Research Article

Genome-wide map of Apn1 binding sites under oxidative stress in *Saccharomyces cerevisiae*

Lydia P. Morris[1,2,6#] ![ORCID], Andrew B. Conley[3,7†], Natalya Degtyareva[2], I. King Jordan[3] and
Paul W. Doetsch[1,2,4,5]*

[1]*Program in Genetics and Molecular Biology, Emory University Atlanta, GA 30322, USA*
[2]*Department of Biochemistry, Emory University Atlanta, GA 30322, USA*
[3]*School of Biology, Georgia Institute of Technology Atlanta, GA 30332, USA*
[4]*Department of Hematology and Medical Oncology, Emory University Atlanta, GA 30322, USA*
[5]*Winship Cancer Institute, Emory University Atlanta, GA 30322, USA*
[6]*WebbWrites, LLC, Durham, NC USA*
[7]*Cedars-Sinai Medical Center, Los Angeles, CA USA*

*Correspondence to:
Paul W. Doetsch, Winship Cancer
Institute, Emory University Atlanta,
GA 30322, USA.
E-mail: medpwd@emory.edu

#Present address: WebbWrites,
LLC, Durham, NC USA

†Present address: Cedars-Sinai
Medical Center, Los Angeles, CA
USA*

Abstract

The DNA is cells is continuously exposed to reactive oxygen species resulting in toxic and mutagenic DNA damage. Although the repair of oxidative DNA damage occurs primarily through the base excision repair (BER) pathway, the nucleotide excision repair (NER) pathway processes some of the same lesions. In addition, damage tolerance mechanisms, such as recombination and translesion synthesis, enable cells to tolerate oxidative DNA damage, especially when BER and NER capacities are exceeded. Thus, disruption of BER alone or disruption of BER and NER in *Saccharomyces cerevisiae* leads to increased mutations as well as large-scale genomic rearrangements. Previous studies demonstrated that a particular region of chromosome II is susceptible to chronic oxidative stress-induced chromosomal rearrangements, suggesting the existence of DNA damage and/or DNA repair hotspots. Here we investigated the relationship between oxidative damage and genomic instability utilizing chromatin immunoprecipitation combined with DNA microarray technology to profile DNA repair sites along yeast chromosomes under different oxidative stress conditions. We targeted the major yeast AP endonuclease Apn1 as a representative BER protein. Our results indicate that Apn1 target sequences are enriched for cytosine and guanine nucleotides. We predict that BER protects these sites in the genome because guanines and cytosines are thought to be especially susceptible to oxidative attack, thereby preventing large-scale genome destabilization from chronic accumulation of DNA damage. Information from our studies should provide insight into how regional deployment of oxidative DNA damage management systems along chromosomes protects against large-scale rearrangements. Copyright © 2017 John Wiley & Sons, Ltd.

Introduction

Reactive oxygen species (ROS) are highly reactive molecules and an inevitable byproduct of aerobic metabolism. ROS are also important for a variety of cellular processes, including cell signalling and protection against invading pathogens (Ježek and Hlavatá, 2005; Moncada, 1999; Guengerich, 2006; Segal and Shatwell, 1997; Orient *et al.*, 2007; Bedard and Krause, 2007). Dysregulation

of cellular ROS levels (Jones, 2006) or exposure to environmental agents that induce increases in intracellular ROS levels can cause oxidative stress, resulting in damage to different cellular components. ROS-induced damage to cellular DNA is one of the most frequently occurring types of endogenous DNA damage, producing base modifications and strand breaks within the genome (Wang, 2008; Bjelland and Seeberg, 2003; Henner *et al.*, 1983a, 1983b) at an estimated rate of 90 000 lesions per mammalian cell each day (Fraga *et al.*, 1990). Genomic instability is an important hallmark of cancer (Luo *et al.*, 2009; Hanahan and Weinberg, 2011) and results from the accumulation of genetic alterations from point mutations and small insertion/deletions to gross chromosomal aberrations. Although many studies have shown that exposure to ROS causes large-scale genome destabilization, the molecular details of how base damage is related to the acquisition of chromosome-level instability have not yet been elucidated.

The base excision repair (BER) pathway is responsible for repairing various types of non-bulky DNA damage (Robertson *et al.*, 2009). BER typically proceeds through the recognition of DNA damage by a lesion-specific glycosylase that removes the damaged base, leaving an apurinic/apyrimidinic (AP) site. Repair continues with recognition and cleavage of AP sites by an AP endonuclease. This is followed by end processing, repair DNA synthesis and ligation of the DNA backbone, restoring the undamaged state. The nucleotide excision repair (NER) pathway is primarily responsible for the repair of bulky, helix-distorting lesions, and can function as a backup pathway when BER is defective, demonstrating the importance of having multiple, overlapping mechanisms in place to protect against spontaneously and exogenously induced DNA damage.

Previous genetic studies in BER/NER-defective *Saccharomyces cerevisiae* strains revealed that the accumulation of spontaneous oxidative DNA damage leads to a profound increase in the frequencies of spontaneous mutations and gross chromosomal rearrangements. This genomic instability results from the handling of DNA damage by tolerance mechanisms (Degtyareva *et al.*, 2008), including recombination and translesion synthesis, respectively, which promote cell survival but may not

repair damage *per se*. Further, the genomic instability in such repair-deficient strains is characterized by the emergence of chromosome rearrangement hotspots, suggesting that certain regions are more susceptible to destabilization. We hypothesized that underlying genomic features at DNA repair sites, such as base composition, directly influence the location and frequency of oxidative damage-induced destabilization. Identification of such features would provide important insights into understanding why certain regions are particularly susceptible to damage-induced instability under oxidative stress.

A detailed understanding of the genomic context of repair can be gained by assessing DNA repair at the chromosomal level. Chromatin immunoprecipitation (ChIP) assays combined with genome-wide analyses have provided a wealth of information regarding the protein–DNA interactions of DNA binding proteins, such as transcription factors (Kuras and Struhl, 1999; Li *et al.*, 1999) and histones (Solomon *et al.*, 1988). However, no studies to date have characterized the genomic occupancy of base excision repair enzymes using such high-resolution methods, presumably owing to the transient and dynamic manner in which repair enzymes interact with the DNA.

In the present study, we sought to determine if the underlying base content influences localization of DNA repair pathway components and how the location of DNA repair machinery (and hence, where repair is preferentially targeted) influences genome destabilization. We used the model eukaryote *S. cerevisiae* and used the major yeast AP endonuclease Apn1 as a representative DNA repair protein for the analysis of the genome-wide localization of the BER machinery. The repair of base damage processed by BER, including hydrolysis, oxidation, alkylation and deamination in DNA, proceeds through an AP site repair intermediate. Thus, AP endonucleases are a convergence point and play a central role in the repair of DNA damage through the BER pathway. We performed chromatin immunoprecipitation combined with DNA microarray analysis (ChIP-chip) to generate genome-wide maps of Apn1 binding in response to different levels of oxidative stress and assessed the underlying genomic landscape of these regions. Our ultimate goals were to determine which genomic features

are predictive of oxidative damage-induced destabilization.

Methods

Yeast strains and culture conditions

Standard yeast cell culture conditions utilizing either YPD (yeast extract, peptone, dextrose) or YPG (yeast extract, peptone, 2% galactose) culture medium as described previously (Griffiths *et al*., 2009). Yeast cell transformation was performed using the lithium acetate method, as described previously (Schiestl and Gietz, 1989).

Strain construction

Parental strain DSC0320 [*MATa* (*lys2::Alu-DIR-LEU2-lys2D5′*) *ade5–1 his7–2 leu2–3112 trp1–289 ura3–52 APN1*] was isolated as a haploid spore of the diploid strain hDNP16 (Degtyareva *et al*., 2008). Strain DSC0436 [*MATa* (*lys2::Alu-DIR-LEU2-lys2D5′*) *ade5–1 his7–2 leu2–3112 trp1–289 ura3–52 TRP1 GAL1–TAP–APN1*], which contains a construct with the *GAL1* promoter and the tandem affinity purification (TAP)-tag integrated directly upstream of the chromosomal *APN1* coding sequence (P_{GAL1}–TAP–*APN1*), was constructed as follows: a P_{GAL1}–TAP–*APN1* fragment was PCR amplified with primers APN1TAPf (AAACACAAAACGCAAC ATTAATAAGCTTTTGG CATATCGGAAC CATCGTAGAACAAAAGCTGGAGCTCAT) and APN1TAPr (AATTT GTATTTCGAGACAGCA GATCTAACAAAGCTAGGTGTCGAAGGCATC TTATCGTCATCATCAAGTG) using plasmid pBS1761 (Puig *et al*., 2001) as a template. Strain DSC0320 was transformed with the resulting PCR fragment. Correct integration of the P_{GAL1}–TAP construct at the *APN1* locus was confirmed by PCR with primers APN1chkTAPf (TCTGGGAACTTGAA CGTGGA ATT) and TRP1TAPchk (CGTGGT ACAGTTGAAGGACATCATC).

Analysis of mutation rates

Standard mutant accumulation assays were performed to determine mutation rates, as described previously (Drake, 1991; Meisel, 1971; Wierdl *et al*., 1996). Briefly, eight independent colonies

of both wild-type *APN1* strain DSC320 and *pGAL1*–TAP–*APN1* strain DSC436 were grown overnight in YPG liquid media to induce overexpression of *APN1* in cells of strain DSC436. Cells were harvested via centrifugation, washed twice with sterile H_2O and re-suspended in water. Dilutions of cells were plated onto SD medium lacking arginine and containing 60 mg/L L-canavanine (Sigma) to determine the number of canavanine-resistant (*can^r*) cells. Appropriate dilutions were also plated onto YPD medium to determine numbers of viable cells. Colonies were counted after 2–4 days of growth at 30°C to determine the *can^r* mutation rates as previously described. Each experiment was repeated at least 3 times for a total of at least 24 independent cultures per strain.

Analysis of H_2O_2 cytotoxicity

Yeast cell cultures were prepared by inoculating YPG media with cells of strain DSC0436, incubating at 30°C to OD_{600} = ~0.8, and then splitting into three aliquots. Cultures were harvested and cells were washed twice in sterile water and resuspended in either sterile water for mock treatment or 0.5 mM hydrogen peroxide (H_2O_2) solution. Samples were incubated for 90 min at 30°C. Cells were harvested, washed twice with sterile water and resuspended in sterile water. Appropriate dilutions of cells were plated onto YPD in duplicate, and numbers of viable colonies were assessed after 2 days of incubation at 30°C. Percentage survival was calculated based on the number of colonies that grew from cultures exposed to H_2O_2 compared with the number from mock-treated cultures. Results are the average of at least four to six independent experiments.

Analysis of mutation frequency following H_2O_2 exposure

To determine the mutation frequency following exposure to H_2O_2, we performed fluctuation tests. Six independent colonies of the wild-type *APN1* strain were grown non-selectively overnight in YPG liquid media and processed as described in the section 'Analysis of mutation rates'. After washing and resuspending, cells were exposed to 0.5 mM H_2O_2 as described in the section 'Analysis of H_2O_2 cytotoxicity'. Appropriate dilutions of cells were plated onto canavanine-containing

medium and YPD medium to determine the number of can^r cells and viable cell numbers, respectively. After 2–4 days of growth at 30°C, colonies were counted and the number of can^r mutant colonies was compared with the viable cell numbers to determine mutation frequency. The median mutation frequencies of at least 20 independent cultures were reported. p-Values were calculated for confidence limits as previously described (Dixon and Massey, 1969). A p-value >0.05 was considered statistically significant.

Chromatin immunoprecipitation

Chromatin immunoprecipitation experiments were carried out in unsynchronized *S. cerevisiae* cultures as previously described (Luthra *et al.*, 2007) with the following modifications. For each replicate, a culture of YPD media inoculated with cells from strain DSC0436 was grown to $OD_{600} \sim 0.8$, and then split into two equivalent volume aliquots. One culture was exposed to 0.5 mM H_2O_2, and the other was subjected to a mock (sterile H_2O) exposure condition, as described above. Following glycine quenching, cells were washed with sterile water. Harvested cell pellets were stored at -80°C until used. For immunoprecipitation, pellets were sonicated in FA lysis buffer (50 mM Hepes–KOH, pH 7.5; 300 mM NaCl; 1 mM EDTA; 1% Triton-X; 0.1% sodium deoxycholate) and the soluble fraction was incubated overnight with anti-TAP antibody (Thermo). The chromatin–antibody mixture was incubated with protein A agarose beads (Invitrogen) for 2 h at room temperature. The beads were then washed and antigens eluted in elution buffer (1% SDS, 0.1 M NaHCO₃). Cross-link reversal was performed in the absence of proteinase K. Three independent experiments were performed for each experimental condition. One sample from the 0.5 mM condition was used for assay optimization. Thus, there were three biological replicates for the 0 mM (basal oxidative stress) condition and two replicates for the 0.5 mM dose (mild oxidative stress) condition.

ChIP-Chip procedure

ChIP DNA was labelled and hybridized to the Affymetrix *S. cerevisiae* Tiling Array 1.0 consisting of 3.2 million 25-mer oligo probes with an overlap of ~20 base pairs on adjacent probes tiled through the entire *S. cerevisiae* genome (Gresham *et al.*, 2006). Labelling and hybridization were done according to the Affymetrix Chromatin Immunoprecipitation Assay Protocol (Affymetrix Chromatin Immunoprecipitation Assay Protocol, 2017) with the following modifications. For the PCR amplification of immunoprecipitated DNA, the DNA was first amplified using the GenomePlex Single Cell Whole Genome Amplification Kit (Sigma-Aldrich). The amplified DNA was then purified using the GenElute PCR Clean-Up Kit (Sigma-Aldrich). Next, the purified DNA was amplified again using the GenomePlex WGA Reamplification Kit (Sigma-Aldrich), and then purified as just described. The tiling array was then scanned according to the Affymetrix ChIP Assay Protocol.

ChIP-Chip data analysis

The raw ChIP-chip probe intensity values from the resulting CEL files (available in the Gene Expression Omnibus data repository), for each experimental condition, were normalized using the Loess normalization procedure implemented in the R *Starr* ChIP-chip analysis package (Zacher *et al.*, 2010). Ratios of normalized probe intensities for Apn1 immunoprecipitated chromatin vs. input chromatin were then calculated using *Starr*. These ratios were used to measure Apn1 occupancy along the chromosomes. Chromosomal locations of the probes were taken from the Affymetrix probe annotations for the yeast 2003 genome build (Liu *et al.*, 2003).

For each of the experimental conditions, enriched regions, or 'peaks', of Apn1 binding were called using the R *Cmarrt* software package (Kuan *et al.*, 2008) with the normalized ratios generated via the *Starr* package as inputs. Analysis was initiated with three replicates for each experimental condition. One replicate from the oxidative stress condition was used for trouble shooting prior to subjecting the DNA to microarray analysis. Of the remaining replicates, those that produced peaks include two from the basal oxidative stress condition and one from the mild oxidative stress condition. Thus, data from a total of three replicates from the two experimental conditions was analysed further.

The average GC content (the fraction of a sequence made of G and C residues) was determined for the Apn1 ChIP-chip peaks for each of the three experimental conditions. These averages were compared against a genomic background distribution of average GC content generated via simulation analysis using the overlap between 10 000 sets of random genomic loci of the same size as the Apn1 peaks. Standard errors were determined on the expected average GC content values computed over the 10 000 simulated peaks and used to compute the significance of the difference between the observed and expected GC content values (see Table 2). For the genome-wide GC content analysis results shown in Figure S1 in the Supporting Information, sliding windows of 10 kb were used, with a step size of 5 kb, to define the GC content. For the Apn1 binding site peak GC content enrichment analysis shown in Fig. 3, sliding windows of 100 BP were used, with a step size of 50 BP in order to define GC content. GC content values were evaluated for observed Apn1 binding site peaks and their flanking regions along with a randomly simulated set of 10 000 genomic sites used to calculate the expected GC content values at and around Apn1 binding site peaks. Microarray data and processed data files are available at https://submit.ncbi.nlm.nih.gov/geo/submission.

Results

Assessment of H_2O_2-induced cell killing

We set out to characterize the distribution of Apn1 binding sites along the yeast chromosomes under normal and mild oxidative stress conditions. Although Apn1 is present at a higher steady-state concentration than the other yeast BER proteins, there are only ~7000 copies of Apn1 per cell (Johnson and Demple, 1988; Ghaemmaghami et al., 2003). Thus, we reasoned that overexpression of Apn1 from the endogenous locus (using the galactose-inducible GAL1 promoter) would increase the likelihood of identifying Apn1 genomic target sites. Because no prior knowledge about the genomic localization of Apn1 exists, we added a TAP tag, which allowed us to use a well-established ChIP protocol for the identification of TAP-tagged yeast DNA binding protein target sites (Luthra et al., 2007). Overexpression of the

yeast BER protein Mag1 has been shown to be mutagenic (Glassner et al., 1998; Hanna et al., 2004). To ensure this was not the case for Apn1, which could indicate abnormalities in BER caused by DNA binding beyond the normal Apn1 distribution, we determined spontaneous mutation rates as a measure of the level of DNA damage (Swanson et al., 1999) and found that the mutation rates for the strain expressing Apn1 at endogenous levels and the strain with Apn1 overexpression were not significantly different (Table 1).

To assess the relationship between oxidative stress level and localization of BER machinery in the genome, we used an acute treatment with hydrogen peroxide (H_2O_2) to induce oxidative DNA damage. Following exposure to 0.5 mM H_2O_2 for 90 min, 80% survival resulted compared with the mock treatment condition. As a biologically relevant endpoint for the level of induced DNA damage in yeast cells, we measured the mutation frequency following exposure to H_2O_2 in wild-type cells without the GAL1p–TAP construct. H_2O_2 treatment induced a mutation frequency of 8.6×10^{-7} compared with 0.38×10^{-7} for no treatment, an ~22-fold increase. Based on these levels of cell killing and mutation frequencies, we regarded exposure to 0.5 mM H_2O_2 for 90 min as mild stress and the mock exposure as the normal, basal oxidative stress condition.

Identification of Apn1 binding sites across the yeast genome

For microarray analysis of the enrichment of genomic DNA in the ChIP DNA samples, we utilized the Affymetrix S. cerevisiae Tiling Array 1.0 (Materials and Methods). We identified peaks of Apn1 binding using the R Cmarrt algorithm based on normalized probe intensity ratios produced by

Table 1. Median mutation rates in cells with *APN1* at endogenous or overexpressed levels

Strain	$Can^r \times 10^{-8}$ (95% confidence interval)	Fold change
Wild type	3.2 (0.74–3.3)	1
pGAL1-APN1	2.3 (1.4–2.8)	0.7

Assays for determination of mutation rates were performed in complete media where galactose was used as the carbon source to induce overexpression at the *APN1* locus.

the R *Starr* software package. We performed each ChIP experiment in triplicate, but we used one sample from the mild oxidative stress experimental condition for microarray quality control and optimization and thus it was not hybridized to the DNA chip. Two replicates from the basal level of oxidative stress conditions and a single replicate from the mild stress condition produced peaks. Without exposure to exogenous ROS, we identified 916 total Apn1 peaks across all chromosomes (Fig. 1A, Fig. S1). Following mild oxidative stress we identified 211 peaks across all chromosomes (Fig. 1A, Fig. S1).

The average heights and widths of the peaks were the same between the two replicates for the basal oxidative stress condition (Fig. 1B, C), demonstrating the technical reproducibility of the data. In addition, the average heights (level of Apn1 occupancy) and widths (length of DNA that makes up an Apn1 binding sites) of the peaks were
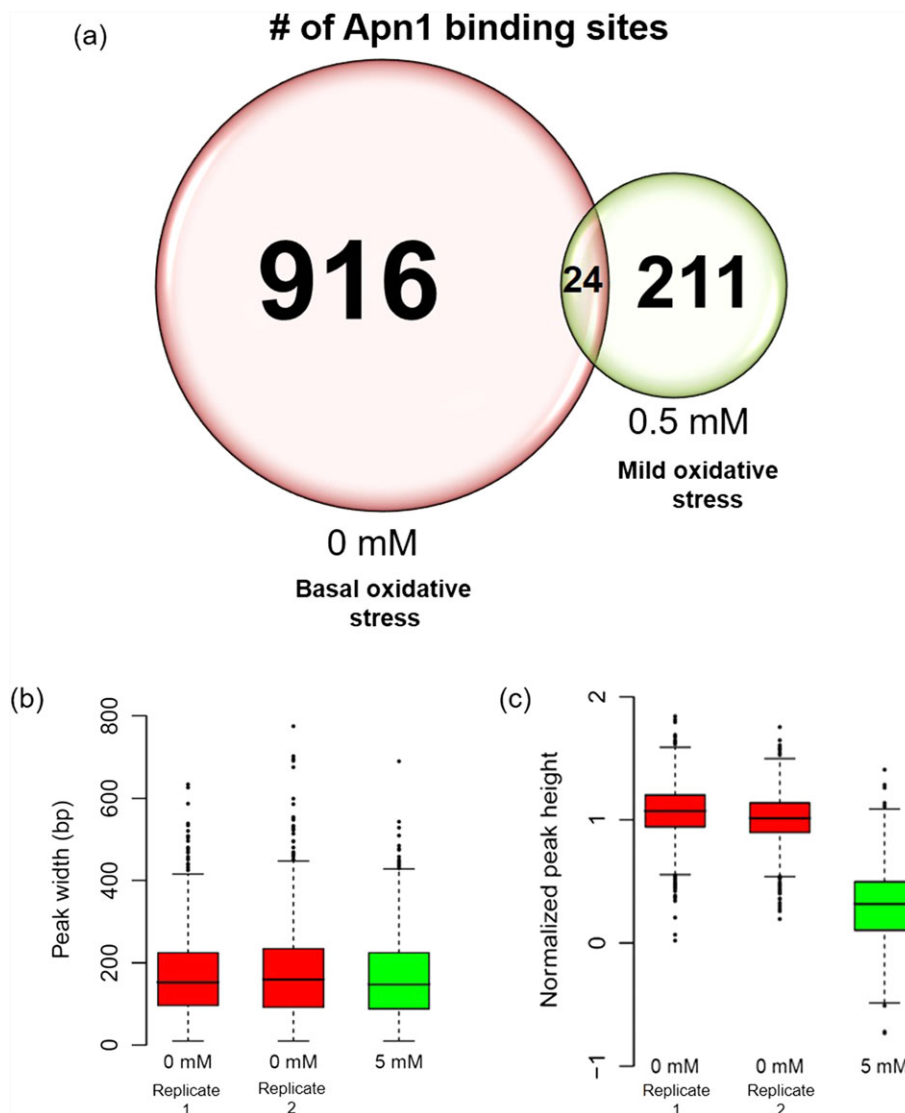
Figure 1. Characteristics of Apn1 binding peaks. (A) Venn diagram displaying the numbers of Apn1 binding peaks. Overlaps between the numbers of Apn1 binding peaks under basal oxidative stress (red) and mild oxidative stress (green). The average Apn1 binding peak (B) heights (ratio) and (C) widths were determined for each of the experimental conditions for which there were a significant number of peaks

the same among all of the replicates for both experimental conditions, indicating that Apn1 may bind the DNA with similar density regardless of oxidative stress level and independently of the amount of DNA damage present (Fig. 1B). The average peak heights were significantly higher in the 0 mM basal oxidative stress condition compared with the 0.5 mM condition ($p \approx 0$).

We predicted that the acute oxidative stress induced in our studies might produce peaks in a region on chromosome II that we previously showed is subject to chronic stress-induced destabilization (Degtyareva et al., 2008). However, neither experimental condition produced peaks that were enriched in the chromosome II region (Fig. 2, Table S1).

### Apn1 preferentially targets GC-rich sequences

To determine whether the underlying DNA sequence influences the locations of Apn1 binding sites, we assessed the base-content within Apn1 binding sites (see 'Materials and methods' section). For both the basal oxidative stress and the mild oxidative stress condition, the GC fraction of the Apn1 binding site peaks was found to be significantly higher than the genomic background GC fraction (Table 2). The genomic background GC content and the p-values in Table 2 were computed using 10 000 randomly simulated sets of genomic loci of the same size as the Apn1 peaks as described in Materials and Methods.

In addition, evaluation of GC content levels in the genomic regions surrounding Apn1 binding sites shows clear peaks of GC content that are coincident with the Apn1 binding site locations and higher than those seen for the local genomic background (Figure 3). GC content values at observed Apn1 binding sites can also be seen to be higher than expected based on comparison with randomly simulated binding site locations (Fig. 3). These results suggest that BER preferentially targets GC-rich DNA. This finding is supported by the fact that Gs and Cs are thought to be more susceptible to spontaneous damage and exogenous oxidative damage than the other bases (Kreutzer and Essigmann, 1998; Schaaper and Dunn, 1991; Friedberg et al., 2006). Not all GC-rich regions overlap with Apn1 binding sites (Figure S1), which suggests that BER may preferentially target particular subsets of GC-rich DNA.

Discussion

To our knowledge, the present study is the first effort towards high-resolution mapping of the
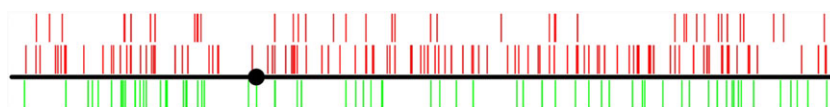
Figure 2. Locations of Apn1 binding site peaks on chromosome II. Locations are shown in red above the chromosome ideogram for replicate 1 (upper row) and replicate 2 (lower row) of the 'basal oxidative stress' experimental condition and in green below the chromosome for the mild oxidative stress condition

Table 2. Genome-wide enrichment of GC content for Apn1 binding sites[a]

Experimental condition	GC content fractions for Apn1 peaks[a]		
—	Observed[b]	Expected[c]	p-Value[d]
0 mM	0.433	$0.383 \pm 1.35 \times 10^{-4}$	$<10^{-5}$
0.5 mM	0.401	$0.383 \pm 1.35 \times 10^{-4}$	$<10^{-5}$

[a]Genome-wide enrichment of GC content for Apn1 binding sites was assessed by comparing the observed vs. expected GC content fractions for Apn1 ChIP-chip peaks as described in the 'Materials and methods' section.
[b]Average GC fractions observed for all Apn1 binding sites (i.e. ChIP-chip peaks) genome-wide. The average GC fractions are significantly different ($p = 1.1 \times 10^{-25}$) between the 0 and 0.5 mM conditions.
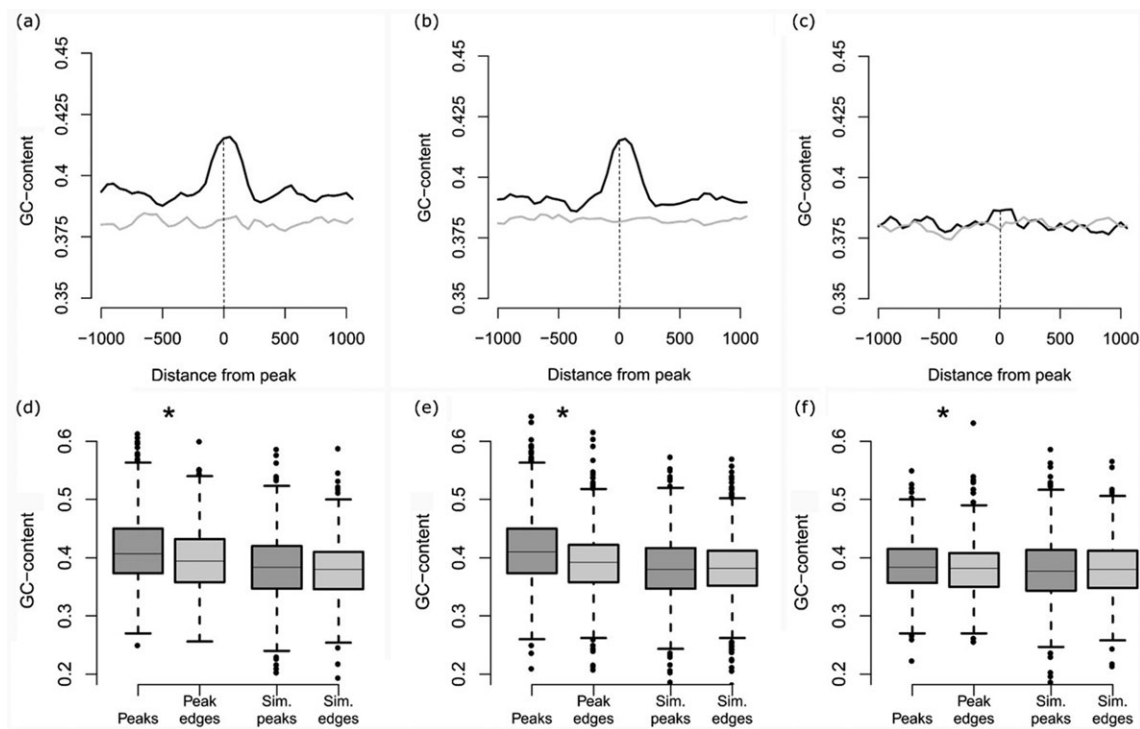[c]Expected GC fractions for Apn1 binding sites, ± standard errors, based on simulations of 10 000 random genomic sites.
[d]Significance of the difference between the observed vs. expected GC content fractions for Apn1 binding sites.

genomic binding sites of a BER protein utilizing the ChIP approach. We have identified the binding sites of Apn1 across the genome for different oxidative stress conditions and have assessed the underlying base content, the patterns of which may influence the genomic localization of BER enzymes.

We found that, regardless of the oxidative stress level, GC content was significantly enriched within peaks of Apn1 binding. Of the four major bases in DNA, guanine is the most common target of methylation as well as hydrolytic and oxidative base damage (Bolin and Cardozo-Pelaez, 2007; Lindahl and Nyberg, 1972; Lindahl and Barnes, 2000). It is reasonable to predict that BER promotes genomic stability by protecting certain regions with a higher

content of guanine bases, where more spontaneous DNA damage may occur, or where the underlying DNA is more susceptible to oxidative damage-induced breakage and subsequent destabilization. These data are in line with our previous studies, which demonstrated that, in yeast strains deficient in the ability to repair spontaneous base damage (BER$^-$/NER$^-$), damage tolerance mechanisms, such as homologous recombination (and probably non-homologous end joining), are engaged in handling the persistent damage. This results in a high level of gross chromosomal rearrangements, some of which occur at hotspots within the genome (Degtyareva *et al.*, 2008). In addition, chronic H_2O_2 exposure has also been shown to induce genome rearrangements in *S. cerevisiae* (Ragu *et al.*,

Figure 3. GC content for Apn1 binding sites. The observed GC content levels were measured for all genomic regions flanking Apn1 binding site peaks and compared with the expected GC content levels based on randomly simulated peak locations. Observed (black) vs. expected (grey) GC content levels are shown for genomic regions flanking Apn1 binding sites (peaks) for replicate 1 (A) and replicate 2 (B) of the 'basal oxidative stress' experimental condition and for the mild oxidative stress condition (C). Dashed vertical lines indicate the 0 position of the distance from the peak on the *x*-axes. Box-plots showing distributions of GC content fractions at and around Apn1 binding sites. The average values for 'peaks' (observed) and 'simulated peaks' (expected) correspond to GC content fractions ±100 BP from the centre for the Apn1 binding site peaks from (A)–(C), whereas the GC content fractions for 'peak edges' and 'simulated edges' correspond to two 100 BP regions at the far edges of the plots shown in (A)–(C). Distributions of GC content fractions for these regions are shown for replicate 1 (D) and replicate 2 (E) of the 'basal oxidative stress' experimental condition and for the mild oxidative stress condition (F)

2007). Such 'tolerance-induced' destabilization probably occurs in particular regions because of the accumulation of damage at these loci that is normally repaired by BER and NER. As guanine bases in DNA are highly susceptible to different types of endogenous and exogenous damage, our results suggest that the BER machinery localizes to areas of the genome that are most susceptible to genotoxic insult to prevent large-scale genome destabilization.

The facts that there were fewer peaks in the mild oxidative stress condition and that there was little overlap of peaks between the two replicates for the basal oxidative stress condition (Fig. 1) were unexpected and led us to speculate that we may not have identified all of the potential AP endonuclease binding sites under the two conditions used in this study. While Apn1 represents 97% of the major AP endonuclease activity in yeast cells under normal growth conditions, yeast also possess

a minor DNA damaging agent-inducible AP endonuclease, Apn2, whose activity becomes important when Apn1 is absent or when DNA damage stress increases substantially (Johnson *et al.*, 1998; Bennett, 1999). It is possible that Apn1 localizes to certain regions under normal cellular conditions, but following exogenous insult-induced DNA damage, Apn1 and Apn2 divide and conquer to make sure the prioritized regions and regions of induced damage are repaired to preserve genome stability. To test this idea, future studies will involve comparing the genome-wide binding profiles of Apn1 and Apn2 under basal and mild oxidative stress conditions.

A major remaining question is how oxidative DNA damage normally handled by excision repair is tolerated to produce large-scale chromosomal changes. Our chromosome II analysis (Figure 3) showed no overlap between Apn1 binding sites and the major region of spontaneous oxidative
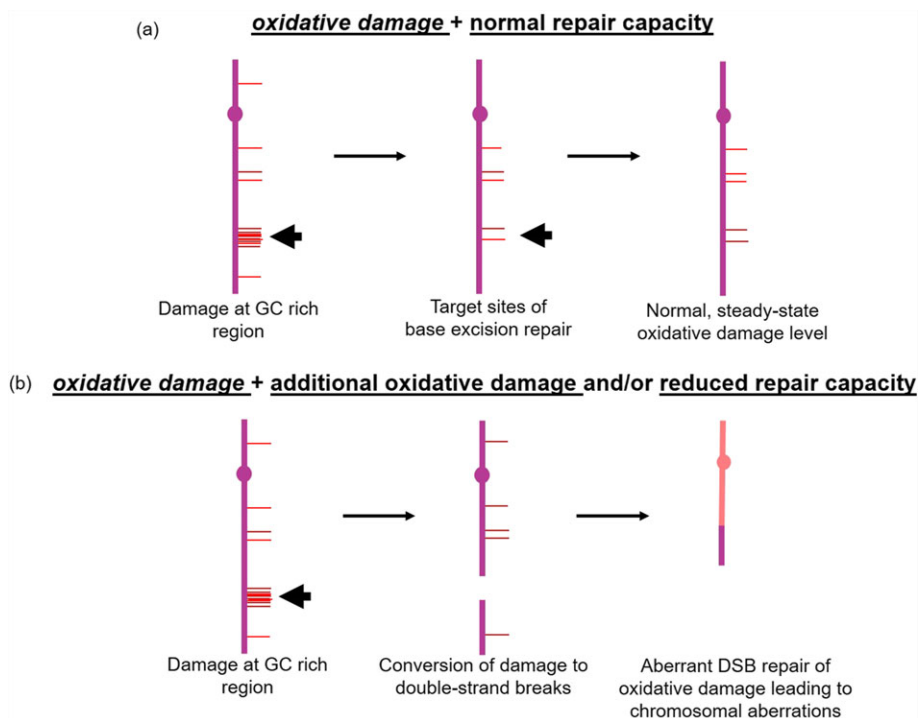
Figure 4. A model for oxidative damage-induced genome destabilization. (A) When cells exhibit normal DNA repair capacity, the level of endogenous oxidative DNA damage remains low (estimated 550 lesions/genome in yeast at stationary phase (Evert, et al., 2004)) because the major oxidative targets (Gs and Cs) are prioritized for repair via the base excision repair (BER) pathway in certain regions of the genome. Virtually no large-scale rearrangements occur under these conditions (Degtyareva, et al., 2008). (B) When there is more oxidative damage than can be efficiently repaired, and/ or when cells exhibit reduced repair capacity due to defects in the DNA repair proteins, damage tolerance pathways (TLS, HR) allow the replication machinery to by-pass damage, thereby promoting cell survival, but this occurs at the expense of increased genomic instability (arm loss and translocation depicted as examples)

DNA damage-induced destabilization. The lack of overlap is probably because our experiments were performed under acute stress conditions while chromosome II rearrangements are a result of chronic oxidative stress and are produced following accumulation of spontaneous damage after serial passaging over multiple cell generations. Based on the results presented here and our previously published work (Degtyareva *et al.*, 2008), we propose a model whereby AP endonuclease molecules are distributed along yeast chromosomes at certain locations (garage sites) to repair spontaneous damage under basal oxidative stress levels (Figure 4). When the amount of damage is increased (0.5 mM H_2O_2 exposure), Apn1 binding sites are also enriched for G and C nucleotides, but the distribution pattern changes. The propensity with which a certain DNA damage pattern dictates conversion of the lesions into double-strand breaks may be influenced by the precise location of Apn1 binding sites because, in the absence of DNA repair machinery, or when the amount of damage exceeds the capacity to repair it, the unrepaired damage is converted into double-strand breaks (Harrison *et al.*, 2006; Karanjawala *et al.*, 2002), the mis-repair of which leads to gross chromosomal rearrangements (Ragu *et al.*, 2007; Duell *et al.*, 1995; Limoli and Giedzinski, 2003).

The interactions of excision repair enzymes with the DNA are expected to be quite long for an enzyme (Schermerhorn and Delaney, 2014), but transient when compared with other types of DNA binding proteins, such as transcription factors. The major advantage of our study, the unbiased approach in identifying binding sites for a protein for which no binding information exists, is also a drawback because there are no positive or negative control regions to optimize various aspects of our experimental protocol by PCR. An alternative strategy would be to identify an Apn1 variant that binds more strongly (i.e. for a longer period of time) to the DNA to increase the number of Apn1-DNA complexes pulled down in each ChIP experiment. The catalytic amino acid substitution variant Apn1(D192G) was shown to bind DNA more efficiently than wild type (Jilani *et al.*, 2003). Recently, we used a mutagenesis screen approach to identify and characterize functional amino acid substitution variants of Apn1 (Morris *et al.*, 2012). We will use this approach in future studies to identify 'strong binder' variants of

Apn1 to further characterize the genome-wide occupancy.

With the above experimental system, it should be possible to generate a complete map of Apn1 binding under different environmental stress conditions as well as identifying other *cis*-features in addition to base content, such as chromatin conformation and histone modification, that contribute to oxidative stress-induced genome instability. As large-scale genomic instability is a hallmark of cancer, elucidating factors that influence location of DNA damage-induced rearrangements will be an important step in understanding the role of another cancer hallmark, oxidative stress, in carcinogenesis.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

L.P.M. participated in the study design, carried out all yeast molecular biology experiments, performed statistical analysis and drafted the manuscript. A.C. performed statistical analysis and helped draft the manuscript. K.G. participated in the statistical analysis. P.D. and N.D. conceived of the study, participated in its design, and helped draft the manuscript. All authors read and approved the final manuscript.

Acknowledgments

References

Affymetrix Chromatin Immunoprecipitation Assay Protocol. 2017. tools.thermofisher.com/content/sfs/manuals/chromatin_immun_ChIP.pdf (accessed 27 February 2017).

Bedard K, Krause K-H. 2007. The NOX family of ROS-generating NADPH oxidases: physiology and pathophysiology. *Physiol Rev* **87**: 245–313.

Bennett RAO. 1999. The *Saccharomyces cerevisiae* ETH1 gene, an inducible homolog of exonuclease III that provides resistance to DNA-damaging agents and limits spontaneous mutagenesis. *Mol Cell Biol* **19**: 1800–1809.

Bjelland S, Seeberg E. 2003. Mutagenicity, toxicity and repair of DNA base damage induced by oxidation. *Mutat Res Fund Mol Mech Mut* **531**: 37–80.

Bolin C, Cardozo-Pelaez F. 2007. Assessing biomarkers of oxidative stress: analysis of guanosine and oxidized guanosine nucleotide triphosphates by high performance liquid chromatography with electrochemical detection. *Biomed Applic* **856**: 121–130.

Degtyareva NP, Chen L, Mieczkowski P, *et al*. 2008. Chronic oxidative DNA damage due to DNA repair defects causes chromosomal instability in *Saccharomyces cerevisiae*. *Mol Cell Biol* **28**: 5432–5445.

Dixon WJ, Massey FJ, Jr. 1969. Introduction to Statistical Analysis. McGraw-Hill: New York; 349.

Drake JW. 1991. A constant rate of spontaneous mutation in DNA-based microbes. *Proc Natl Acad Sci* **88**: 7160–7164.

Duell T, Lengfelder E, Fink R, *et al*. 1995. Effect of activated oxygen species in human lymphocytes. *Mutat Res DNA Repair* **336**: 29–38.

Evert BA, Salmon TB, Song B, Jingjing L, Siede W, Doetsch PW. 2004. Spontaneous DNA damage in *Saccharomyces cerevisiae* elicits phenotypic properties similar to cancer cells. *J Biol Chem* **279**: 22585–22594.

Fraga CG, Shigenaga MK, Park JW, *et al*. 1990. Oxidative damage to DNA during aging: 8-hydroxy-2′-deoxyguanosine in rat organ DNA and urine. *Proc Natl Acad Sci* **87**: 4533–4537.

Friedberg EC, Walker GC, Siede W, *et al*. 2006. DNA Repair and Mutagenesis, 2nd edn. American Society of Microbiology Press: Washington, DC.

Ghaemmaghami S, Huh W-K, Bower K, *et al*. 2003. Global analysis of protein expression in yeast. *Nature* **425**: 737–741.

Glassner BJ, Rasmussen LJ, Najarian MT, *et al*. 1998. Generation of a strong mutator phenotype in yeast by imbalanced base excision repair. *Proc Natl Acad Sci* **95**: 9997–10002.

Gresham D, Ruderfer DM, Pratt SC, *et al*. 2006. Genome-wide detection of polymorphisms at nucleotide resolution with a single DNA microarray. *Science* **311**: 1932–1936.

Griffiths LM, Swartzlander D, Meadows KL, *et al*. 2009. Dynamic compartmentalization of base excision repair proteins in response to nuclear and mitochondrial oxidative stress. *Mol Cell Biol* **29**: 794–807.

Guengerich F. 2006. Cytochrome P450s and other enzymes in drug metabolism and toxicity. *AAPS J* **8**: E101–E111.

Hanahan D, Weinberg RA. 2011. Hallmarks of cancer: the next generation. *Cell* **144**: 646–674.

Hanna M, Chow BL, Morey NJ, *et al*. 2004. Involvement of two endonuclease III homologs in the base excision repair pathway for the processing of DNA alkylation damage in *Saccharomyces cerevisiae*. *DNA Repair* **3**: 51–59.

Harrison L, Brame KL, Geltz LE, *et al*. 2006. Closely opposed apurinic/apyrimidinic sites are converted to double strand breaks in *Escherichia coli* even in the absence of exonuclease III, endonuclease IV, nucleotide excision repair and AP lyase cleavage. *DNA Repair* **5**: 324–335.

Henner WD, Grunberg SM, Haseltine WA. 1983a. Enzyme action at 3′ termini of ionizing radiation-induced DNA strand breaks. *J Biol Chem* **258**: 15198–15205.

Henner WD, Rodriguez LO, Hecht SM, *et al*. 1983b. Gamma ray induced deoxyribonucleic acid strand breaks. 3′ Glycolate termini. *J Biol Chem* **258**: 711–713.

Ježek P, Hlavatá L. 2005. Mitochondria in homeostasis of reactive oxygen species in cell, tissues, and organism. *Int J Biochem Cell Biol* **37**: 2478–2503.

Jilani A, Vongsamphanh R, Leduc A, *et al*. 2003. Characterization of two independent amino acid substitutions that disrupt the DNA repair functions of the yeast Apn1. *Biochemistry* **42**: 6436–6445.

Johnson AW, Demple B. 1988. Yeast DNA diesterase for 3′-fragments of deoxyribose: purification and physical properties of a repair enzyme for oxidative DNA damage. *J Biol Chem* **263**: 18009–18016.

Johnson RE, Torres-Ramos CA, Izumi T, *et al*. 1998. Identification of APN2, the *Saccharomyces cerevisiae* homolog of the major human AP endonuclease HAP1, and its role in the repair of abasic sites. *Genes Dev* **12**: 3137–3143.

Jones DP. 2006. Redefining oxidative stress. *Antioxid Redox Sign* **8**: 1865–1879.

Karanjawala ZE, Murphy N, Hinton DR, *et al*. 2002. Oxygen metabolism causes chromosome breaks and is associated with the neuronal apoptosis observed in DNA double-strand break repair mutants. *Curr Biol* **12**: 397–402.

Kreutzer DA, Essigmann JM. 1998. Oxidized, deaminated cytosines are a source of C → T transitions in vivo. *Proc Natl Acad Sci* **95**: 3578–3582.

Kuan PF, Chun H, Keles S. 2008. CMARRT: a tool for the analysis of ChIP-chip data from tiling arrays by incorporating the correlation structure. *Pac Symp Biocomput*: 515–526.

Kuras L, Struhl K. 1999. Binding of TBP to promoters in vivo is stimulated by activators and requires Pol II holoenzyme. *Nature* **399**: 609–613.

Li X-Y, Virbasius A, Zhu X, *et al*. 1999. Enhancement of TBP binding by activators and general transcription factors. *Nature* **399**: 605–609.

Limoli CL, Giedzinski E. 2003. Induction of chromosomal instability by chronic oxidative stress. *Neoplasia* **5**: 339–346.

Lindahl T, Barnes DE. 2000. Repair of endogenous DNA damage. *Cold Spring Harb Symp Quant Biol* **65**: 127–134.

Lindahl T, Nyberg B. 1972. Rate of depurination of native deoxyribonucleic acid. *Biochemistry* **11**: 3610–3618.

Liu G, Loraine AE, Shigeta R, *et al*. 2003. NetAffx: Affymetrix probesets and annotations. *Nucl Acids Res* **31**(1): 82–86.

Luo J, Solimini NL, Elledge SJ. 2009. Principles of cancer therapy: oncogene and non-oncogene addiction. *Cell* **136**: 823–837.

Luthra R, Kerr SC, Harreman MT, *et al*. 2007. Actively transcribed GAL genes can be physically linked to the nuclear pore by the SAGA chromatin modifying complex. *J Biol Chem* **282**: 3042–3049.

Meisel P. 1971. The molecular basis of mutation. *Mol Nutr Food Res* **15**: 602–603.

Moncada S. 1999. Nitric oxide: discovery and impact on clinical medicine. *JRSM* **92**: 164–169.

Morris LP, Degtyareva N, Sheppard C, *et al*. 2012. *Saccharomyces cerevisiae* Apn1 mutation affecting stable protein

expression mimics catalytic activity impairment: implications for assessing DNA repair capacity in humans. *DNA Repair* **11**: 753–765.

Orient A, Donkó Á, Szabó A, *et al*. 2007. Novel sources of reactive oxygen species in the human body. *Nephrol Dial Transplant* **22**: 1281–1288.

Puig O, Caspary F, Rigaut G, *et al*. 2001. The Tandem Affinity Purification (TAP) method: a general procedure of protein complex purification. *Methods* **24**: 218–229.

Ragu S, Faye G, Iraqui I, *et al*. 2007. Oxygen metabolism and reactive oxygen species cause chromosomal rearrangements and cell death. *Proc Natl Acad Sci* **104**: 9747–9752.

Robertson A, Klungland A, Rognes T, *et al*. 2009. DNA repair in mammalian cells. *Cell Mol Life Sci* **66**: 981–993.

Schaaper RM, Dunn RL. 1991. Spontaneous mutation in the *Escherichia coli* lacI gene. *Genetics* **129**: 317–326.

Schermerhorn KM, Delaney S. 2014. A chemical and kinetic perspective on base excision repair of DNA. *Acc Chem Res* **47**: 1238–1246.

Schiestl RH, Gietz RD. 1989. High efficiency transformation of intact yeast cells using single stranded nucleic acids as a carrier. *Curr Genet* **16**: 339–346.

Segal AW, Shatwell KP. 1997. The NADPH oxidase of phagocytic leukocytesa. *Ann N Y Acad Sci* **832**: 215–222.

Solomon MJ, Larsen PL, Varshavsky A. 1988. Mapping proteinDNA interactions in vivo with formaldehyde: evidence that histone H4 is retained on a highly transcribed gene. *Cell* **53**: 937–947.

Swanson RL, Morey NJ, Doetsch PW, *et al*. 1999. Overlapping specificities of base excision repair, nucleotide excision repair, recombination, and translesion synthesis pathways for DNA base damage in *Saccharomyces cerevisiae*. *Mol Cell Biol* **19**: 2929–2935.

Wang Y. 2008. Bulky DNA lesions induced by reactive oxygen species. *Chem Res Toxicol* **21**: 276–281.

Wierdl M, Greene CN, Datta A, *et al*. 1996. Destabilization of simple repetitive DNA sequences by transcription in yeast. *Genetics* **143**: 713–721.

Zacher B, Kuan P, Tresch A. 2010. Starr: simple tiling ARRay analysis of Affymetrix ChIP-chip data. *BMC Bioinformat* **11**(1): 194:1–7.

Supporting information

Additional Supporting Information may be found online in the supporting information tab for this article.

Table S1 Supporting information item

Figure S1A Supporting information item

Figure S1B Supporting information item

Figure S1C Supporting information item