recombination events that presumably reflect the repair of endogenous DNA damage, we also analyze recombination events induced by two exogenous sources: γ rays and ultraviolet light (UV).

Both γ rays and UV cause a variety of different types of DNA damage. γ rays cause DSBs, single-stranded DNA nicks, and base damage (FRIEDBERG et al. 2006; WARD 1990). UV results in pyrimidine dimers (FRANKLIN et al. 1985; SETLOW 1966), DNA-DNA or DNA-protein crosslinks (PEAK and PEAK 1986), and single-stranded DNA nicks resulting from the dimer excision (BREEN and MURPHY 1995).

In yeast, as in most eukaryotes, there are two recombination pathways that are used to repair DSBs: non-homologous end-joining (NHEJ) and homologous recombination (HR). In NHEJ events, as the name implies, broken DNA molecules are re-joined by a mechanism that requires little or no homology (DALEY et al. 2005). This mechanism is most active in haploid yeast cells during G1 of the cell cycle (SHRIVASTAV et al. 2008). In diploid cells, the dominant pathway is HR. HR uses an intact homologous DNA molecule, either the sister
chromatid or the homologous chromosome, as a template for repair of the broken chromosome.

DSBs can be repaired by several different HR pathways (HEYER et al. 2010). The repair of a DSB by gene conversion unassociated with a crossover is shown in Fig. 1A. This process involves the non-reciprocal transfer of sequences from the intact donor molecule to the broken chromosome in several steps: 1) invasion of one broken end into the intact template molecule, followed by DNA synthesis primed by the invading 3’ strand, 2) removal of the invading end and reannealing of this end back to the other broken end, forming a heteroduplex with mismatches, and 3) repair of the mismatches. This mechanism (synthesis-dependent strand-annealing or SDSA) was first suggested to explain some features of meiotic recombination in yeast (ALLERS and LICHTEN 2001). In the second pathway (Fig. 1B), gene conversion may be associated with crossovers. In this pathway, a double Holliday junction (dHJ) is formed that can be resolved to yield a crossover or non-crossover. In this pathway, heteroduplexes flank the original position of the DSB. Although the heteroduplex regions have the same size in Fig. 1B, in both meiosis (JESSOP et al. 2005; MERKER et al. 2003) and mitosis (MITCHEL et al. 2010; TANG et al. 2011), the conversion tracts flanking the DSB are often of different lengths. The dHJ can also be dissolved without nucleolytic cleavage of DNA strands to yield non-crossover products with heteroduplexes located in cis on one of the two interacting chromosomes (HEYER et al. 2010). In the third pathway (Fig. 1C), one part of the broken DNA molecule is lost and a complete chromosome is then reconstructed by break-induced replication (BIR). In this mechanism, one of the broken ends invades the intact template molecule and a replication fork is set up that duplicates the template from the site of invasion to the telomere.
If HR involves an interaction between two homologues that can be distinguished by single-nucleotide polymorphisms (SNPs), conversions without crossovers will produce a small patch of LOH within a chromosome that is otherwise heterozygous, whereas both crossovers and BIR result in LOH that extends from the site at which the event initiates to the end of the chromosome. Repair of a DSB by HR in which the homologous chromosome is used as a template will result in LOH, but repair events involving the sister-chromatid will not. Although most sister-chromatid recombination events are genetically silent, unequal sister-chromatid exchanges can be detected in yeast by a variety of different systems (Petes and Hill 1988). Using one of these systems, Kadyk and Hartwell (Kadyk and Hartwell 1992) showed that, in diploid cells, sister chromatids are the preferred substrate for the repair of DSBs generated by X-rays. Despite this preference, it is clear that ionizing radiation and UV strongly stimulate HR events (both mitotic crossovers and gene conversions) between homologous chromosomes (Fabre 1978; Nakai and Mortimer 1969).

One problem with studying spontaneous mitotic recombination is that most analytic systems do not allow the selection of both daughter cells that contain the recombinant chromosomes. Several years ago, we developed a method of selecting reciprocal crossovers on chromosome V that surmounts this difficulty (Barbera and Petes 2006; Lee et al. 2009). One copy of chromosome V has the can1-100 allele (an ochre mutation) and, in the other copy, the CAN1 gene is replaced by SUP4-o, a gene encoding an ochre-suppressing tRNA (Fig. 2A). In the absence of the suppressor, strains with the can1-100 allele are resistant to canavanine, but because of the suppressor, the diploid used in our experiments is canavanine-sensitive. In addition, the diploid is homozygous for the ade2-1 mutation (an ochre mutation). Strains with this mutation, in the absence of the SUP4-o gene form red
colonies, but form pink colonies if one copy of the SUP4-o gene is present. Thus, the diploid strain is canavanine-sensitive and forms pink colonies. If a crossover occurs between the centromere of chromosome V and the can1-100/SUP4-o markers (a distance of about 120 kb), a canavanine-resistant red/white colony is formed (Fig. 2A).

Although this method was first used in diploid strains lacking polymorphisms, subsequent studies were done in which a diploid was constructed using two haploid strains that had about 0.5% sequence divergence (Lee et al. 2009; Lee and Petes 2010), resulting in about 55,000 SNPs distributed throughout the genome. Crossovers and associated gene conversions on chromosome V were mapped to a resolution of about 4 kb by using PCR, restriction digests, and gel electrophoresis to look for LOH. Although a few of the crossovers had no associated conversion, most of the crossovers were associated with an adjacent conversion event (boxed regions in Fig. 2). In the 3:1 class of events (Fig. 2B), in the boxed region, three of the four chromosomes had one of the two forms of the SNP and one of the chromosomes had the other form (one sector being homozygous for a SNP with the other sector being heterozygous). In addition, about 40% of the crossovers were associated with a conversion event in which the same SNP was homozygous in both sectors (Fig. 2C); we term these events “4:0 conversions.” The observation of the 4:0 events argues that about half of mitotic crossovers result from the repair of two sister chromatids that are broken at approximately the same positions. One simple mechanism for generating this intermediate is to have the DSB occur in G1, and the broken chromosome replicate to form two broken chromatids (Lee et al. 2009). This proposed mechanism was confirmed by analysis of the types of conversion events stimulated by γ rays in synchronized G1 and G2 cells (Lee and Petes 2010). In addition to 3:1 and 4:0 conversion events, 3:1/4:0 hybrid
tracts are also observed (Lee et al. 2009; Lee and Petes 2010). Such tracts can also be explained as a consequence of repair of two broken sister chromatids (Fig. 3).

In the current study, we use SNP microarrays, and HTS to map selected events on chromosome V as well as unselected events throughout the genome at a resolution of about 1 kb. To our knowledge, these studies are the first to measure the numbers and types of LOH events throughout the genome induced by doses of ionizing radiation (100 Gy) and UV (10-15 J/m²) that have no significant effect on cell viability. We also determined the number of mutations induced in the genomes by these treatments.
MATERIALS AND METHODS

Strains and genetic methods: All experiments were conducted with the diploid strain PG311 (Lee et al. 2009). The relevant genotype of PG311 is $MAT\alpha::NAT^R/MAT\alpha URA3/ura3-\!1$ ade2-1/ade2-1 TRP1/trp1-1 HIS3/his3-11,15 GAL2/gal2 SUP4-o/can1-100 V9229/V9229::HYG V261553/V261553::LEU2. This diploid was generated by crossing the haploid strains PSL2 and PSL5 which are isogenic with strains W303a and YJM789, respectively, except for alterations introduced by transformation (Lee et al. 2009). Below, we will refer to the haploid parents of PG311 as W303a and YJM789. In general, PG311 has the SNPs predicted from the haploid parents. The disruption of $MAT\alpha$ in PG311 allows synchronization of this diploid by alpha factor. Although diploids that lack $MAT\alpha$ do not sporulate under normal conditions, such strains can be sporulated on plates containing 5 mM nicotinamide (J. Rine, personal communication). For experiments in which we analyzed meiotic products of PG311, the stains were pre-grown on YPD plates with 5 mM of nicotinamide, and then transferred to sporulation plates containing 5 mM nicotinamide. Plates were incubated at 25° for 2-4 days before tetrad dissection.

Standard media were used (Guthrie and Fink 1991) unless noted. To detect spontaneous crossovers, we first isolated single colonies of PG311 grown on rich growth medium (YPD) at 30° for two days. Individual colonies were suspended in 400 μl of dH₂O, and 100 μl of this mixture was plated on canavanine-containing plates (SD-arg + 120 μg/ml canavanine). The plates were incubated four days at room temperature, followed by incubation for 16 hours at 4°; the 4° incubation allows better visualization of the red sectors. We purified cells from the red and white sectors for subsequent analysis.
In the experiments in which cells were irradiated, we synchronized cells in G1 using alpha factor (Lee and Petes 2010). The synchronized cells were treated with either γ radiation in a Shepherd Mark 1 Cesium-137 irradiator at 100 Gy or with UV (254 nm) derived from a TL-2000 Ultraviolet Translinker at a dosage of 10 or 15 J/m². Following radiation, the cells were plated either on non-selective plates (SD- arg) or plates that lacked arginine and contained 120 μg/ml canavanine. The subsequent growth of the cells and the analysis of sectors were done as described above for the spontaneous selection with the exception that sectored colonies for the UV-treated samples were isolated from non-selective plates grown at 30° instead of room temperature.

**SNP microarrays: design and optimization:** We designed the SNP arrays based on genomic sequence information available from SGD for S288c (very closely related to W303a) (Winzeler et al. 2003) and YJM789 (Wei et al. 2007). Microarrays that were capable of detecting LOH for SNPs in PG311 were designed based on principles outlined by Gresham et al. (Gresham et al. 2010). For each SNP represented on the array, four 25-base oligonucleotides were used: one for each strand of the W303a SNP, and one for each strand of the YJM789 SNP. The SNP was located in the middle of the 25-base oligonucleotide. Although there are approximately 55,000 SNPs in PG311, about three-quarters of these SNPs were not used for our analysis. We excluded most of the SNPs found in repeated genes. We also screened out oligonucleotides that had a melting temperature for the perfectly-matched 25 base pair duplex that was less than 55° or greater than 59°. The remaining oligonucleotides (representing about 15,000 SNPs) that were printed on the microarray are listed in Table S3. We also included on the array about 120 oligonucleotides that were not different between W303a and YJM789; these are listed in Table S4.
Oligonucleotides were printed onto the microarrays by Agilent Technologies in slides containing about 105,000 oligonucleotides. Many of the oligonucleotides are represented more than once in the microarrays. Following experiments to determine which oligonucleotides resulted in the most specific hybridization signals (described in Supporting Information), we reduced the number of SNPs used in our analysis to 13,000. This final set of oligonucleotides is presented in Table S5.

**Methods used for microarray analysis: sample preparation, hybridization conditions, and data analysis:** The methods used for sample preparation, hybridization conditions, and data analysis are similar to those described previously (LEMOINE et al. 2005; McCULLEY and PETES 2010). A detailed description of these procedures is presented in Supporting Information. In brief, genomic DNA from the experimental strain was labeled with Cy5-dUTP, DNA from the control strain (PG311) was labeled with Cy3-dUTP, and the two labeled samples were competitively hybridized to the microarrays. The arrays were then scanned at wavelengths of 635 and 532 nm using a GenePix scanner and GenePix Pro software using settings recommended by the manufacturer. The ratio of the medians (635 nm/532 nm; RM) for each probe was used for analysis, and replicate RMs were averaged. The data were centered around a value of one by dividing each probe RM by the average of all of the probe RMs in order to normalize for differences in the hybridization levels for the control and experimental strain samples.

We calculated 95% confidence intervals on the median sizes of conversion tracts using Table B11 of Altman (ALTMAN 1990). Comparisons of conversion tract lengths under different experimental conditions were done using the Mann-Whitney test on the VassarStats Website (http://faculty.vassar.edu/lowry/VassarStats.html).
Generation and analysis of HTS data: Samples were prepared for HTS as described above for the SNP microarray sample preparation with the exception that genomic DNA was sonicated to 300-700 bp fragments. The DNA was then prepared for sequencing using the protocol recommended by Illumina for the Genome Analyzer IIx. The samples were sequenced using an Illumina GAIIx machine, generating 67- to 75-bp paired end reads. For the eight sequenced samples, coverage varied from 90- to 180-fold.

The details of the HTS data analysis are presented in Supporting Information. In brief, we detected regions of LOH by identifying SNPs in the experimental strains in which at least 90% of the “reads” that were originally heterozygous were now identical to one of the original alleles. We identified new mutations by finding bases that were identical in the original diploid, but had a novel base in at least 40% of the “reads” in the irradiated diploid; we use the 40% criterion because we expect that any new mutation will be heterozygous. Mutations that appeared in more than one independent isolate were not counted as de novo mutations, since such mutations presumably arose in the strain before treatment with the DNA damaging agent.
RESULTS

As described in the Introduction, we previously selected spontaneous mitotic crossovers, as well as crossovers induced by \( \gamma \) rays, that occurred between the centromere of chromosome V and the \textit{can1-100/SUP4-o} markers, an interval of about 120 kb (Fig. 2) \cite{lee2009, lee2010}. The diploid used in these studies (PG311) was constructed by a cross between two haploids that are isogenic with W303a and YJM789, and is heterozygous for about 55,000 SNPs. In our previous analysis, the positions of the crossovers and associated gene conversion events were mapped to a resolution of about 4 kb using a PCR-based strategy that determined whether the SNPs were heterozygous or homozygous. This procedure is impractical for genome-wide mapping of recombination events. Below, we describe the use of SNP arrays to map spontaneous, UV-induced and \( \gamma \) ray-induced crossovers selected on chromosome V, as well as unselected crossovers and gene conversion events throughout the genome.

SNP arrays have been used previously to map LOH events in tumor cells \cite{lin2000}, to map meiotic recombination events in \textit{S. cerevisiae} \cite{mance2008}, to characterize chromosome rearrangements and chromosome loss in \textit{C. albicans} \cite{abbe2011}, and in a variety of other experiments. Using principles outlined by GRESHAM et al. \cite{gresham2010}, we developed a SNP array to examine LOH events throughout the genome in the diploid PG311. This array has oligonucleotides that distinguish over 13,000 SNPs, resulting in an average density of one oligonucleotide per kb of genomic DNA. The details of the array design and the specific sequences of the probes are in Supporting Materials and Methods and in Tables S3-S5.
In our experiments, we labeled genomic DNA from a diploid with a recombination event with one fluorescent nucleotide and DNA from the control diploid PG311 with a different fluorescent nucleotide. The samples were mixed and competitively hybridized to the SNP arrays. If SNPs retained heterozygosity in the experimental strain, then the hybridization signal for the oligonucleotides representing that SNP was similar to that of the control strain. LOH was detected by an increase in hybridization to oligonucleotides that had one form of the SNP (for example, the W303a form) and a decrease in hybridization to oligonucleotides that had the other form (for example, the YJM789 form).

As described above, crossover events between CEN5 and the can1-100/SUP4-o markers produce canavanine-resistant red/white sectored colonies. Thus, all of the samples analyzed had a selected recombination event on chromosome V, and we found that samples treated with UV or γ rays also had unselected events on other chromosomes. We isolated genomic DNA from both the red and white sides of the sectors and examined the DNA by SNP arrays for diploids untreated with recombinogenic agents, as well as for diploids treated with UV or γ rays. An example of the analysis for chromosome V for a selected spontaneous crossover (PG311-2A) is shown on Fig. 4. In Fig. 4A, a low-resolution depiction of the hybridization levels is shown for both the red (top) and white (bottom) sectors. In both sectors, the hybridization pattern indicates that the transition from heterozygous SNPs to homozygous SNPs is at approximately SGD coordinate 55000. As expected, the DNA that is centromere-distal to the crossover junction from the red sector hybridizes well to the W303a-specific probes (red line) and poorly to the YJM789-specific probes (blue line), since the red sector is generated by LOH events that include the can1-100 marker that is derived from the W303a-related homologue (Fig. 2). Genomic DNA from
the white sector shows the reciprocal pattern of hybridization. The “spike” of increased hybridization in the red sector for YJM789 SNPs near SGD coordinate 30000 is an artifact resulting from a deletion of YJM789 sequences during the insertion of the SUP4-o gene into the YJM789-derived chromosome.

In Fig. 4B, we show the same crossover event at higher resolution. Each square in this figure shows the hybridization ratio to a specific oligonucleotide. In the red sector, the transition between the homozygous SNPs and the heterozygous SNPs is between SGD coordinates 57170 and 60701. In the white sector, the transition occurs between 51915 and 53692. Thus, there is a region (boxed in Fig. 4B) in which one sector is homozygous for SNPs and the other is heterozygous. This region is a 3W:1Y gene conversion tract (W and Y indicating W303a-derived and YJM789-derived SNPs, respectively), equivalent to the boxed region in Fig. 2B; in our subsequent discussions, a 3:1 conversion event indicates 3W:1Y conversion and 1:3 indicates 1W:3Y conversion. We estimate the length of the gene conversion tract by averaging the maximal length of the tract (the distance between markers that are not within the gene conversion tract, 8.8 kb) and the minimal length of the tract (the distance between the converted markers, 3.5 kb). For the tract shown in Fig. 4B, this length is about 6.2 kb. It is important to emphasize that the presence and extent of gene conversion tracts can only be identified when the patterns of LOH are analyzed in genomic DNA from both sectors of the sectored colony.

With the genome-wide SNP analysis, we found that the parental diploid PG311 and all of its subsequent derivatives had two LOH events that were unexpected from the sequence of the parental haploids. All strains were homozygous for W303a-derived SNPs centromere-distal to SGD coordinate 685 kb on chromosome XIII, and were homozygous
for W303a-derived SNPs between coordinates 412715 and 414085 on chromosome X. Since these events, presumably generated during sub-culturing of PG311, were present in all strains, they were excluded from our analysis.

**Analysis of spontaneous selected LOH events on chromosome V by SNP arrays:**

We examined by SNP arrays genomic DNA from both the red and white sectors from fourteen independent canavanine-resistant colonies. The crossover events in five of the isolates had been mapped previously by PCR-amplifying regions along the left arm of chromosome V that contained polymorphic restriction enzyme sites and testing for homozygosity or heterozygosity by a restriction enzyme digest analysis (Lee et al. 2009). We also mapped an additional nine events solely by SNP arrays. In the previous study, we used 34 markers in the 120 kb *CEN5-can1-100/SUP4-o* interval. We monitored 172 markers in this same interval by SNP arrays.

In our previous study, we found that most of the spontaneous crossovers were crossovers without conversions (Fig. 2A), crossovers with associated 3:1 conversions (Fig. 2B), crossovers with 4:0 conversions (Fig. 2C), or crossovers with hybrid conversions (3:1/4:0 or 3:1/4:0/3:1 tracts; Fig. 3) (Lee et al. 2009). A comparison of the mapping of recombination events by the PCR-based method and SNP arrays for four of the sectored colonies is shown in Table 1. In this table, we define the position of the crossovers with only two SGD coordinates: the position of the centromere-proximal heterozygous marker that is closest to the crossover/conversion event and the position of the centromere-distal homozygous markers representing the crossover. Although the agreement between the two methods was reasonably good, as expected, the SNP array mapped events with better resolution and also revealed that some of the conversion events were more complex than
previously determined. For example, in PG311-1.4, we previously mapped a crossover between SGD coordinates 125754 and 133080 that appeared to be unassociated with gene conversion. With the SNP arrays, we mapped the transition at higher resolution between SGD coordinates 127038 and 130096, and we detected a SNP in this region that had undergone gene conversion. The complete description of all spontaneous crossovers and associated gene conversion events is given in Table S6.

One exception to the generally good agreement between the two mapping methods is isolate PG311-1.6. This event was originally classified as crossover associated with a conversion tract that extended from SGD coordinate 31694 to 63936. SNP analysis demonstrated that the white sector had a terminal deletion on chromosome V, beginning near coordinate 62000. The same sector also had a large terminal duplication on chromosome VII. Although this rearrangement has not been fully characterized, since there are a cluster of delta elements near the breakpoint on chromosome V and Ty elements at the breakpoint on chromosome VII, it is possible that the strain has a chromosome V-VII translocation, similar to those that we have characterized previously (ARGUESO et al., 2008). No alterations were detected on either chromosome V or VII in the red sector. Sequence analysis indicated that the red sector retained the SUP4-o gene. It is possible that the cell that gave rise to the red sector lost the prion PSI, that affects the efficiency of ochre suppressors (SHKUNDINA and TER-AVANESYAN 2007), although other possibilities cannot be excluded. Whatever the details of the genetic alterations in PG311-1.6, the event does not represent a conventional allelic crossover on chromosome V and, therefore, is excluded from our analysis.
Of the 13 colonies with spontaneous reciprocal recombination events analyzed by SNP arrays, the numbers of colonies of various classes were: 1) two crossovers without detectable conversions, 2) two crossovers with 3:1 conversion events, 3) one crossover with a 0:4 conversion, 4) one crossover associated with a hybrid tract (1:3 and 0:4 segments), and 5) seven crossovers with complex conversion tracts (Table 2). The complex conversion tracts will be discussed further below.

The locations of the spontaneous crossovers and associated conversion tracts are shown in Fig. 5A. Each sectored colony is depicted as a pair of lines with the upper line representing the red sector and the lower line representing the white sector. The red and black line segments indicate that the sector is homozygous for the W303a-associated SNPs and the YJM789-related SNPs, respectively. The green line segment indicates that the sector is heterozygous for the SNPs. We show the two chromosomes within each sector as a single line because our analysis does not allow us to determine the coupling relationships for heterozygous SNPs between the two homologues. The median length of all crossover-associated conversion tracts was 6.1 kb, similar to the median observed in our previous study of spontaneous conversion tracts (6.5 kb, (Lee et al. 2009). Only one unselected LOH was observed in unirradiated cells. Both sectors in PG311-7B had a gene conversion event on chromosome VIII unassociated with a crossover (Table S6). Thus, the frequency of spontaneous unselected LOH events/cell is very low (about 0.08) as expected.

For the events shown in Fig. 5A, 3:1 conversion events could reflect an initiating DNA lesion occurring anywhere within the tract, since events can be propagated bidirectionally from the DSB (Tang et al. 2011). For 4:0 or 3:1/4:0 hybrid tracts, the initiating lesion presumably occurs within the 4:0 region of the tract (Fig. 3). Although we do not see any
strong hotspots for spontaneous events with this limited dataset, in a larger sample, we
found that the region between SGD coordinates 41,000 and 60,000 had a significantly
elevated level of crossovers and the region near CEN5 had a significantly reduced level of
events (Lee et al. 2009).

As shown in Fig. 5A, many of the recombination events are associated with multiple
transitions between heterozygosity and homozygosity. In Table S6, for each sectored
colony, we assigned a letter to represent each transition point; for each transition point, we
also show two SGD coordinates, indicating the positions of the closest SNPs on the array
that flank the transition points. The simplest events (crossovers without gene conversion,
6B and 1.7 in Fig. 5A) have a single transition point. The recombination event shown in Fig.
4 (which corresponds to event 2A in Fig. 5A) has two transition points at different
positions, one in the red sector and one in the white sector. In contrast to these relatively
simple events, the sectored colony 18A (Fig. 5A) has six transitions, one in the red sector
and five in the white sector. In our analysis, if the transition point is identical in both
sectors, it is only counted once. In Table S6, we also assign a class (A to L) for all events. In
Tables S1 and S2, each class of event is diagrammed using the same approach employed in
Fig. 5A. In these figures, we also indicate the number of events in each class, and which
supplementary figure (Figs. S1-S40) shows the pattern of DNA repair events consistent
with the specific conversion event. The same methods are used to describe the
recombination events induced by DNA damage as were used to describe the spontaneous
events. Multiple transitions within conversion tracts could reflect “patchy” repair of
mismatches within long heteroduplexes (discussed further below) or template switching
between homologues. For the spontaneous conversion events in Fig. 5A, we did not find a correlation between SNP density and the number of transitions within the tract ($r^2 = 0$).

**Analysis of LOH events by SNP arrays in cells treated with γ rays:** We analyzed PG311 sectors that were generated in a previous study (Lee and Petes 2010) by treatment of G1-arrested cells with 100 Gy of γ radiation, followed by selection of red/white sectored colonies on canavanine-containing plates; this dose of γ rays elevated the frequency of sectoring about 26-fold. All of the colonies examined had a crossover on chromosome V. We analyzed seven of these sectored colonies with SNP microarrays, and two of these were also examined by HTS. The SNP array data are shown in Table S7 with depiction of the recombination events in Tables S1 and S2.

The positions of the selected crossovers and associated gene conversion events in the CEN5-can1-100/SUP4-o interval are shown in Fig. 5B. Our mapping of these events by SNP arrays is in reasonably good agreement with our PCR-based mapping method (Lee and Petes 2010). All of the conversion events had at least one SNP that was homozygous on both sides of the sector (4:0 conversion) as expected if the recombination events were a consequence of repair of two sister chromatids broken at the same position (Figs. 2 and 3).

In addition to the selected events, from our genome-wide analysis, we observed 17 unselected events on other chromosomes among the seven colonies: four crossovers associated with conversion (Table 2), eleven conversions that were not associated with crossovers, and two break-induced replication (BIR) events (Table 3). Since the frequency of unselected crossovers in unirradiated samples is very low (less than 0.1/cell), it is likely that all of the events in the irradiated cells reflect the repair of γ ray-induced DNA damage. The locations of these unselected events are shown as blue symbols in Fig. 6. The events
appear randomly distributed in this small data set. The SNP arrays for radiation-induced unselected crossovers and associated conversions have patterns similar to the selected crossover shown in Fig. 4. In addition, we observed many conversion events unassociated with crossovers (Fig. 7). This figure shows at low and high resolution a 0:4 conversion event in which both sectors have gained YJM789 SNPs and lost W303a SNPs. This pattern could represent an event that occurred prior to radiation. However, since such events were observed commonly in irradiated cells but not in control diploids, we assume that most (or all) were induced by γ rays.

Since the red and white sectors are produced by the two daughter cells resulting from the division of radiation treatment of a G1-synchronized cell, the analysis of genomic DNA from both sectors gives valuable mechanistic information even for unselected events. For example, if we observed an interstitial LOH event by examining only the white sector, we would not know whether this event was a consequence of a 3:1 conversion, a 4:0 conversion, or a two-strand double crossover. This ambiguity can be resolved by examining genomic DNA from the red sector.

We observed two sectored colonies that had BIR events. In single BIR events (such as Class L2 in Table S1), one sector has an LOH event that extends from an internal site on the chromosome to the telomere, whereas the other sector is heterozygous for the same SNPs. In double BIR events, both sectors have LOH events extending from an internal site to the telomere (Class L1 in Table S1). Interestingly, in the colony with the single BIR event, there is a conversion event on the chromosome that was originally the sister chromatid of the one involved in BIR. This result argues that both sister chromatids had DSBs at...
approximately the same position. The molecular interactions required to produce Classes L1 and L2 are shown in Figs. S39 and S40.

All of the selected and unselected reciprocal crossovers induced by γ rays were associated with conversion tracts. The median length of all conversion tracts (both associated and unassociated with crossover) was 12.9 kb (95% confidence limits of 5.2-20.4 kb). The median lengths of conversion tracts associated and unassociated with crossovers were 18.4 kb (10.8-25.3) and 8.4 kb (2.6-13.3), respectively. By the Mann-Whitney test, the median lengths of crossover-associated and crossover-unassociated γ ray-induced conversion tracts were significantly different (p=0.01).

**Analysis of LOH events by SNP arrays in cells treated with UV:** G1-synchronized PG311 cells were treated with a UV dose of 10-15 J/m². This dose resulted in no significant loss of viability but stimulated the frequency of sectors by about 1000-fold. We examined three sectored colonies by SNP arrays and two of these colonies were also analyzed by HTS. In addition to the selected crossover on chromosome V, among the three colonies, there were seven unselected crossovers, 33 unselected conversion events, and one BIR event (Tables 2 and 3). Thus, there were about fourteen unselected LOH events per UV-treated cell. The locations of selected chromosome V events and the unselected LOH events are shown in Figs. 5 and 6, respectively. The UV-induced LOH events are distributed fairly evenly throughout the genome (Fig. 6). As observed for the γ ray-induced BIR events, the UV-induced BIR event is located close to the telomere. The detailed information about breakpoints in UV-treated cells is shown in Table S9 and Tables S1 and S2.

All of the crossover events had an associated conversion. In most of the conversion events, there was at least one SNP with the 4:0 or 0:4 pattern, suggesting that UV-induced
damage in G1 may result in DSBs (Fig. 2B). The median length of all UV-induced conversion tracts was 9.2 kb (6.5-10.3), whereas the median lengths of conversion tracts associated and unassociated with crossovers were 10.3 (7.0-18.9) and 7.5 (4.5-10.2) kb, respectively. Although the crossover-associated conversions are longer than the crossover-unassociated conversions, this difference is not significant (p=0.08 by Mann-Whitney test).

**HTS analysis of LOH in γ-ray- and UV-treated cells:** Since G1-arrested yeast cells treated with 100 Gy of ionizing radiation have about 35 DSBs per cell (Lee and Petes 2010), the average number of unselected LOH events per cell (two to three) indicates most events must be repaired by mechanisms that do not generate LOH. An alternative possibility is that a substantial fraction of the events have short conversion tracts that are not detectable by the SNP arrays. Since HTS can detect LOH events for all of the 55,000 SNPs existing in the diploid strain, rather than the 13,000 SNPs represented on the SNP array, we sequenced genomic DNA samples from both red and white sectors of two sectored colonies of γ-ray-treated samples (PG311-GR-37R/W and PG311-GR-40R/W) and two colonies of UV-treated samples (PG311-UV-8R/W and PG311-UV-9R/W). The details of this analysis are described in Supporting Information. All of the LOH events detected by SNP arrays were also found by HTS. LOH events that had not been previously detected by the SNP arrays were confirmed by re-sequencing the relevant PCR fragment. The patterns of LOH as identified by HTS in the γ-ray- and UV-treated samples are in Tables S8 and S10.

Figure 8 shows a comparison between SNP microarray and HTS data for an LOH event on chromosome II in a UV-treated sample (PG311-UV-8R). The SNP microarray indicates that a transition between heterozygous and homozygous SNPs occurs somewhere between SGD coordinates 450919 and 452926, whereas the HTS data refine the mapping of this
transition between SGD coordinates 451337 and 451581. Additionally, the HTS data showed that the recombination event was more complex than indicated by the microarray data. By HTS, we found that the region between coordinates 448628 and 450279 had undergone LOH; this event was not detected by microarrays because there were no oligonucleotides between 448488 and 450919 on the microarray. A summary of the comparison of data from SNP arrays and HTS for the same samples is given in Table 4.

Although more LOH events were observed with HTS than with SNP arrays, the difference was not large. For example, in the two UV-treated samples, we observed 32 LOH events by SNP arrays and 35 events by sequencing. In the γ ray-treated samples, we found five events by SNP arrays and six events by sequencing. Since more than 80% of the events detected by HTS were also detected by microarrays, it is unlikely that our estimates of LOH events are substantially affected by a high frequency of gene conversion events with short conversion tracts. We cannot rule out, however, the possibility of gene conversion events with very short (<100 bp) tracts. As shown in Table 4, a number of the gene conversion tracts analyzed by HTS were more complex than the same tracts examined by the SNP arrays. The frequencies of complex tracts, as determined by HTS and SNP microarrays, were 0.37 and 0.22, respectively. Despite the differences in the numbers and types of LOH events detected by HTS and the SNP microarrays, it is clear that most of the LOH events are detectable by the SNP microarrays.

**HTS analysis of mutations induced in γ ray- and UV-treated cells:** About 99.5% of the bases between the two homologues of PG311 are identical. The HTS data generated for the same colonies examined for LOH events were analyzed for radiation-induced mutations. We analyzed both sectors of two sectored colonies induced by UV and two
sectored colonies induced by γ rays for new mutations (Table S11). There were three and twelve de novo point mutations detected in the γ ray- and UV-treated samples, respectively.

All three of the mutations induced by γ rays and six of the twelve mutations induced by UV were in both red and white sectors of the colony. The presence of the mutation in both sectors indicates that the mutation induced in G1 by the radiation was represented in both strands of the duplex prior to replication. The mutations in the UV-treated cells that were present in only one sector could reflect a mutant base in only one of the two strands. These two types of events have been observed previously in UV-treated cells (Eckardt and Haynes 1977; James and Kilbey 1977). Nine of twelve of the UV-induced mutations and two of three of the γ ray-induced mutations were transitions. In much more extensive study of spontaneous and UV-induced mutations at the SUP4-o locus (Kunz et al. 1987), spontaneous mutations had a ratio of transitions:transversions of 4:6, whereas UV-induced mutations were biased toward transitions (4:1).

Most (11 of 15) of the induced point mutations were located within genes rather than between genes (Table S11). By Chi-square analysis, the distribution of mutations throughout the genome is non-random (p = 0.002). Five of the fifteen mutations are located on the left arm of XI and, remarkably, two mutations (one induced by γ rays and one induced by UV) are within the NUP120 gene. By similar methods used to detect new base substitutions, we failed to detect any insertion/deletion mutations (in/dels) in the eight sequenced samples. It should be pointed out, however, that detection of in/dels in HTS data with short-paired reads is challenging, particularly in a diploid that is heterozygous for many pre-existing in/dels.
DISCUSSION

In this study, we mapped both selected and unselected mitotic recombination events in a genome-wide analysis. Most of the events were mapped using only SNP microarrays, but four events were examined by both SNP microarrays and HTS. To our knowledge, this study is the first to examine spontaneous and DNA damage-induced LOH events throughout the yeast genome. The conclusions from this study are: 1) the gene conversion tracts analyzed by SNP arrays and HTS were often more complex than inferred from our earlier lower-resolution mapping studies (Lee et al. 2009; Lee and Petes 2010), 2) doses of radiation that result in little or no loss of cell viability in G1-synchronized diploid cells resulted in multiple unselected LOH events, and 3) the same doses of γ rays and UV used in the LOH study result in very low levels of de novo mutations. In addition, we conclude that, although HTS has four-fold better resolution than SNP microarrays, the SNP arrays detect most of the same events identified by HTS. These conclusions will be further discussed below.

**Lengths of gene conversion tracts:** The median lengths (95% confidence limits shown in parentheses), as measured by SNP arrays, of gene conversion tracts associated with crossovers for spontaneous, UV-induced, and γ ray-induced events were 6.1, 10.3, and 18.5 kb, respectively. As we have observed previously (Lee et al. 2009), mitotic gene conversion tracts are substantially longer than meiotic conversion tracts (Mancera et al. 2008). In addition, for the γ ray-induced conversion events, the conversion tracts associated with crossovers were significantly longer than those unassociated with crossovers as has been observed previously (Aguilera and Klein 1989; Mancera et al. 2008). One simple explanation of this observation is that conversion events unassociated
with crossovers usually are a consequence of SDSA, and such events might involve limited processing of the broken chromosome ends. In contrast, crossovers likely proceed through formation of a double Holiday junction. Formation of this dHJ intermediate may require more extensive processing of broken DNA ends and/or more extensive DNA synthesis primed from the invading end. It is also possible that branch migration of the dHJ intermediate could extend the length of the heteroduplex associated with the crossover; this possibility will be further discussed below.

**Recombinogenic DNA damage:** Although it is clear from a variety of studies that DSBs stimulate mitotic recombination, the DNA lesion responsible for initiating spontaneous recombination events is not certain. We previously showed that about half of mitotic crossovers on chromosome V are associated with a gene conversion tracts that are exclusively 4:0 or 0:4 or that have a region of 4:0 or 0:4 (hybrid tracts). Such conversion tracts indicate that both sister chromatids have breaks at approximately the same position and one simple mechanism consistent with this property is that these spontaneous conversion events reflect a DSB formed in G1 of the cell cycle. Supporting this conclusion, many (44%) of the conversion events induced by γ rays in G1 of the cell cycle have regions of 4:0 or 0:4 segregation, whereas none of the conversion events induced by γ rays in G2 had this pattern (LEE and PETES 2010). Among the mechanisms that could produce the spontaneous lesions required to initiate recombination are: 1) closely-opposed nicks generated during removal of adducts caused by oxidative DNA damage, 2) DSBs caused by Top2p or other cellular endonucleases, 3) lesions resulting from collisions between replication forks and the transcription machinery, 4) DSBs resulting from the collision of converging replication forks, and 5) nuclease-dependent processing of secondary DNA
structures ("hairpins" and cruciforms). Thus far, we have been unable to associate the spontaneous recombination events with any of these mechanisms. For example, the two positions that represent the convergence of replication forks on the left arm of chromosome V (FACHINETTI et al. 2010) are not hotspots for recombination in our limited dataset. If there are several different mechanisms that can produce recombinogenic DNA lesions, we will need to map many events to detect significant associations.

In our current analysis of γ ray-induced gene conversion events in G1-synchronized cells by SNP microarrays, we found that ten of the eleven conversion tracts associated with crossovers had a 4:0 or 0:4 segment, and eight of the eleven conversion tracts that were unassociated with crossovers had such a segment. This observation is consistent with the possibility that most of the observed recombination events in IR-treated cells reflected a DSB introduced by γ rays in G1.

The recombinogenic effect of UV-induced DNA damage is less clear. One possibility is that small gaps resulting from the removal of UV-induced dimers are the recombinogenic lesion. Galli and Schiestl (GALLI and SCHIESTL 1999) found that UV did not stimulate mitotic recombination between direct repeats (single-strand annealing) in G1-arrested cells unless the cells were allowed to enter the S-period after the UV treatment. In contrast, when G1-arrested cells were treated with IR, stimulation of single-strand annealing was observed without requiring the cells to enter S. If the recombinogenic DNA lesion is a DSB, the likely explanation of the different results is that IR directly creates DSBs whereas the repair of UV lesions results in nicks that result in DSBs when the nicked molecule is replicated (GALLI and SCHIESTL 1999). By this explanation, it is surprising that many of the conversion events induced by UV in G1 in our experiments had 4:0 or 0:4 segments, suggesting that these
exchanges were a consequence of a G1-stimulated DSB. Such a DSB could be generated in G1 if the removal of dimers on opposite DNA strands resulted in a very short (<10 bp) duplex region separating the 30 bp gaps. Based on previous estimates of the number of dimers induced by 20 J/m² of UV (DAIGAKU et al. 2010), we calculate that there are about 7500 dimers/diploid genome induced by a UV dose of 15 J/m² and, based on a Poisson distribution, there would be about 35 regions per genome in which two dimers are on opposite strands within 75 base pairs of each other. Since the DSB would require two closely-opposed gaps rather than two closely-opposed dimers, the kinetics of gap formation and gap repair affect the probability of DSB formation by this mechanism. Another complicating factor is that the frequency of closely-opposed dimers is higher than expected if dimer formation is random (REYNOLDS 1987).

A second possibility is that DSB formation is initiated by gaps on opposite strands that are relatively close together (≤ 500 bp apart), but too far apart to generate a G1 DSB. If a DNA molecule with such gaps is replicated, the product would be two sister chromatids with DSBs located 500 bp or less apart. Processing of the broken ends to yield single-stranded DNA regions 500 bp or greater would preclude formation of a dHJ involving the two sister chromatids. Thus, such molecules would likely be repaired using the intact homologue as the template, generating a 4:0 conversion. This model is consistent with the GALLI and SCHIESTL interpretation. We calculate that cells irradiated with 15 J/m² would have about 234 dimers on opposite strands within 500 bp of each other.

An alternative possibility is that recombination events are a consequence of DSBs occurring at replication forks stalled at unexcised pyrimidine dimers. Unrepaired UV-induced damage has been demonstrated to block replication forks and replication of such
damaged templates promotes sister-chromatid recombination (BRANZEI and FOIANI 2010). Although we cannot exclude this model, the observed UV-induced 4:0 events would require that the replication fork stalled at the dimer result in two broken sister chromatids, perhaps by increasing the probability of a replication fork collision. It should be emphasized that although UV very efficiently stimulates crossovers between homologues, most of the UV-induced recombination events are likely to represent sister-chromatid interactions (KADYK and HARTWELL 1992).

**Relationship between mitotic gene conversion and crossovers:** In meiosis in yeast, about half of conversion events are associated with crossovers (MANCERA et al. 2008; PETES et al. 1991). In our previous mitotic studies, we selected crossovers and found that most (>80%) of these events were associated with an adjacent tract of gene conversion (LEE et al. 2009; LEE and PETES 2010); conversion events unassociated with crossovers could not be selected with our system. In the current study, for unselected events stimulated by radiation, we can estimate the fraction of conversion events that are associated with crossovers.

For IR-treated cells, including all non-selected events except BIR, we found four conversions associated with crossovers and eleven conversions unassociated with crossovers (Tables 2 and 3). In these experiments, we detect the associated crossover because it generates LOH from the conversion tract to the end of the chromosome. As shown in Fig. 2, depending on the pattern of chromosome segregation, we expect that only half of the crossovers will lead to LOH of markers distal to the point of exchange, and this expectation has experimental support (CHUA and JINKS-ROBERTSON 1991). In addition, as discussed in the Supporting Information and Table S12, we found preliminary evidence in
our experiments for conversion events associated with crossovers that did not result in LOH.

Thus, we calculate that of the fifteen conversion events induced by γ rays, it is likely that eight were associated with crossovers (53% association). Similarly, among UV-induced recombinants, since there were seven unselected conversions associated with crossovers and 33 unselected conversions not associated with LOH (Tables 2 and 3), we calculate that about 35% of the UV-induced conversion events are associated with crossovers. Our conclusion that the frequency of crossovers associated with conversions is not very different for mitotic and meiotic conversion events is consistent with other recent studies (Ho et al. 2010). In yeast studies in which conversion events are limited in size, the association between conversion and crossovers is weaker (Pâques and Haber 1999). Also, in Drosophila and mammalian cells, conversion events are only rarely associated with crossovers (Andersen and Sekelsky 2010).

**Complex gene conversion tracts and BIR events:** Previously, we classified conversion tracts as “simple” if the markers within the tract had one of the following patterns: 1) all markers were 3:1 or 1:3 (not mixtures of 3:1 and 1:3 in same tract), 2) all markers were 4:0 or 0:4, or 3) hybrid tracts of the form 3:1/4:0, 1:3/0:4, 3:1/4:0/3:1, or 1:3/0:4/1:3. All such tracts can be explained as the consequence of the repair of one or two broken chromatids by one of the HR pathways shown in Fig. 1. There were, however, conversion tracts that were more complicated (termed “complex tracts”), which will be described below. In the UV-treated samples, six of ten of the crossover-associated conversion tracts were complex, although only three of thirty-three tracts were complex in conversions unassociated with crossovers (Tables 2 and 3); this difference is significant.
(p<0.01) by the Fisher exact test. In the IR-treated samples, the conversion events associated with crossovers were usually more complex than those that were not (Table 2), although the difference was not significant. Mancera et al. (Mancera et al. 2008) reported that 11% of meiotic crossovers had complex conversion tracts, whereas the frequency of complex tracts among conversions unassociated with crossovers was 3%. One explanation of this difference could be that crossovers that proceed through the pathway shown in Fig. 1B are associated with two regions of heteroduplex, while conversions resulting from SDSA or dHJ dissolution have only a single region of heteroduplex (Fig. 1A and 1C). Second, because gene conversion tracts associated with crossovers are usually longer than those unassociated with crossovers, there may be a greater chance to observe patchy repair of mismatches (as defined below) in tracts associated with crossovers.

Diagrams of all recombination events in our study are shown in Tables S1 and S2, and figures showing the patterns of DSB repair required to produce the recombination events are shown in Figs. S1-S40. Most of the complex conversion tracts could be divided into two categories: those that had multiple transitions between 3:1, 4:0, and heterozygosity within the tract and those tracts in which 3:1 and 1:3 or 4:0 and 0:4 segments occurred within one tract. Examples of conversion tracts with multiple transitions are strains 18A (Class J9, Table S1) and 4.1 (Class J8, Table S1); both 18A and 4.1 are also depicted in Fig. 5A. The complex tract in 4.1 is consistent with the repair of two DSBs with “patchy” repair of mismatches in two of the resulting heteroduplexes (Fig. S29). Heteroduplexes will often contain multiple mismatches that can be repaired to produce either a conversion event or a restoration event (Kirkpatrick et al. 1998). For example, in Fig. 1A, repair of the heteroduplex resulting in a duplex with two “red” strands would represent conversion-
type repair, since this pattern produces 3:1 segregation; repair of the mismatch to produce a duplex with two “blue” strands represents restoration-type repair, since this pattern generates two cells that retain heterozygosity at the position of the original heteroduplex. Although multiple mismatches within one heteroduplex are generally converted in a concerted manner yielding a continuous conversion tract, tracts with mixtures of conversion-type and restoration-type repair have been detected in both meiosis (Mancera et al. 2008; Symington and Petes 1988) and mitosis (Mitchel et al. 2010; Nickoloff et al. 1999).

The pathway of DSB repair shown to explain the pattern of markers in the strain 18A conversion event (Fig. S30) invokes patchy repair and branch migration of the dHJ. During recombination in *E. coli*, a Holliday junction can be translocated by branch migration, resulting in symmetric heteroduplexes (West 1997). Although genetic evidence argues against the formation of symmetric heteroduplexes during meiotic recombination in *S. cerevisiae* (Petes et al. 1991), symmetric heteroduplexes have been invoked previously to explain certain classes of mitotic gene conversions (Esposito 1978; Nickoloff et al. 1999; Roitgrund et al. 1993). Branch migration can also generate patterns of repair in which a single DSB can produce both 3:1 and 1:3, or 4:0 and 0:4 events as shown in Fig. S10B.

Our data do not allow us to determine unambiguously the pathways required to generate the observed conversion tracts. However, we can state that many of the complex tracts are inconsistent with the simplest form of the recombination models shown in Fig. 1. In particular, it is likely that patchy repair of mismatches is a relatively common feature of mitotic gene conversion tracts. A detailed discussion of all of the conversion tracts in our studies is given in the Supporting Information.
There were three unselected BIR events observed in our study (Class L, Fig. S39 and S40). For two of the three events, we observed a region of conversion associated with the BIR event. This pattern is consistent with the repair of two DSBs, one by SDSA and one by BIR (Fig. S40). The BIR events were about three-fold less frequent than unselected crossovers, as expected from previous studies (Ho et al. 2010; McMurray and Gottschling 2003)

**Relationship between the level of DNA damage and the frequency of LOH events:**

The 100 Gy dose of IR used in our experiments is expected to produce about 35 DSBs/diploid genome (Lee and Petes 2010). Since we observed only 2.4 LOH events/irradiated cell, most of these DSBs must be repaired by a mechanism that does not produce a detectable LOH event. Since the cells in our experiments were irradiated in G1, the DSBs must have been repaired either by an interaction with the homologous chromosome or by NHEJ. We suggest several possible explanations. First, it is possible that the repair of the DSB frequently involves an interaction with the homologue that is associated with a very short conversion tract. Tracts shorter than 50 bp would be rarely detected, even by HTS. Such a repair event would likely involve very limited processing of broken DNA ends as well as short excision repair tracts. A system of short-patch (often less than 12 base pairs) mismatch repair that is independent of the classical mismatch repair system in *S. cerevisiae* was described by Coic et al. (Coic et al. 2000), although the genes involved in this type of repair have not been identified. In addition, conversion tracts less than 53 bp have been detected among HO-induced events (Palmer et al. 2003). A related possibility is that gene conversion events occur non-randomly in regions of the genome that are not represented on our microarrays (regions that are identical between W303a...
and YJM789 or regions with repeated genes). A third possibility is that the repair of the DSB is associated with restoration-type repair of mismatches within the heteroduplexes. Since most of the crossovers in our study are associated with detectable gene conversion tracts, we would have to hypothesize that conversion events that are not associated with crossovers are much more prone to restoration-type repair than conversion events that are associated with crossovers. A fourth possibility is that the IR-induced DSBs are frequently (Daley et al. 2005) repaired by NHEJ events. Although NHEJ events are repressed in MATα/MATα diploids, since PG311 lacks the MATα locus, NHEJ events will occur. Although NHEJ events will not produce LOH, depending on the nature of DNA ends (compatible single-strand overhangs or blunt), some NHEJ events would be expected to result in loss or gain of a few base pairs. Although we did not observe in/dels in our HTS analysis, this observation does not rule out the possibility that some repair events reflect NHEJ. It is also possible, of course, that all four possibilities described above are partly responsible for the “missing” LOH events.

One explanation that we can exclude as a major contributor to the discrepancy between the number of lesions and the number of LOH events is chromosome loss. Chromosome loss can be readily detected by the SNP microarrays, and no losses were observed in cells treated with γ rays or UV. In experiments in which eight-fold higher doses of γ rays were used, about 10% of the treated yeast cells had chromosome loss (Argueso et al. 2008).

Although we detected more than 50 unselected LOH events in cells treated with γ rays and UV, no duplications or deletions were detected. Thus, SNP arrays that can detect both LOH and changes in copy number are a much more efficient method of detecting
recombinogenic DNA lesions than comparative genomic hybridization (CGH) arrays. In our previous analysis of γ ray-treated diploid cells by CGH (ARGUESO et al. 2008), we found that most of the irradiated cells had one or more chromosome rearrangements, usually non-reciprocal translocations with retrotransposons at the breakpoints. In these experiments, we treated G2-synchronized cells with doses of radiation that were eight-fold higher than the doses used in our current study.

**Mutations induced by γ rays and UV:** We found only a few mutations induced by γ rays and UV, averages of 1.5 and 6 mutations/irradiated cell, respectively. Although there are no genome-wide studies of the frequencies of mutations induced by γ rays, extrapolating from the frequency of induction of X-ray-induced mutations at the *CAN1* locus (Goeke and Manney 1979) and the rate of spontaneous mutations per base pair at *CAN1* (Lang and Murray 2008), we calculate that the expected frequency of mutations per genome is about two/diploid cell, close to our observed number. The most direct comparison for the UV-induced mutations is with data obtained from HTS of UV-treated stationary-phase haploid yeast strains (Burch et al. 2011). These strains in these studies had a temperature-sensitive mutation in *CDC13*. However, three of the sequenced isolates were treated at the permissive temperature. By extrapolating their data to our UV dose, we would expect about 14 mutations/diploid cell, only two-fold different from our observed frequencies. In summary, our HTS data detected roughly the expected number of mutations per irradiated strain.

As described in the Results, the mutations induced by UV and γ rays are non-randomly distributed among the yeast chromosomes. Although this non-random distribution needs to be verified with a large dataset, it is possible that the mutagenic DNA damage is
distributed non-randomly because of the specific position of different chromosomes within the nucleus or chromosome-specific chromatin domains. Since the UV-irradiated strains have about 7500 DNA lesions (as discussed above), the vast majority of these lesions must be repaired by nucleotide excision repair in a manner that does not result in LOH or mutations.

We assume that most of the UV-induced mutations reflect errors introduced during the bypass of pyrimidine dimers by Rev1p and Polζ, since 90% of UV-induced mutations require these activities (Lawrence 2002). The source of the mutations in the γ-ray-treated samples is less clear. Since the mutations are not associated with regions of LOH, the mutations probably do not reflect errors introduced by DSB repair. It is possible that bases damaged by γ rays are bypassed by error-prone polymerases by a mechanism similar to that associated with UV-induced DNA damage.

All three of the mutations introduced by γ irradiation and about half of the mutations caused by UV were found in both sectors of sectored colonies. This result indicates that the introduced mutation was placed into both strands of the duplex before DNA replication. Such events, which have been observed previously for UV-induced DNA damage (Eckardt and Haynes 1977; James and Kilbey 1977), have been termed “two-strand” mutations (Abdulovic et al. 2006). One model for such events is that they reflect the repair of two closely-opposed DNA lesions by nucleotide excision repair (NER). During the repair of one lesion, a mutation is introduced. The repair of the second lesion on the opposite strand results in a gap that includes the mutant substitution, and filling-in of the gap results in mutant substitutions in both strands of the duplex. Whatever the explanation of two-strand events, both UV and γ rays efficiently produce this type of mutation.
The repair of DSBs is associated with a 100-fold elevation in the frequency of reversion of a closely-linked mutation (Strathern et al. 1995), and approximately 1000-fold elevated rates of mutation have been observed during BIR (Deem et al. 2011) and other gene conversion events that result in two newly-synthesized strands (Hicks et al. 2010). In addition, the frequency of UV-induced mutagenesis is elevated more than 100-fold in regions of single-stranded DNA next to DSBs or abnormal telomeres (Yang et al. 2008).

Based on these observations, we checked whether the de novo mutations were non-randomly associated with LOH regions associated with gene conversion or BIR. There were 15 base substitutions observed among four sectored colonies resulting from the irradiation of G1 cells. The total lengths of the unselected gene conversion and BIR events among these strains were 163 kb (PG311-UV-8R/W), 271 kb (PG311-UV-9R/W), 22 kb (PG311-IR-37R/W), and 50 kb (PG311-IR-40R/W). The fraction of the genome with these LOH regions was about 0.01. Two of the fifteen (0.13) mutations were in regions of LOH. Although this calculation suggests that the LOH regions may have a significantly elevated frequency of mutations, most of the induced mutations are located outside of the LOH regions.

Comparison among methods of physically mapping recombination events: In our previous studies, we mapped recombination events by a PCR-based technique (described in the Introduction). As employed in our analysis of chromosome V events, this approach was time-consuming and expensive, and mapped events with relatively poor resolution (about 4 kb). More importantly, this method could not be easily used to map events throughout the genome. In addition, the PCR-based approach did not allow us to examine changes in gene dosage (deletions or duplications). For example, we found that an event
classified as a crossover on chromosome V by the PCR-based method was actually a terminal deletion on V when examined by SNP arrays.

In contrast, both SNP arrays and HTS allow analysis of events throughout the genome. The advantages of SNP arrays compared to HTS are: 1) relatively low cost (about $100/sample), 2) speed of analysis (about four hours for SNP arrays versus a week for HTS), and 3) relative ease in detecting changes in gene dosage. The major advantages of HTS are: 1) higher resolution (1 kb for SNP arrays versus 250 bp for HTS), and 2) the ability to detect de novo mutations. In addition, diagnosis of LOH by HTS can be done with any diploid in which the progenitor haploid strains have been sequenced, whereas diagnosis of LOH by SNP arrays requires the construction of strain-specific microarrays. Although SNP arrays are probably a more cost-effective and faster approach for mapping large numbers of recombination events at present, as HTS becomes cheaper and analysis of HTS data becomes faster, HTS is likely to be the method of choice in the future. Neither SNP microarrays nor HTS, however, can map recombination events that do not involve LOH (for example, sister chromatid exchanges).

**Summary:** In conclusion, we have used SNP microarrays and HTS to map crossovers and gene conversion events at high resolution throughout the yeast genome. These studies represent the first genome-wide measurement of the number and types of unselected LOH events induced by UV and γ rays. In G1-synchronized cells treated with either UV or γ rays, 4:0 conversion events are common, suggesting that many of the LOH events reflect the repair of two sister chromatids broken at approximately the same position. In addition, the high-resolution analysis of recombination events by SNP arrays and HTS reveals that gene
conversion tracts, particularly those associated with crossovers, are more complex than was previously recognized by low-resolution studies.
ACKNOWLEDGMENTS

We would like to thank members of the PETES and JINKS-ROBERTSON labs for discussions, and LUCAS ARGUESO, KAT MITCHEL, SUE JINKS-ROBERTSON, and WEI SONG for comments on the manuscript. We thank members of J. McCUSKER’s lab for technical help, and L. ARGUESO and J. RINE for useful suggestions. The research was supported by NIH grants GM24110 (TDP), GM52319 (TDP and PM), 5RC1ES18091 (TDP and PM), and funds from the Duke University School of Medicine.
REFERENCES


FACHINETTI, D., R. BERMEJO, A. COCITO, S. MINARDI, Y. KATOU et al., 2010 Replication termination at eukaryotic chromosomes is mediated by Top2 and occurs at genomic loci containing pausing elements. Mol Cell 39: 595-605.


MCCULLEY, J. L., and T. D. PETES, 2010 Chromosome rearrangements and aneuploidy in yeast strains lacking both Tel1p and Mec1p reflect deficiencies in two different mechanisms. Proc Natl Acad Sci U S A 107: 11465-11470.


REYNOLDS, R. J., 1987 Induction and repair of closely opposed pyrimidine dimers in *Saccharomyces cerevisiae*. Mutat Res 184: 197-207.


**Figure legends**

Figure 1. Pathways of DSB repair by homologous recombination.

In this figure, we show accepted models of DSB repair by homologous recombination. DNA strands from two different homologues are shown in red and blue; light red and blue lines indicate newly-synthesized DNA. Regions of the duplex that have strands of different colors represent heteroduplexes. These pathways are described in detail in the text.

A. Synthesis-dependent strand annealing (SDSA) pathway. Following processing of the DSB, the 3’ end of the left end of the broken DNA molecule invades the other duplex. Following DNA synthesis, the invading strand is displaced, and hybridizes to the right end of the broken chromosome. This pathway results in conversion events unassociated with crossovers.

B. Double-strand break repair (DSBR) pathway. In this pathway, a double Holliday junction (dHJ) is formed. Resolution of these junctions by resolvase cleavage can result in two different crossover products (middle panel) and two different non-crossover products (right panel). These products have two regions of heteroduplex located in trans. Alternatively, the dHJ can be dissolved by the action of topoisomerases/helicases resulting in a non-crossover product with heteroduplexes located in cis.

C. Break-induced replication (BIR) pathway. One of the broken ends invades the homologous chromosome, and duplicates sequences from the point of invasion to the telomere. The net result of BIR events is an apparent long terminal gene conversion event.

Figure 2. Genetic system used to select mitotic crossovers and associated conversions on the left arm of chromosome V.
The starting diploid strain PG311 has the ochre-suppressible *can1-100* on one copy of chromosome V (shown in red) and the *SUP4-o* gene (encoding an ochre suppressor tRNA gene) at an allelic position on the other homologue (shown in black). The strain is homozygous for the ochre-suppressible *ade2-1* allele. Strains with an unsuppressed *ade2-1* mutation form red colonies. The starting diploid strain is canavanine-sensitive and forms pink colonies.

A. Reciprocal crossover without an associated gene conversion initiated by a single DSB in G2. This type of event produces a canavanine-resistant red/white sectored colony (Barbera and Petes 2006). The transition from heterozygous markers to LOH is identical in the two sectors.

B. Reciprocal crossover with an associated conversion event initiated by a single DSB in G2. If a DSB forms on one of the black chromatids, a conversion associated with the crossover may occur. This event will also result in a canavanine-resistant red/white sectored colony in which the transitions between heterozygous markers and LOH are different in the two sectors. The region of conversion is indicated by the blue rectangle.

C. Reciprocal crossover and conversion resulting from a DSB formed in G1. A black chromosome with a DSB is replicated to form two sister chromatids that are broken at the same place. One chromatid is repaired to yield a reciprocal crossover and an associated conversion; the second is repaired to yield a conversion without a crossover. The resulting red and black sectors will have a 4:0 conversion event, a region in which both sectors are homozygous for SNPs derived from the red chromatid (included within the blue rectangle).
Figure 3. Production of hybrid conversion tracts by repair of two broken sister chromatids. The black chromosome is broken in G1 and replicated to yield two broken sister chromatids.

A. Production of a 3:1/4:0 hybrid tract. If the two DSBs are repaired to yield conversion tracts that have the same centromere-proximal boundary, but different centromere-distal boundaries, a 3:1/4:0 hybrid will be generated (shown in the blue rectangle).

B. Production of a 3:1/4:0/3:1 hybrid tract. If one conversion event is extended beyond the other at both the centromere-proximal and centromere-distal boundaries, a 3:1/4:0/3:1 tract will be formed.

Figure 4. Analysis of a spontaneous reciprocal crossover (PG311-2A) on the left arm of chromosome V by SNP microarrays. Most of the details concerning this figure are explained in the text. In brief, DNA samples isolated from the red and white sectors were labeled with one fluorescent nucleotide and DNA from a heterozygous control strain was labeled with a different fluorescent nucleotide. The samples were competitively hybridized to the SNP array and we measured the ratio of hybridization of the probes to SNP-specific oligonucleotides. The red and blue colors indicate hybridization to the W303a- and YJM789-specific oligonucleotides, respectively. CEN5 is located approximately at SGD coordinate 152000.

A. Low-resolution depiction of the samples derived from the red and white sectors. In the boxed region, the red sector has a region of LOH whereas the white sector is heterozygous at the same position. This pattern indicates a 3:1 conversion associated with
the crossover. Centromere-distal to the conversion event, the red and white sectors are homozygous for the W303a- and YJM789-specific SNPs, respectively.

B. High-resolution depiction of the samples derived from the red and white sectors. Each blue and red square represents hybridization to a single oligonucleotide on the array; the converted region is boxed.

Figure 5. Mapping of crossovers and associated conversion events on the left arm of chromosome V in untreated cells, and cells treated with γ rays or UV by SNP microarrays.

Each red/white sectored canavanine-resistant colony is given a number and is depicted as a pair of lines with the upper line representing the red sector and the lower line showing the white sector. The colored segments signify heterozygosity for the markers (green), homozygosity for the YJM789-derived SNPs (black), or homozygosity for the W303a-derived SNPs (red). The green arrows show that the markers are heterozygous from the position that the green segment begins to the end of the chromosome, and the red and black arrows indicate that the markers are homozygous for the W303a- or the YJM789-derived SNPs, respectively, from the point that the segment begins to the telomere of the chromosome. Internal regions of heterozygosity and homozygosity are shown as line segments without arrows and are drawn approximately to scale. The numbers at the top of the figure are SGD coordinates and the region between CEN5 and the can1-100/SUP4-o markers is about 120 kb in length.

A. Spontaneous crossovers and conversions.

B. γ ray-induced crossovers and conversions.

C. UV-induced crossovers and conversions.
Figure 6. Genomic locations of unselected recombination events and \textit{de novo} mutations in untreated cells and in cells treated with UV or $\gamma$ rays as determined by SNP microarrays and HTS.

The horizontal black bars depict each chromosome and are proportional to the chromosome length except for chromosome XII. The black circles depict the centromere of each chromosome. Short horizontal bars above each chromosome depict conversion events unassociated with crossovers and the length of each bar is approximately proportional to the length of the conversion tract. All conversion tracts are shown as single solid lines without regard to the complexity of the event (for example, transitions between 4:0 and 3:1). Single arrowheads depict reciprocal crossovers and double arrowheads depict BIR events. Asterisks located on the chromosome indicate the approximate positions of mutations induced by UV or $\gamma$ rays; two of the mutations (located at SGD coordinates 171529 and 301552 on X) are in regions of LOH. Events observed in untreated cells, cells treated with UV, and cells treated with $\gamma$ rays are shown in green, red, and blue, respectively. None of the events selected on the left arm of V are shown in this figure.

Figure 7. SNP array analysis of a gene conversion event unassociated with a crossover.

In cells treated with $\gamma$ rays, one of the canavanine-resistant red/white sectored colonies (43RW) had an unselected gene conversion event on chromosome IV. As shown at low- (Fig. 7A) and high- (Fig. 7B) resolution, both sectors had an LOH region in which YJM789-derived SNPs became homozygous (0:4 conversion event). The depiction of the SNP array data is the same as in Fig. 4. The length of the conversion tract is about 3 kb. \textit{CEN4} is located approximately at SGD coordinate 450000.

Figure 8. Analysis of the same recombination event by both SNP arrays and HTS.
This figure shows the analysis of the unselected recombination event on chromosome II in the red sector of the UV-induced sectored colony 8. Our standard SNP array analysis (top parts of Fig. 8A and 8B) showed a single transition between heterozygosity and homozygosity at about SGD coordinate 452000. In the bottom panels of the figure, we show HTS data for the same genomic sample. For the HTS data, the Y axis represents the frequency of YJM789 SNP (blue) or W303a SNP (red) “reads” for the experimental sample when assembled to the PSL2 genome. For heterozygous regions, there should be approximately equal frequencies of the two types of SNPs. It is clear from the high-resolution depictions of the HTS data that there is a short LOH region (boxed) located near SGD coordinate 450000 that was not detected by the SNP arrays. This region was not detected because oligonucleotides containing these SNPs were not present on the array. In the low-resolution depiction of the HTS data, within the LOH region, there is a small region near SGD coordinate 800000 in which SNPs appear to be heterozygous. These signals are artifacts based on “reads” from the repeated diverged MAL and MPH genes that were incorrectly mapped by the genome analysis software to chromosome II. CEN2 is located near SGD coordinate 238000.
<table>
<thead>
<tr>
<th>Strain name</th>
<th>PCR-based method</th>
<th>SNP microarrays</th>
</tr>
</thead>
<tbody>
<tr>
<td>PG311-1.4</td>
<td>133080</td>
<td>125754</td>
</tr>
<tr>
<td>PG311-1.7</td>
<td>151440</td>
<td>146855</td>
</tr>
<tr>
<td>PG311-4.1</td>
<td>99267</td>
<td>60163</td>
</tr>
<tr>
<td>PG311-4.11</td>
<td>133080</td>
<td>94329</td>
</tr>
</tbody>
</table>
TABLE 2. SUMMARY OF ALL CROSSTOVERS DIAGNOSED BY SNP MICROARRAY

<table>
<thead>
<tr>
<th>Types of reciprocal crossovers</th>
<th>Spontaneous</th>
<th>(\gamma)-ray</th>
<th>UV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Selected</td>
<td>Unselected</td>
<td>Selected</td>
</tr>
<tr>
<td>No detectable conversions</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3:1 or 1:3 conversions</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4:0 or 0:4 conversions</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Hybrid conversions&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Complex conversions</td>
<td>7</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td><strong>Total crossovers</strong></td>
<td><strong>13</strong></td>
<td><strong>0</strong></td>
<td><strong>7</strong></td>
</tr>
</tbody>
</table>

<sup>a</sup>In this table, we summarize data from selected crossovers and associated conversion events on the left arm of chromosome V as well as unselected crossovers and associated conversions on other chromosomes. For this table, the data obtained with high-throughput DNA sequencing was not used.

<sup>b</sup>3:1/4:0, 3:1/4:0/3:1, 1:3/0:4, or 1:3/0:4/1:3 conversion events.
TABLE 3. SUMMARY OF ALL UNSELECTED CONVERSION EVENTS UNASSOCIATED WITH LOH OR BIR EVENTS AS DIAGNOSED BY SNP MICROARRAYS.

<table>
<thead>
<tr>
<th>Type of event</th>
<th>Spontaneous</th>
<th>γ-ray</th>
<th>UV</th>
</tr>
</thead>
<tbody>
<tr>
<td>3:1 or 1:3 conversions</td>
<td>0</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>4:0 or 0:4 conversions</td>
<td>1</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>Hybrid conversions&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>Complex conversions</td>
<td>0</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>BIR</td>
<td>0</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td><strong>Totals</strong></td>
<td><strong>1</strong></td>
<td><strong>13</strong></td>
<td><strong>34</strong></td>
</tr>
</tbody>
</table>

<sup>a</sup>3:1/4:0, 3:1/4:0/3:1, 1:3/0:4, or 1:3/0:4/1:3 conversion events.
TABLE 4. COMPARISON BETWEEN RECOMBINATION EVENTS DETECTED BY HIGH-THROUGHPUT SEQUENCING OR BY SNP MICROARRAYS

<table>
<thead>
<tr>
<th>Type of event</th>
<th>γ-ray Array</th>
<th>HTS</th>
<th>UV</th>
<th>Array</th>
<th>HTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crossovers with no detectable conversion tracts</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Crossovers with 3:1 or 1:3 conversion tracts</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Crossovers with 4:0 or 0:4 conversion tracts</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Crossovers with hybrid conversion tracts</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Crossovers with complex conversion tracts</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>3:1 or 1:3 conversion tracts without crossovers</td>
<td>1</td>
<td>0</td>
<td>8</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>4:0 or 0:4 conversion tracts without crossovers</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Hybrid conversion tracts without crossovers</td>
<td>0</td>
<td>1</td>
<td>8</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Complex conversion tracts without crossovers</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>BIR</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total recombination events</strong></td>
<td><strong>5</strong></td>
<td><strong>6</strong></td>
<td><strong>32</strong></td>
<td><strong>35</strong></td>
<td></td>
</tr>
</tbody>
</table>

*aThis table includes data from two UV-induced sectored colonies and two sectored colonies induced by gamma radiation that were analyzed by both SNP arrays and HTS. The table includes both selected recombination events on chromosome V and unselected events on other chromosomes. All conversion events unassociated with LOH were unselected.
A. Reciprocal crossover without an associated conversion (G2 DSB)

B. Reciprocal crossover with associated conversion tract (G2 DSB)

C. Reciprocal crossover resulting from a G1 DSB with associated conversion tract
A. 4:0/3:1 hybrid tract

B. 3:1/4:0/3:1 hybrid tract

White Can^R  Red Can^R

White Can^R  Red Can^R