

Genome-wide mapping of nuclear mitochondrial DNA sequences links DNA replication origins to chromosomal double-strand break formation in *Schizosaccharomyces pombe*

Sandrine Lenglez,¹ Damien Hermand,² and Anabelle Decottignies^{1,3}

¹Catholic University of Louvain, Faculty of Medicine, de Duve Institute, Genetic and Epigenetic Alterations of the Genome Group (GENEPI), 1200 Brussels, Belgium; ²University of Namur (FUNDP), Molecular Genetics Laboratory (GEMO), 5000 Namur, Belgium

Chromosomal double-strand breaks (DSBs) threaten genome integrity and repair of these lesions is often mutagenic. How and where DSBs are formed is a major question conveniently addressed in simple model organisms like yeast. NUMTs, nuclear DNA sequences of mitochondrial origin, are present in most eukaryotic genomes and probably result from the capture of mitochondrial DNA (mtDNA) fragments into chromosomal breaks. NUMT formation is ongoing and was reported to cause de novo human genetic diseases. Study of NUMTs is likely to contribute to the understanding of naturally occurring chromosomal breaks. We show that *Schizosaccharomyces pombe* NUMTs are exclusively located in noncoding regions with no preference for gene promoters and, when located into promoters, do not affect gene transcription level. Strikingly, most noncoding regions comprising NUMTs are also associated with a DNA replication origin (ORI). Chromatin immunoprecipitation experiments revealed that chromosomal NUMTs are probably not acting as ORI on their own but that mtDNA insertions occurred directly next to ORIs, suggesting that these loci may be prone to DSB formation. Accordingly, induction of excessive DNA replication origin firing, a phenomenon often associated with human tumor formation, resulted in frequent nucleotide deletion events within *ORI3001* subtelomeric chromosomal locus, illustrating a novel aspect of DNA replication-driven genomic instability. How mtDNA is fragmented is another important issue that we addressed by sequencing experimentally induced NUMTs. This highlighted regions of *S. pombe* mtDNA prone to breaking. Together with an analysis of human NUMTs, we propose that these fragile sites in mtDNA may correspond to replication pause sites.

[Supplemental material is available online at <http://www.genome.org>.]

Nuclear insertions of mitochondrial DNA (mtDNA) sequences have been detected in most eukaryotic genomes, from yeast to human (Bensasson et al. 2001) and probably result from the fortuitous capture of mtDNA fragments during double-strand break (DSB) repair (Ricchetti et al. 1999; Leister 2005). In *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*, NUMTs (nuclear DNA sequences of mitochondrial origin) are not only present in the nuclear genome (Ricchetti et al. 1999; Decottignies 2005; Sacerdot et al. 2008) but are also produced de novo during the repair of experimentally induced chromosomal (Ricchetti et al. 1999; Yu and Gabriel 1999) or extrachromosomal DSBs (Decottignies 2005). In human, sequencing of the genome revealed the presence of several hundreds of NUMTs, ranging from tens of bases to more than 14 kb and amounting to 0.01% of the total nuclear genome (Richly and Leister 2004). About one-third of human NUMTs are believed to be associated with mtDNA insertion events, while the other two-thirds probably originated as duplications of preexisting NUMTs (Tourmen et al. 2002; Hazkani-Covo et al. 2003; Pamilo et al. 2007). Strikingly, nuclear insertion of human mtDNA fragments is not only the result of an ancient colonization event but is an ongoing process for which the frequency was estimated at one integration in the germline each 180,000 yr (Ricchetti et al. 2004), and de novo mtDNA insertions were found associated with

human genetic diseases (Willett-Brozick et al. 2001; Borensztajn et al. 2002; Turner et al. 2003). Whether they occurred in evolutionary recent or ancient times, transfer of DNA from mitochondrion to the nucleus has undoubtedly contributed significantly to the evolution of eukaryotic genomes (Leister 2005). A recent study further suggested that NUMTs may contribute to genome integrity maintenance as sealing of chromosomal DSBs by mtDNA prevents nuclear processing of the breaks that could result in otherwise deleterious repair (Hazkani-Covo and Covo 2008). Beyond the impact of mtDNA insertions on genome stability, investigating the mechanisms of NUMT formation offers an opportunity to understand naturally occurring chromosomal DSBs in eukaryotic species. Hence, in this study, a genome-wide analysis of *S. pombe* NUMT insertion sites was carried out to get insight into the mechanisms of DSB formation in chromosomes.

Another unsolved issue concerns the mechanisms of mtDNA fragmentation. Here, we sequenced about 200 experimentally induced *S. pombe* NUMTs to get clues about how and where mtDNA is fragmented. This analysis revealed that some regions of the mtDNA genome are probably more prone to breaking.

Results

S. pombe NUMTs are exclusively detected in noncoding regions of the genome but show no preference for gene promoters

Our computational analysis identified a total of 16 *S. pombe* NUMTs (≥ 25 bp) distributed over 12 insertion sites, with no

³Corresponding author.

E-mail anabelle.decottignies@uclouvain.be.

Article published online before print. Article and publication date are at <http://www.genome.org/cgi/doi/10.1101/gr.104513.109>.

evidence for nuclear duplication (Table 1; Fig. 1A,B; Supplemental Fig. 1). The 12 independent mtDNA insertion events occurred on *S. pombe* chromosomes I and II, but not on chromosome III. The genomic distribution of *S. pombe* NUMTs revealed that they are absent from coding regions that, however, occupy 60% of the total *S. pombe* genome (Fig. 1C; Wood et al. 2002). This indicates that NUMTs are preferentially located in noncoding regions of the *S. pombe* genome (Fisher's exact probability test, $P = 0.002$). Similarly, the 26 independent insertions of mtDNA into *S. cerevisiae* chromosomes (Ricchetti et al. 1999) are mainly located in noncoding regions of the genome (Supplemental Fig. 2), and this genomic distribution is significantly different from the expected one (Fisher's exact probability test, $P = 1.1 \times 10^{-6}$).

Previously, retrotransposon insertion mechanisms were elucidated through the analysis of their genomic insertion sites, and this revealed the existence of an active mechanism of insertion (Chalker and Sandmeyer 1992; Yieh et al. 2000). Although NUMTs are unlikely to be the result of active insertions, we anticipated that a careful analysis of their genomic insertion sites may help to get more insight into NUMT insertion mechanisms. We therefore classified the noncoding regions comprising *S. pombe* NUMTs according to the orientation of the flanking pairs of genes as tandem, divergent, or convergent (Fig. 2A). No decrease in the frequency of NUMT insertions into noncoding regions between convergent pairs of genes was observed, suggesting that *S. pombe* NUMTs are not preferentially located into gene promoters (Fig. 2A). This observation was further supported by the similarity between NUMT insertion site distribution and the expected distribution calculated from promoter and 3' region sizes of the *S. pombe* genome (Fig. 2B). Conversely, retrotransposon insertion sites described previously (Behrens et al. 2000) clustered in a narrow window upstream of the ATG, in agreement with RNA polymerase-dependent insertion of these transposable elements. Despite the fact that NUMTs are not preferentially inserted into gene promoters, five independent NUMT integrations occurred into intergenic regions associated with divergent pairs of genes that therefore comprise gene promoters. To test the impact of *NUMT14* and *NUMT4/5/6* on transcription level of, respectively, *SPAC1F12.02c/p23fy* and *SPBC3B9.17/isa2* genes, we constructed plasmids comprising either WT or NUMT-deleted promoters upstream of the *lacZ* reporter gene (Fig. 2C). Constructs were then integrated into their respective chromosomal contexts, and beta-

galactosidase activities were measured. We found that *NUMT14* deletion had no effect on *SPAC1F12.02c* gene expression level (Fig. 2C). Similarly, deleting either *NUMT4/5* or *NUMT6* did not affect the transcription level of the *SPBC3B9.17* gene (Fig. 2C). These data suggest that *S. pombe* NUMTs located into promoter regions do not affect gene transcription level.

S. pombe NUMTs are enriched in chromosomal loci containing a DNA replication origin

Next, a comparison with the genomic position of the 557 *S. pombe* DNA replication origins identified in two distinct studies (Heichinger et al. 2006; Hayashi et al. 2007) revealed that 92% (11/12) of NUMT insertion sites are located within intergenic regions associated with an ORI. This value is significantly different from the expected one (Fisher's exact probability test, $P = 0.0003$) since, at most, 11% of *S. pombe* intergenic regions comprise a DNA replication origin (557 ORIs spread over 5000 intergenic regions) (Fig. 3A–C). Conversely, and in agreement with distinct mechanisms of integration, the 26 de novo insertion sites of the long terminal repeat-containing retrotransposon *Tf1* (transposon of fission yeast *1*) identified previously (Behrens et al. 2000) showed an expected frequency of association with *S. pombe* predicted ORIs (Heichinger et al. 2006; Hayashi et al. 2007) (Fig. 3A). A fivefold enrichment of NUMTs next to ORIs was also observed in *S. cerevisiae* genome. Indeed, 35% of budding yeast NUMTs are located within 600 bp of a DNA replication origin although the expected frequency amounts to only 7% (429 ORIs/6286 intergenic regions) (Supplemental Fig. 3).

The striking colocalization of NUMTs and ORIs that we observed in the *S. pombe* genome may be explained in two ways: either NUMTs act as chromosomal ORIs or mtDNA fragments were preferentially inserted next to chromosomal DNA replication origins. To test whether NUMTs act as ORIs in *S. pombe* chromosomes, we checked the binding of origin recognition complex protein Orc1 (also known as Orp1) over two chromosomal regions comprising double NUMT insertions: *NUMT10/11* (70 and 31 bp) and *NUMT2/3* (311 and 33 bp). Chromatin immunoprecipitation experiments using antibodies against Orc1-Flag were performed in *orp1-5flag* cells (Hayashi et al. 2007), followed by PCR amplifications of consecutive ~300-bp fragments along these two chromosomal loci. PCR products obtained from the immunoprecipitated

chromatin were quantified and normalized with the amplifications obtained from total chromatin extracts (Fig. 4A). As positive control, we screened *ORI2050/ARS2004* genomic locus and detected Orc1 binding onto fragments G and H, as expected (Wuarin et al. 2002). Orc2 (also known as Orp2) and Cdc21 (also known as Mcm4) were also reported to bind to these two fragments (Wuarin et al. 2002). At *ORI3007*, we detected Orc1 binding over fragments A to C, in agreement with Orc1/Mcm6 chromatin immunoprecipitation (ChIP) profiles obtained by Hayashi et al. (2007) (Fig. 4A; Supplemental Fig. 4). Screening of a 2.9-kb chromosomal region encompassing both *NUMT10/11* and predicted *ORI2038* revealed that Orc1 binds mainly to fragment D (Fig. 4A). Strikingly, although *NUMT10/11* is present in fragment E, the strongest Orc1 binding was detected in

Table 1. *S. pombe* NUMTs

	Size (bp)	Chromosome	Chromosome coordinates	Mitochondrial genome coordinates	Identity (%)
<i>NUMT1</i>	364	II	36900–37263	7223–7580	95
<i>NUMT2</i>	311	I	2667777–2668087	2557–2871	89
<i>NUMT3</i>	33	I	2667618–2667650	6064–6031	94
<i>NUMT4</i>	175	II	4019667–4019841	13800–13622	84
<i>NUMT5</i>	65	II	4019842–4019906	15266–15201	92
<i>NUMT6</i>	115	II	4020012–4020126	15681–15795	81
<i>NUMT7</i>	148	II	3360977–3361124	11514–11665	92
<i>NUMT8</i>	89	I	72348–72436	2598–2686	92
<i>NUMT9</i>	85	I	79544–79628	4377–4461	94
<i>NUMT10</i>	70	II	1244073–1244142	15512–15441	88
<i>NUMT11</i>	31	II	1244144–1244174	16070–16100	93
<i>NUMT12</i>	62	II	636435–636496	15470–15531	87
<i>NUMT13</i>	59	I	1487018–1487076	1597–1539	94
<i>NUMT14</i>	56	I	3810379–3810434	18833–18778	100
<i>NUMT15</i>	55	II	2433877–2433931	9702–9756	96
<i>NUMT16</i>	25	I	4263907–4263931	7302–7326	92

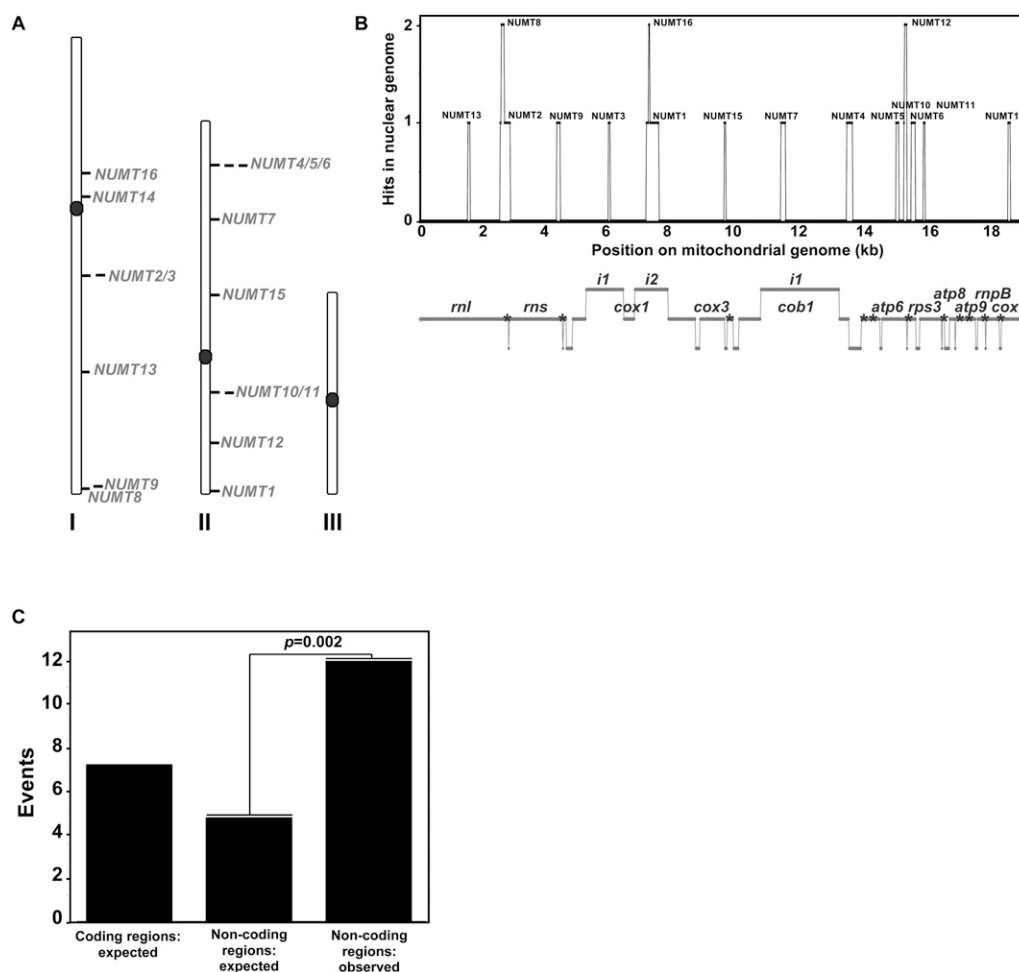


Figure 1. The 16 *S. pombe* NUMTs are distributed over noncoding regions of chromosomes I and II. (A) Chromosomal mapping of *S. pombe* single and multiple NUMTs. The total number of independent NUMT insertion sites is estimated at 12. Position of the centromeres is indicated. (B) *S. pombe* NUMTs are not duplicated. Position of NUMTs is shown on the linear map of *S. pombe* mitochondrial genome (bottom panel). The graph represents the number of hits in the nuclear genome for the entire mtDNA sequence. (Bottom panel) mtDNA genes are indicated on mtDNA map; i1 and i2 refer to introns; intergenic regions of mtDNA genome are indicated by gray bars downstream from the genes; asterisks indicate the position of tRNAs. (C) The 12 independent NUMT insertion sites from *S. pombe* were classified according to their genomic localization: coding or noncoding regions. NUMT distribution was compared with the expected frequency based on the contribution of coding and noncoding DNA sequences in *S. pombe* genome. Fisher's exact probability test was performed on <http://faculty.vassar.edu/lowry/odds2x2.html>.

fragment D, located directly upstream of *NUMT10/11*. These data suggest that *NUMT10/11* is probably not acting as ORI on its own. Similar observations were made for the 3.4-kb region encompassing both *NUMT2/3* and predicted *ORI1081*. In the latter case, although Orc1 binding was weak, the peak was detected in fragment D, directly upstream of *NUMT2* (Fig. 4A). The Orc1 ChIP profiles we obtained for *NUMT10/11* (*ORI2038*) and *NUMT2/3* (*ORI1081*) chromosomal loci also match the Orc1/Mcm6 binding site maps reported by Hayashi et al. (2007) (Supplemental Fig. 4). To control for the absence of a putative bias due to the overrepresentation of DNA coming from highly transcribed genes in the ChIP DNA preparation as previously suggested (Giresi et al. 2007), we PCR-amplified the constitutively expressed *cdc2* gene in both total and immunoprecipitated chromatin extracts. We found no enrichment in the latter fraction (Supplemental Fig. 5).

To test further our hypothesis that chromosomal NUMTs result from the insertion of mtDNA fragments directly next to ORIs, rather than acting as ORIs themselves, we deleted *NUMT10/11* and

probed the chromosomal locus again for Orc1 binding. *NUMT10/11* deletion was performed in a two-step process so that the marker used for selection of the integrants was then removed, leaving only four nucleotides (see Methods). Strikingly, we found that Orc1 binding to fragment D in the *ORI2038* locus was not affected by deletion of *NUMT10/11* (Fig. 4B), thereby supporting our hypothesis. Note that, although at the two loci tested, the nucleotide sequence of chromosomal NUMTs are unlikely to ensure chromosomal replication origin firing, it is reasonable to believe that insertion of longer mtDNA fragments may create new ORIs in chromosomes as the A+T content of *S. pombe* mitochondrial genome amounts to 70% (Bullerwell et al. 2003), a value close to the nucleotide content of DNA replication origins in *S. pombe* chromosomes which consist of 0.5- to 1-kb-long A+T-rich regions (Segurado et al. 2003). In this respect, we showed that, in an extrachromosomal context, 0.4- to >1-kb-long mtDNA fragments were able to act as efficient replication origins for episomes (Supplemental Fig. 6).

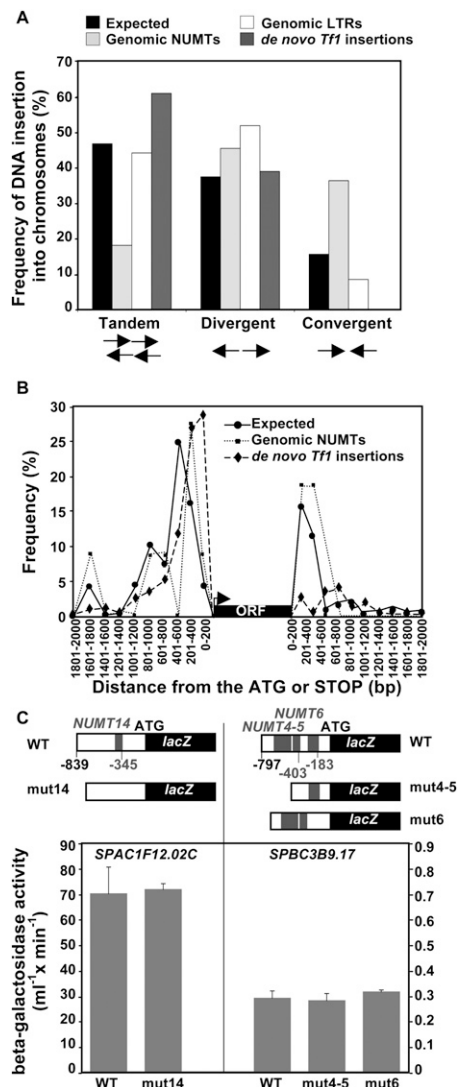


Figure 2. *S. pombe* NUMTs are not enriched into gene promoters. (A) Insertion sites of *S. pombe* NUMTs (pale gray), genomic LTRs (white; based on Bowen et al. 2003) and de novo *Tfi1* (dark gray; based on Behrens et al. 2000 and Singleton and Levin 2002) were classified into three categories of intergenic regions based on transcription direction of the genes flanking the insertion site: tandem, divergent, or convergent. Expected distribution for *S. pombe* (black) was calculated from http://www.sanger.ac.uk/Projects/S_pombe/genome_stats.shtml and is expressed as percentage of total noncoding DNA. (B) (■) Distribution of the distance from each independent NUMT insertion site to the nearest ORF. All NUMTs from intergenic regions between tandem gene pairs were located 5' of the nearest gene. (◆) Distribution of the 26 intergenic de novo *Tfi1* insertion sites described in Behrens et al. (2000). (●) Expected distribution of insertion sites for *S. pombe* genome was obtained by calculating the size of promoters and 3' regions from a total of 92 intergenic regions in either divergent or convergent pairs of genes from all three chromosomes. The insertion sites were binned in regions of 200 bp in windows spreading from -2000 to ATG and from STOP to +2000. (C) Transcriptional activity of *SPAC1F12.02C* and *SPBC3B9.17* promoters was performed as described in Methods. The effect of NUMT(s) in these promoters was assessed by measurement of beta-galactosidase activity driven by mutated promoters lacking the NUMT(s). (Gray boxes) NUMT positions. Error bars indicate standard deviations.

Excessive firing of DNA replication origins induces Rad22-YFP foci next to ORIs

The hypothesis that NUMT insertions may be favored near DNA replication origins postulates that regions surrounding ORIs are prone to DSB formation. This idea fits with the proposal that, in *S. cerevisiae*, prereplication complex binding at ORIs may slow down the progression of a replication fork starting from a nearby origin, a process possibly associated with an increased risk of fork collapse and DSB formation (Tourriere and Pasero 2007).

To test the hypothesis that prereplication complex binding at *S. pombe* origins may increase the risk of DSB formation at these loci during S phase, we followed Rad22-YFP foci formation, a marker of DSB (Meister et al. 2003, 2005), in cells experiencing excessive firing of DNA replication origins (Rad22 is the homolog of RAD52 in other species, including *S. cerevisiae* and human). This was achieved by overexpression of DNA replication initiation factor *cdc18* (the *S. pombe* *CDC6* homolog) from the strong regulatable *nmt1* promoter (pREP3X::*cdc18*) which is repressed by the addition of thiamin in the culture medium. Overexpression of *cdc18* is known to induce DNA rereplication (Nishitani and Nurse 1995), DNA replication checkpoint activation, and cell cycle arrest (Greenwood et al. 1998; Fersht et al. 2007). Accordingly, incubation of cells in the absence of thiamin resulted in increased DNA content and cell elongation (Fig. 5A–C). Bright Rad22-YFP foci were detected in elongated cells (Fig. 5A), suggesting that excessive firing of DNA replication origins was indeed associated with DNA break formation. To see whether these Rad22-YFP foci are formed next to origins, we overexpressed *cdc18*-CFP from *nmt1* promoter (pREP5::*cdc18*-CFP) in cells expressing Rad22-YFP (Fig. 5C). Although fluorescence of overexpressed Cdc18-CFP was often very bright and distributed throughout the *S. pombe* nucleus (Fig. 5C,h), we could detect Cdc18-CFP foci in a fraction of cells (Fig. 5C,b,e) that, presumably, corresponded to Cdc18-CFP proteins bound to ORIs. In such cells, about two-thirds (27/48) of Cdc18-CFP dots were found to colocalize with Rad22-YFP foci (Fig. 5C,c,f). Hence, these data suggest that, in conditions of excessive origin firing, DSBs arise frequently in ORI regions of *S. pombe* nucleus.

Although capture of mtDNA fragments within ORI regions may be achieved through nonhomologous end-joining (NHEJ), it is also likely that Rad22-dependent microhomology-mediated end-joining (MMEJ) (Decottignies 2007) may drive mtDNA insertions. Indeed, we showed earlier that NHEJ-deficient *S. pombe* cells are able to capture extrachromosomal DNA fragments within *ura4* episomes providing that microhomologies between extrachromosomal DNA and *ura4* gene are present at both ends (Decottignies 2005). Hence, we postulate that recruitment of Rad22 at ORIs may ensure MMEJ-dependent capture of mtDNA.

Replication stress induces chromosomal deletions at subtelomeric *ORI300I* locus

Our analysis suggested that ORI chromosomal loci may be prone to DSB formation. In particular, *NUMT1*, *NUMT8*, and *NUMT9* insertion sites were not only located within ~1 kb of a predicted ORI but were also found at very distal positions on chromosomes I or II (37, 72, and 79 kb from chromosome ends, respectively) (Fig. 3B,C), suggesting that terminal ORI regions of chromosomes may be even more subjected to replication stress-induced DSBs. To assess DSB formation at distinct *S. pombe* chromosomal loci, we set up an experimental system based on the measurement of *ura4* gene

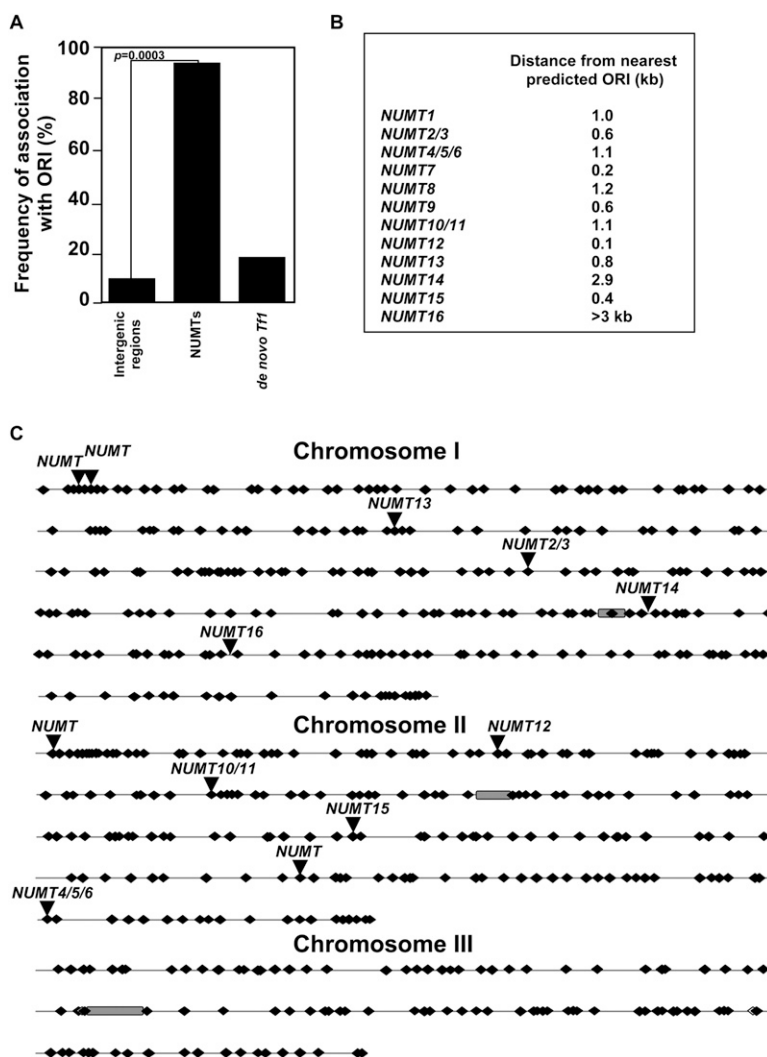


Figure 3. *S. pombe* NUMTs are preferentially located next to DNA replication origins. (A) Based on two distinct studies (Heichinger et al. 2006; Hayashi et al. 2007), the number of *S. pombe* DNA replication origins may be close to 557, leading to a frequency of replication origin-containing intergenic regions of ~11% (557/5000). Eleven out of the 12 independent NUMT insertion sites are located in the proximity of a prereplication complex binding site. For comparison, the 26 de novo *Tfi* insertion sites identified in another study (Behrens et al. 2000) show an expected frequency of association with DNA replication origins. Fisher's exact probability test was performed on <http://faculty.vassar.edu/lowry/odds2x2.html>. (B) Distance between each NUMT and the nearest predicted ORI (kb). (C) Replication origins deduced from Hayashi et al. (2007) and Heichinger et al. (2006) studies are represented on *S. pombe* chromosomes. For clarity reasons, diamonds do not reflect properly the size of ORIs, as replication origins cover at most 4.4% of the total *S. pombe* genome. NUMT16 is not associated with an ORI. (Gray rectangles) Centromeres.

mutation frequency. To this end, we inserted the *ura4* gene into five distinct yeast strains either directly next to a DNA replication origin (two strains) or far away from an ORI (three strains) (Fig. 6A; Supplemental Fig. 7). These included the *SPCC1884.01Δ::ura4* strain with an insertion of *ura4* at the *ORI3001* locus located 37 kb away from the left telomere of chromosome III. An additional strain with a WT endogenous *ura4* gene was also included in the analysis, as this locus is located directly next to *ORI3006* DNA replication origin in *S. pombe* genome (Fig. 6A; Supplemental Fig. 7). The six yeast strains were transformed with the inducible *nmt1::cdc18* construct, grown in either the presence (*nmt1* promoter repressed) or the absence (*nmt1* promoter induced) of thia-

min and then plated onto 5-fluoro-orotic acid (5-FOA) (Boeke et al. 1984). Yeast colonies auxotrophic for uracil were next screened for *ura4* mutations.

Strikingly, we found that, although the frequency of *ura*⁻ colonies formation was <10⁻⁸ in cells grown in repressive conditions (+thiamin), it increased to 1.4 × 10⁻⁶ in the *SPCC1884.01Δ::ura4* strain following *cdc18* induction (Fig. 6B). However, this was not the case for the other yeast strains that we tested, whether the *ura4* gene was located next to an ORI or not (Fig. 6B). This suggested that terminal regions of chromosomes, like *ORI3001*, may indeed be more subjected to replication stress-induced mutation. To get insight into the mutational events that gave rise to *ura*⁻ colonies in *SPCC1884.01Δ::ura4* cells and, in particular, to see whether DSBs may have been involved, we tried to PCR amplify the mutated *ura4* gene in *ura*⁻ mutants. Interestingly, we failed to amplify the *ura4* cassette in 55% (34/62 tested) of *ura*⁻ colonies recovered from *SPCC1884.01Δ::ura4* cells incubated without thiamin (Fig. 6C). Further PCR amplifications confirmed the deletion of the *ura4* gene in these cells (Fig. 6D). We analyzed a total of 24 clones with *ura4* deletion and classified them into four groups according to the extent of genomic DNA deletion (Fig. 6D). Strikingly, in group 1, genomic DNA deletions were found to extend over more than 10 kb. On the other hand, the *ura*⁻ colonies recovered from the other yeast strains were not attributable to *ura4* gene deletions as shown by nucleotide sequencing of the *ura4* PCR products that we obtained (data not shown).

Altogether, these data suggest that, in replication stress conditions induced by overexpression of DNA replication initiation factor Cdc18, some chromosomal loci, like subtelomeres, may be subjected to high frequency of DSB formation and nucleotide loss.

Mitochondrial intron DNA splicing and tRNA stretches may be two major causes of mtDNA fragmentation

Another goal of this study was to address the mechanisms of mtDNA fragmentation through nucleotide sequence analysis of mtDNA fragments captured at nuclear DSBs. The number of natural *S. pombe* NUMTs being too small for this kind of analysis, we relied on the *ura4* system of extrachromosomal DSB repair to capture mtDNA fragments produced de novo (Decottignies 2005). A total of 195 mtDNA inserts (83–4004 bp) were recovered in *ura4* circles. Sequence analysis of these inserts revealed an enrichment of sequences originating either from *cox1* gene introns or from the

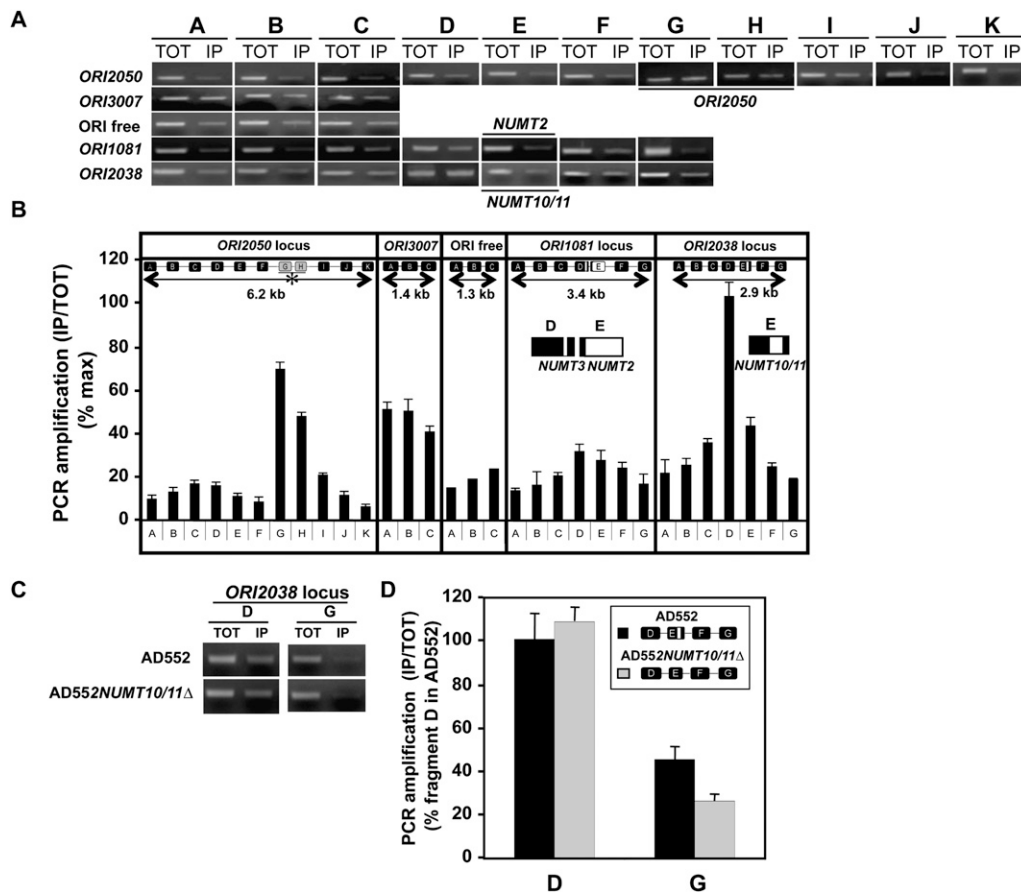


Figure 4. NUMTs are probably not acting as replication origins on their own in *S. pombe* chromosomes. (A) Lysates from *orp1::orp1-5flag* cells (strain AD552) were subjected to ChIP analysis using anti-Flag antibodies. Total chromatin (TOT) or DNA associated with immunoprecipitates (IP) was amplified using primer sets that amplify consecutive ~300-bp fragments along *ORI2050* (ARS2004) (Wuarin et al. 2002) (chromosome II: 1541299–1547496), *ORI3007* (chromosome III: 122883–124270), an ORI-free region (chromosome II: 3034667–3035885), *ORI1081*-NUMT2/3 (chromosome I: 2665851–2669206), and *ORI2038*-NUMT10/11 (chromosome II: 1242241–1245130). The nomenclature used for ORIs is based on Heichinger et al. (2006). (B) The size of the five regions analyzed and the position of PCR fragments are indicated in each case. PCR reaction products were run on gel and quantified with Image Station 440 CF (Kodak). Band intensity of IP was normalized with TOT for each PCR fragment to give IP/TOT values (100% corresponds to IP/TOT value measured for fragment D of *ORI2038* locus). Screening of the *ORI2050* genomic locus was used as positive control and revealed Orc1 binding onto fragments G and H (*), as expected (Wuarin et al. 2002). (C,D) The ChIP experiment was repeated in strains AD552 and AD552NUMT10/11Δ (AD566) to test the impact of NUMT10/11 deletion on Orc1 binding onto fragments D and G of *ORI2038* locus described in panel A. Data result from two independent experiments and measurements were performed in triplicate. Error bars indicate standard deviations.

tRNA-rich regions of the mitochondrial genome (between *cob1* and *atp9* genes) (Fig. 7A). This enrichment of mtDNA fragments originating from *cox1* introns or tRNA-rich regions was not due to sequence homologies with *ura4* DNA. We propose that the process of DNA splicing operating at the level of *S. pombe* mitochondrial introns (Schafer et al. 1991) may produce high amounts of mtDNA linear fragments from *cox1* introns that are good substrates for capture at nuclear DSBs. Alternatively, it is possible that these regions may be preferentially reverse transcribed and transferred to the nucleus similarly to what was observed for the transfer of the flowering plant *coxII* gene following the apparent reverse transcription of an edited RNA intermediate (Nugent and Palmer 1991).

On the other hand, we propose that the enrichment of mtDNA sequences originating from tRNA-rich regions may be related to the fact that tRNA genes represent natural pause sites for replication forks and are presumably more prone to DSB formation (Ivessa et al. 2003; Shimada et al. 2008). In agreement with this, it was previously shown that the tRNA^{Leu(UUR)} gene of the human mtDNA genome, bound by mitochondrial transcription termina-

tion factor mTERF, is a replication pause site (Hyvärinen et al. 2007). Accordingly, an analysis of 37 human-specific NUMTs (Ricchetti et al. 2004) revealed the presence of an mtDNA breakage site at this tRNA locus (Fig. 7B). Moreover, we detected additional breakage sites at the origin of heavy-strand synthesis (O_H), the origin of light-strand synthesis (O_L), and alternative O_L (alt O_L). As O_H and O_L were recently shown to be associated with replication pausing and breakage (Bailey et al. 2009), this analysis of human NUMTs provided additional evidences in favor of mtDNA replication pausing as a factor contributing to mtDNA breakage and NUMT formation.

Discussion

Fragments of mtDNA can escape from mitochondria and get inserted into nuclear DSBs, giving rise to NUMTs. This phenomenon has been reported in many different eukaryotic species, including human, and is believed to act as a mutagenic process able to cause de novo genetic diseases. Here, we analyzed the genomic

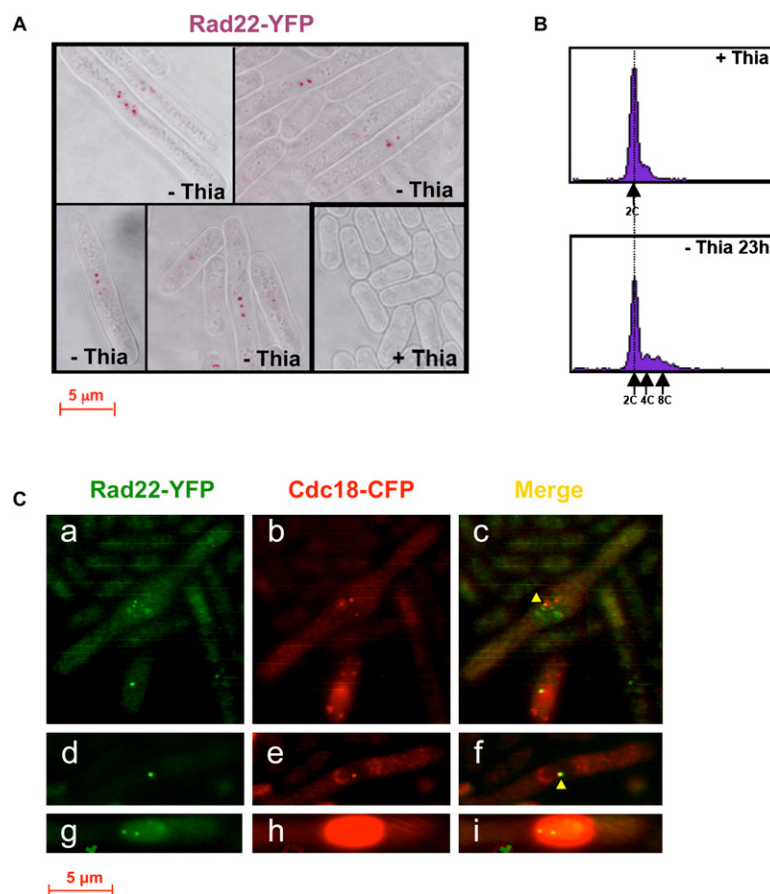


Figure 5. Overexpressed Cdc18-CFP foci are associated with Rad22-YFP. (A) The pREP3x::cdc18 plasmid was introduced into yeast cells expressing the Rad22-YFP fusion. Overexpression of the *cdc18* gene was induced by growing cells for 23 h in the absence of thiamin, and Rad22-YFP foci were visualized by epifluorescence microscopy. Cells grown in the presence of thiamin are shown as controls. (B) Genomic DNA content of the above cells grown in the presence or the absence of thiamin was analyzed by FACS. (C) Observation of Cdc18-CFP was achieved by introducing pREP5::cdc18-CFP plasmid in cells described in A and processed similarly. YFP- (a,d,g) and CFP-tagged (b,e,h) proteins were observed in separate channels as described in Methods, and images were merged with Adobe Photoshop (c,f,i). Representative pictures are shown.

distribution of *S. pombe* NUMTs to get hints into the genomic targeting of mtDNA fragments and, more generally, DSB formation in chromosomes.

Our computational analysis identified 16 *S. pombe* NUMTs distributed over 12 genomic loci. These NUMTs are not duplicated and the 12 genomic loci are therefore likely to represent independent integration events. Conversely, two NUMT loci are clearly duplicated in the nuclear genome of *S. cerevisiae* (data not shown). In human or honeybee, the ratio of independent insertions to nuclear duplications of preexisting NUMTs is thought to be close to 1:2 while, in some rodent species, like *M. rossi* and *M. merionalis*, colonization of the nuclear genome was limited to a few NUMTs that were subsequently subjected to multiple nuclear duplication events (Tourmen et al. 2002; Hazkani-Covo et al. 2003; Pamilo et al. 2007; Triant and DeWoody 2007). Like in many other species, insertions of multiple mtDNA fragments within the same chromosomal locus were detected in *S. pombe* genome. We found that multiple insertions account for 25% of the integration events. For comparison, multiple insertions account for 11% of the independent integration events in *S. cerevisiae* and 7% in human (Ricchetti et al. 1999, 2004). Multiple mtDNA insertions were

also frequently recovered during the repair of experimentally induced DSBs (Ricchetti et al. 1999; Decottignies 2005), suggesting that the limiting step for NUMT insertion may be the formation of chromosomal DSBs rather than the production of mtDNA fragments.

Chromosome III of *S. pombe* was found to be devoided of NUMT. As this latter chromosome is about half the size of chromosome I or II, two or three mtDNA insertions were expected. This distribution of insertion sites is clearly distinct from that observed for retrotransposons, for which the frequency is nearly twofold higher than expected on chromosome III (Behrens et al. 2000; Singleton and Levin 2002; Bowen et al. 2003). Bowen et al. (2003) suggested that 80% of the enrichment of LTRs on chromosome III may be due to their preferential insertion into intergenic sequences flanking *S. pombe*-specific *wtf* (with *Tf*) sequences that belong to a high-copy family of 1-kb-long sequences strongly enriched on chromosome III. It was proposed that this biased distribution of *wtf* sequences in the original 972 strain may be the result of acquisition of chromosome III from an isolated population of *S. pombe* that had *wtf*s distributed on all three chromosomes (Bowen et al. 2003). If chromosome III of strain 972 has indeed originated from another *S. pombe* population, this may explain the absence of detectable NUMT on chromosome III as, based on a previous comparison of three *S. pombe* mitochondrial genomes (Bullerwell et al. 2003), the mtDNA sequences of these two populations may be very different.

In the present study, we addressed the mechanisms of NUMT formation through the analysis of their genomic distribution. This revealed an exclusive location of NUMTs within noncoding regions of the genome. There was no preference for gene promoters and our analysis suggested that, unlike retrotransposons, NUMTs are unlikely to rely on RNA polymerase for their insertion. Besides, when present in gene promoters, we showed that NUMTs do not affect gene expression level. Although the enrichment of *S. pombe* NUMTs into noncoding regions is significant (Fisher's exact probability test, $P = 0.002$), it is expected that NUMT insertion into gene coding regions would be selected against if deleterious. A recent genome-wide study revealed that ~26% of *S. pombe* genes are essential for growth in experimental conditions (Kim et al. 2010), but it is difficult to evaluate the long-term impact of mtDNA insertion into coding regions of "nonessential" genes in the wild. Nevertheless, we believe that the striking association between NUMTs and ORIs that we unraveled in this study is likely to account for the enrichment of NUMTs in noncoding regions of *S. pombe* genome. Indeed, we found that 11 of the 12 independent NUMT insertion sites are located in noncoding regions associated with an ORI, a value that is significantly different from the

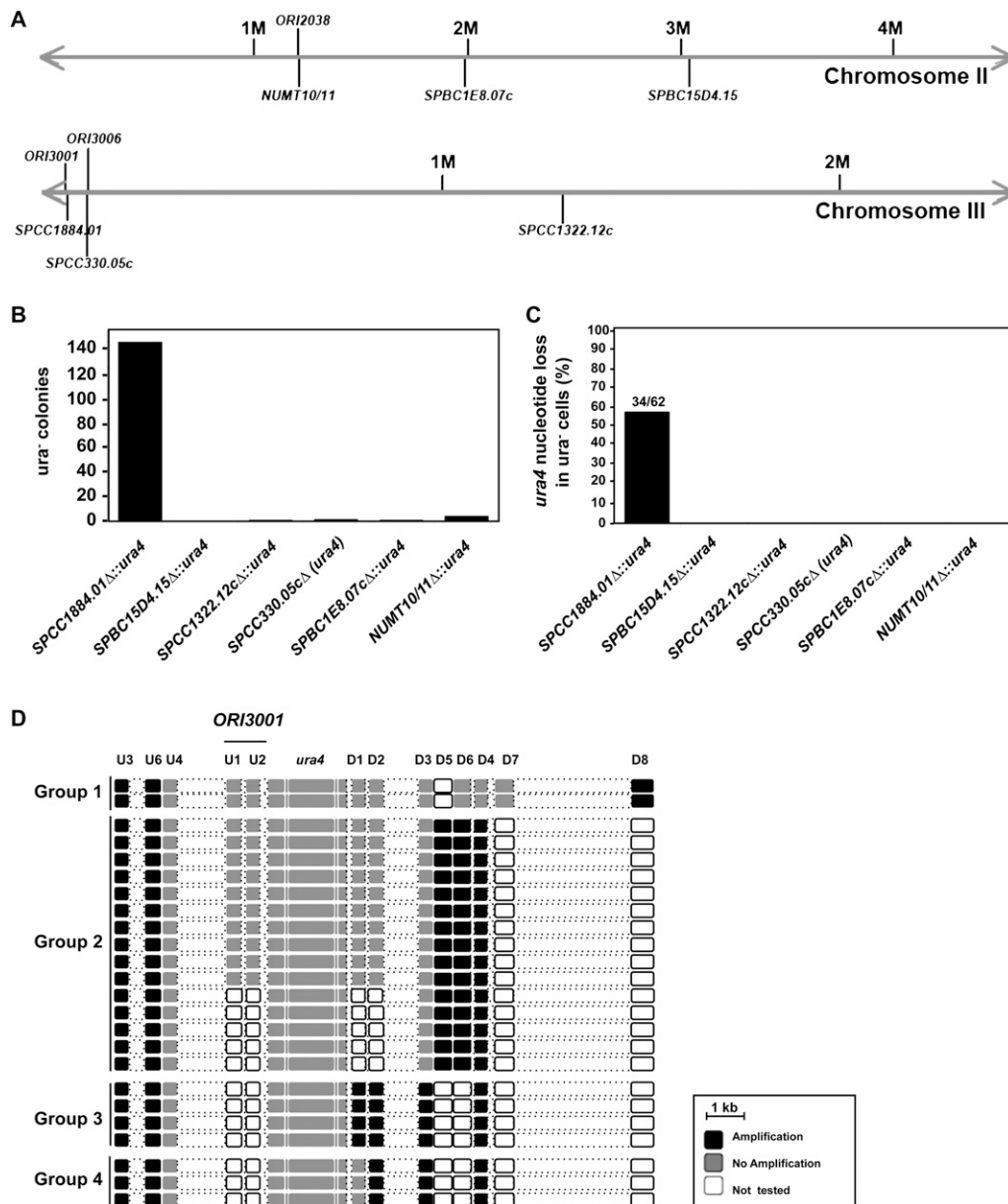


Figure 6. Replication stress induces frequent nucleotide deletion events at subtelomeric *ORI3001* locus. (A) An expression cassette comprising the *ura4* gene was inserted at five distinct chromosomal loci in *S. pombe* to give AD549 (*SPCC1884.01Δ::ura4*), AD496 (*SPBC1E8.07cΔ::ura4*), AD555 (*NUMT10/11Δ::ura4*), PN4095 (*SPCC1322.12cΔ::ura4*), and AD421 (*SPBC15D4.15Δ::ura4*) strains. AD467 strain has an endogenous WT *ura4* gene (*SPCC330.05c*). Positions of *ORI3001*, *ORI3006*, and *ORI2038* were deduced from the peaks of Mcm6 binding onto chromatin obtained by Hayashi et al. (2007) (Supplemental Fig. 7) and the data from Heichinger et al. (2006). (B) pREP3x::*cdc18* plasmid was introduced into all six *S. pombe* strains. *Cdc18* overexpression was achieved by incubating cells for 19 h at 32°C in the absence of thiamin. Cells (0.5×10^9) were then plated onto 5-FOA-containing EMM2 agar plates (50×10^6 cells/plate) and the 5-FOA^R colonies that were recovered after 4–6 d were tested for uracil auxotrophy. The graph gives the number of ura⁻ colonies that were obtained for each strain. (C) The ura⁻ colonies recovered in B were tested for the presence of the *ura4* gene by PCR screen. (D) The extent of genomic DNA deletion was analyzed by PCR in 24 colonies with the *ura4* deletion derived from the *SPCC1884.01Δ::ura4* strain. Each box represents a PCR product (see Supplemental Table III for primers). In ura⁻ mutants, PCR products were either recovered (black boxes) or not (gray boxes). White boxes indicate regions that were not tested by PCR. The mutants were classified into four groups according to the extent of genomic DNA deletion. The position of *ORI3001* is indicated.

expected frequency (Fisher's exact probability test, $P = 0.0003$). Two different hypotheses could explain the colocalization of NUMTs and ORIs: (1) NUMTs act as chromosomal ORIs or (2) mtDNA insertions were favored next to ORIs during evolution.

Chromatin immunoprecipitation experiments on the intergenic region comprising both *NUMT10/11* and predicted *ORI2038* revealed that the peak of Orc1 binding was located directly 5' of

NUMT10/11. Similar conclusions were drawn from the analysis of the *NUMT2/3* chromosomal locus associated with predicted *ORI1081*. Moreover, removal of *NUMT10/11* did not change the binding efficiency of Orc1 at *ORI2038*. Hence, we propose that mtDNA fragments inserted next to preexisting ORIs in the *S. pombe* nucleus. Although we believe that the NUMTs present in *S. pombe* chromosomes are unlikely to act as chromosomal ORIs

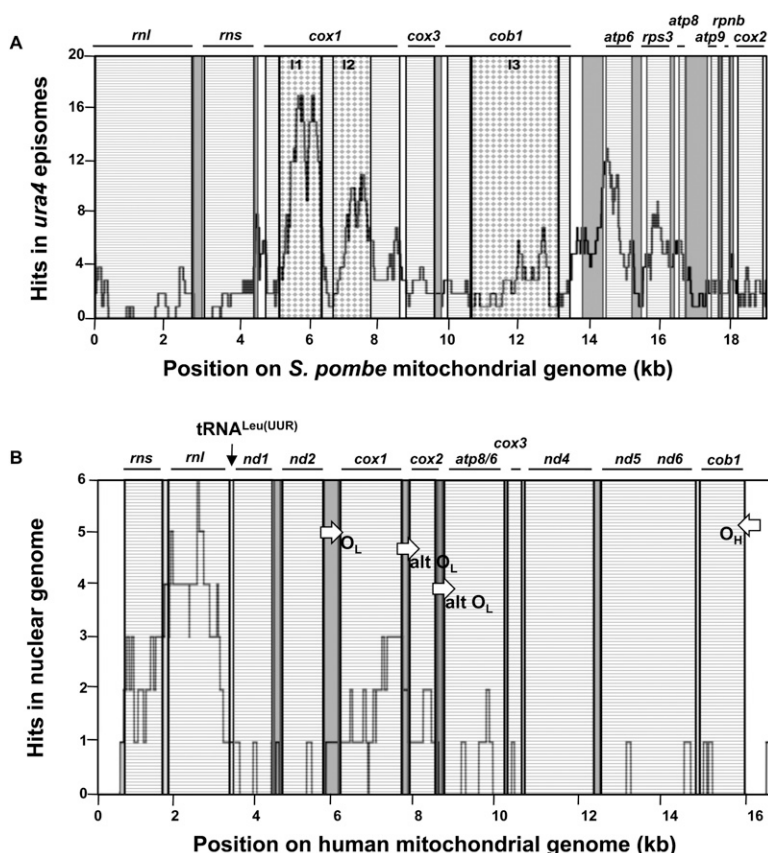


Figure 7. Fragile sites are present in mitochondrial genomes of *S. pombe* and human. (A) Distribution of *S. pombe* mtDNA sequences at repair junctions of circularized *ura4* fragments. One-hundred-ninety-five mtDNA inserts (83–4004 bp) recovered in *ura4* episomes were sequenced. The graph gives the number of times each nucleotide from the *S. pombe* mitochondrial genome was recovered in episomes. (B) Distribution of mtDNA nucleotides in the 37 human-specific NUMTs described previously (Ricchetti et al. 2004). The graph gives the number of times each nucleotide from the human mitochondrial genome was recovered in human-specific NUMTs. (O_H) Heavy-strand synthesis origin; (O_L) light-strand synthesis origin; (alt O_L) alternative O_L . In A and B: (dashed boxes) genes; (gray) tRNAs; (white) intergenic regions. Draughtboards indicate gene introns in *S. pombe* mitochondrial genome. Position of tRNA^{Leu(UUR)} is indicated on human mitochondrial DNA map.

on their own, we do not rule out the possibility that insertion of mtDNA fragments into chromosomes may bring new ORIs as mitochondrial genome and DNA replication origins display similar (A+T)-rich content in *S. pombe* (Segurado et al. 2003). Accordingly, we found that mtDNA fragments of ≥ 0.4 kb are sufficient to mediate the replication of extrachromosomal circular molecules of DNA. These data are in agreement with a previous study reporting that the budding yeast mitochondrial genome contains over 40 ARS elements that are able to ensure nuclear replication of a plasmid (Hyman et al. 1983). These mtDNA fragments may, however, not be able to act as chromosomal ORIs, as it was previously reported that, although more than 50% of *S. pombe* intergenic regions have the propensity to function as origins on extrachromosomal plasmids, only part of them are active in a chromosomal context (Dai et al. 2005). Nevertheless, it remains possible that chromosomal insertion of big mtDNA fragments may be able, in some circumstances, to create new chromosomal ORIs and/or to contribute to the evolution of complex eukaryotic replication origins.

It has been previously reported that excessive origin firing leads to DNA lesions in both *S. pombe* and *S. cerevisiae* (Fersht et al. 2007; Ide et al. 2007). Although the mapping of DSBs has not

been established, a genome-wide analysis revealed that subtelomeric regions have a higher probability of rereplicating in Cdc6-overexpressing *S. cerevisiae* cells (Tanny et al. 2006). Here, we performed two separate experiments that unraveled a link between ORI regions in chromosomes and DSB formation. First, we showed that some of the bright foci of overexpressed DNA replication initiation factor Cdc18-CFP (the CDC6 homolog) colocalized with Rad22-YFP (the RAD52 homolog) dots that mark DNA breaks. Second, we showed that the frequency of *ura4* nucleotide loss reaches 0.8×10^{-6} in cells with a *ura4* cassette inserted within 0.5–1 kb of subtelomeric ORI3001 DNA replication origin and subjected to excessive firing of the ORI through Cdc18 overexpression. The chromosomal deletions that we recovered extended over 10 kb in some mutants. These results support the recent observation that replication stress induces tumor-like microdeletions in human cells (Durkin et al. 2008). On the other hand, using our experimental system, we did not detect any *ura4* gene loss when the reporter gene was present in the vicinity of two other *S. pombe* ORIs, suggesting that some ORIs, like subtelomeric ones, may be more subjected to breakage than others. As expected, no *ura4* gene deletion was detected either when the gene was located far away from an origin, even in conditions of excessive origin firing. These observations, however, do not rule out the possibility that, in an evolutionary context, over years, chromosomal breakage may occur more frequently at any ORI. In support of our data, a recent

study identified the chromosomal region surrounding budding yeast *ARS106* as a locus associated with fragility (γ H2A peak) in chromatin isolated from *top2-1* topoisomerase mutant in S phase (Bermejo et al. 2009). Mapping of Ino80 chromatin remodeling complex binding sites over four *S. cerevisiae* chromosomes further revealed a correlation with sites of DNA replication initiation where the remodeler promotes recovery of stalled replication forks (Shimada et al. 2008). Besides, two recent studies reported that, in *S. cerevisiae*, replication origins are fragile genomic sites (Di Rienzi et al. 2009; Gordon et al. 2009). An interesting observation was that these ORI-related breakpoints in *S. cerevisiae* also tended to be associated with tRNAs and Ty elements, suggesting that they may have participated in the breakage and/or the repair (Di Rienzi et al. 2009; Gordon et al. 2009). It is still not clear why ORIs are associated with an increased risk of chromosomal breakage. One hypothesis would be that, as suggested earlier, prereplication complex binding at ORIs may act as a replication fork barrier (Tourriere and Pasero 2007), arrested forks being unstable structures prone to DSB formation. As sequences that contain Ty transposons, tRNAs, telomeres, and subtelomeres are also sensitive to spontaneous fork pausing (Labib and Hodgson 2007; Tourriere and Pasero 2007), it is

possible that the combination of an ORI together with another site of replication fork pausing increases the risk of breakage. Accordingly, our analysis of *S. cerevisiae* NUMTs revealed that numerous NUMT loci are not only associated with an ORI but are also located next to LTRs, tRNAs, or within subtelomeres (Supplemental Fig. 3). These observations are in line with a very recent genome-wide analysis revealing nonrandom distribution of gammaH2A sites in *S. cerevisiae* chromosomes with clustering at telomeres, DNA replication origins, LTRs, and tRNAs (Szilard et al. 2010). Our hypothesis that the combination of an ORI with another site of replication fork pausing may increase DNA breakage risk would also explain the high frequency of gene deletion events that we detected at the subtelomeric *ORI3001* locus in cells experiencing replication stress. Altogether, we suggest that NUMTs may have been inserted into fragile sites during evolution, and we propose that ORIs may be fragile sites in *S. pombe*, possibly because they represent replication fork pause sites.

Genome-wide mapping of human DNA replication origins has not been completed yet. However, colocalization of NUMTs and DNA replication origins may also exist in human cells as suggested by the presence of a NUMT 5' of *MYC* oncogene exon 2 in HeLa cells (Shay and Werbin 1992), an exon characterized by the presence of a DNA replication initiation site with increased activity in HeLa cells (Tao et al. 2000). In these cells, mtDNA insertion resulted in the formation of a chimeric *MYC* gene (Shay and Werbin 1992). The putative association of NUMTs and DNA replication origins in human is further supported by the fact that both human-specific NUMTs and replication origins are preferentially located within genes (Ricchetti et al. 2004; Lucas et al. 2007). Interestingly, the well-characterized human lamin B2 DNA replication origin lies 500 bp upstream of the gene encoding TIMM13 translocase of inner mitochondrial membrane (Abdurashidova et al. 2000), raising the hypothesis that insertion of entire mitochondrial genes into chromosomes and NUMT formation may have occurred similarly during evolution.

Finally, our data suggest that, in addition to a possible involvement of mechanisms like mtDNA splicing, replication pausing in mtDNA might also be at the origin of mtDNA fragmentation and, thereby, NUMT formation. Indeed, sequence analysis of *S. pombe* mtDNA fragments recovered at extrachromosomal DSBs showed that tRNA stretches may be an important source of DSB formation in the mitochondrial genome. In favor of tRNAs acting as replication pause sites in the mtDNA genome, binding of mitochondrial transcription termination factor mTERF at the tRNA^{Leu(UUR)} locus of the human mtDNA genome was previously reported to create a replication pause site (Hyvärinen et al. 2007). Accordingly, our analysis of 37 human-specific NUMTs (Ricchetti et al. 2004) revealed the presence of a breakage site at this tRNA^{Leu(UUR)} locus. In addition, our analysis also identified two additional breakage sites at O_H and O_L loci, that were recently described as both replication pause and fragile sites in the human mtDNA genome (Bailey et al. 2009).

In summary, this analysis of *S. pombe* NUMT genomic distribution highlighted the risk of DSB formation at chromosomal loci associated with DNA replication origins, providing further evidence for genomic instability induced by DNA replication defects. These findings may be relevant for tumorigenesis as several oncogenes, like RAS (Di Micco et al. 2006), are known to induce replication stress in the cell and, although cells have evolved checkpoint-dependent and -independent mechanisms that cooperate to prevent genomic instability at stalled forks (Tourrière and Pasero 2007), the integrity of these pathways can be altered.

Methods

S. pombe strains and methods

The *S. pombe* strain PN567 (*h*[−], *ura4-D18*, *leu1-32*, *ade6-704*) comes from P. Nurse's lab collection. The *rad22-YFP* strain clone 2.3. (*h*[−], *rad22-YFP* [*kanR*]) was kindly provided by B. Froget and G. Baldacci (Institut Curie, Paris, France) and crossed with PN745 (*h*⁺, *ura4-D18*, *leu1-32*, *ade6-704*) from P. Nurse's lab collection to give AD546 (*rad22-YFP* [*kanR*], *leu1-32*, *ade6-704*). The AD467 strain (strain PHP3 described in Haffter and Fox [1992]) was kindly provided by Thomas D. Fox (Cornell University, New York). PN4095 strain (*h*⁺, *SPCC1322.12cΔ::ura4*, *ura4-D18*, *leu1-32*, *ade6-M210*) was described previously (Yamaguchi et al. 2003). Strain AD552 (*h*[−], *nda3-KM311*, *cdc10-129*, *ura4-D18*, *leu1-32*, *orp1::orp1-5flag*) was kindly provided by Hisao Masukata (Osaka University, Osaka, Japan). The AD552NUMT10/11Δ strain (AD566) strain was obtained by deleting *NUMT10/11* in a two-step process. First, *NUMT10/11* was replaced by *ura4* in AD552 strain to obtain AD555 strain. Deletion was performed as described previously (Decottignies 2007). In brief, two 100 bp-long PCR fragments comprising 80 bp of genomic DNA located 5' and 3' of the genomic locus to be deleted were used as primers for PCR amplification of the deletion cassette on pREP4::*ura4* plasmid (Maundrell 1993). The *ura4* gene from AD555 was then removed by transforming yeast with PCR-amplified DNA comprising two regions of about 500 bp located exactly upstream of and downstream from the initial *NUMT10/11*. Resulting transformants were selected with 1 mg/mL 5-fluoro-orotic acid (Apollo Scientific) and checked by nucleotide sequencing. The *SPCC1884.01Δ::ura4* (AD549) and *SPBC1E8.07cΔ::ura4* (AD496) yeast strains were constructed similarly in PN567 background. The *SPBC15D4.15Δ::ura4* (AD421) yeast strain was derived from PN567 as described previously (Bähler et al. 1998). All primers are listed in Supplemental Table I. Cells were cultured in rich (YE5S) or Edinburgh minimal (EMM2) media with or without addition of 15 μM thiamin to control the *mtt1* promoter, as described in <http://www-bcf.usc.edu/~forsburg/media.html>. Mating and sporulation of cells was induced on EMM2 medium without NH₄Cl. Yeast transformations were performed as described previously (Decottignies 2005).

DNA sequencing

DNA sequencing reactions were done with the BigDye Terminator sequencing kit from PE Applied Biosystems.

Beta-galactosidase activity measurements

For transcriptional activity measurements, promoters from *SPAC1F12.02c/p23fy* and *SPBC3B9.17/isa2* genes were amplified by PCR using, respectively, primer sets 5'-TCCCCCGGGCATGTAATTCTATTCGTGC/5'-CGGGATCCATGGTGATTGCAGATGG and 5'-CGGGATCCATCACTGCAAAGTATTATAGC/5'-GGGGTACCAA CAGTAAAAGATCTTCAAG. The 1-kb PCR-amplified *SPAC1F12.02c* promoter was cloned into BamHI/SmaI-cleaved pUC18. This plasmid was then used as template to amplify pUC18 carrying NUMT-deleted *SPAC1F12.02c* promoter with primers 5'-CAGC CATCATAATGGTTT and 5'-ATTTAACAGTATGTGTATTAATC, and the PCR fragment was circularized with T4 DNA ligase. WT and mutated promoters were excised from pUC18 and cloned into BamHI/SmaI-cleaved pSEYC102 plasmid (kindly provided by M. Ghislain, Catholic University of Louvain, Louvain-la-Neuve, Belgium) to give in-frame fusions with bacterial *lacZ* gene. The 1.1-kb PCR-amplified *SPBC3B9.17/isa2* promoter was cloned into BamHI/KpnI-cleaved pUC18. This plasmid was then used as

substrate for PCR amplification of pUC18::*isa2mut1-2* (5'-TATT AATTTGACTATAAATCATC and 5'-TCGTGATACCAGTGTATGT AC) and pUC18::*isa2mut3* (5'-ACACACTCGTCAAGTTATG and 5'-AGTAGAGAATAAGGGGAATT) plasmids prior to circularization. BamHI/EcoRV fragments containing WT or mutated *isa2* promoters were then cloned into BamHI/SmaI-cleaved pSEYC102 plasmid. The resulting recombinant pSEYC102 plasmids were linearized within promoter sequences by HindIII (*SPAC1F12.02c*) or SphI (*SPBC3B9.17*) and integrated at the endogenous genomic loci of PN567 yeast cells through homologous recombination. Beta-galactosidase activity was measured on cell extracts as described previously (Decottignies et al. 1998).

Live fluorescence microscopy

Plasmids pREP3x::*cdc18* (P. Nurse's lab collection, described in Nishitani and Nurse [1995]) and pREP5::*cdc18-CFP* were introduced into AD546 cells. pREP5::*cdc18-CFP* plasmid was constructed as follows: The 1.7-kb *cdc18* ORF was PCR-amplified with primers 5'-acgcgtcgacgatgtgtgaaactcc and 5'-CGGGATCCCCCT CTTCTGTCAAAAATCATG and digested with SalI and BamHI prior to ligation into SalI/BamHI-cleaved pREP5::*CFP* plasmid described previously (Decottignies et al. 2001). For live fluorescence microscopy, cells were grown at 32°C in EMM2 medium lacking thiamin for the indicated time, washed once with water, and layered on top of low-melting agarose-coated slides as described previously (Decottignies et al. 2001). Rad22-YFP and Cdc18-CFP foci were observed in vivo with a Zeiss Axioplan microscope, using a 100×, 1.3 oil immersion lens. A mercury lamp and the Zeiss filter set No. 487910 (BP 450–490, FT 510, BP 515–565) were used for YFP excitation/emission. CFP was observed with the Chroma 41014 High Q filter set (BP 425–475, LP 480, BP 485–535). Images were captured with a Hamamatsu 3CCD chilled camera and processed with Adobe Photoshop 8.0.1 software (Adobe Systems).

Measurement of *ura4* gene mutation rate in chromosomes

Overexpression of the *cdc18* gene was obtained by introduction of the pREP3x::*cdc18* plasmid into AD549, AD496, AD555, PN4095, AD467, and AD421 strains. Cells were incubated for 19 h at 32°C with or without thiamin before plating onto EMM2 agar (50 × 10⁶ cells/plate) containing 1 mg/mL 5-fluoro-orotic acid. 5-FOA^R colonies were counted after 4–6 days and tested for uracil auxotrophy. PCR screen on ura[−] colonies was performed with either primers C1884-uraF1 and C1884-R2 or C15D4-uraF1 and C15D4-uraR2 (Supplemental Table I). If no amplification was obtained, other pairs of primers were used (Supplemental Table II).

Chromatin immunoprecipitation

AD552 and AD566 cells (*nda3-KM311 cdc10-129 orp1::orp1-5flag*) were grown at the permissive temperature of 28°C. Chromatin immunoprecipitation experiments were performed as described previously (Ghavi-Helm et al. 2008). Chromatin was sonicated for 10 min 30 sec with Bioruptor. Average chromatin fragment size was of ~500 bp. Dynabeads Pan Mouse IgG were purchased from Invitrogen and monoclonal anti-Flag M2 was from Sigma. PCRs were performed in a final volume of 25 µL containing 2.5 µL of either total chromatin from cell lysates (TOT) (dilution 1/100) or DNA associated with immunoprecipitates (IP) (dilution 1/20), 10 pmol of each primer, and 1 U of *Taq* polymerase (Takara). Primers and the number of cycles corresponding to each PCR amplification (nonsaturating conditions) are listed in Supplemental Table III. PCR products were run onto 2% agarose gels, and band intensity was quantified with Image Station 440 CF (Kodak Digital Science).

In each case, band intensity of the IP was normalized with the intensity of the TOT sample.

FACS analysis

DNA content was analyzed as described in <http://www-bcf.usc.edu/~forsburg/yeast-flow-protocol.html>, using a Becton Dickinson FACScan.

Extrachromosomal DSB repair assay

The *ura4* extrachromosomal DSB repair assay was described previously (Decottignies 2005).

BLAST

NUMTs were detected through a BLAST search against the *S. pombe* nuclear genome (<http://www.genedb.org/genedb/pombe/blast.jsp>) using the mtDNA genome (accession number X54421) as query.

Acknowledgments

We thank B. Froget, G. Baldacci, H. Masukata, T.D. Fox, and M. Ghislain for the gift of yeast strains and plasmid. We are grateful to C. De Smet and G. Tilman for critical reading of the manuscript and we thank M. Foiani, P. Pasero, and all members of the GENEPI group for helpful discussions.

References

- Abdurashidova G, Deganuto M, Klima R, Riva S, Biamonti G, Giacca M, Falaschi A. 2000. Start sites of bidirectional DNA synthesis at the human lamin B2 origin. *Science* **287**: 2023–2026.
- Bähler J, Wu JQ, Longtine MS, Shah NG, McKenzie A, Steever AB, Wach A, Philippsen P, Pringle JR. 1998. Heterologous modules for efficient and versatile PCR-based gene targeting in *Schizosaccharomyces pombe*. *Yeast* **14**: 943–951.
- Bailey LJ, Cluett TJ, Reyes A, Prolla TA, Poulton J, Leeuwenburgh C, Holt JJ. 2009. Mice expressing an error-prone DNA polymerase in mitochondria display elevated replication pausing and chromosomal breakage at fragile sites of mitochondrial DNA. *Nucleic Acids Res* **37**: 2327–2335.
- Behrens R, Hayles J, Nurse P. 2000. *S. pombe* retrotransposon Tf1 integration is targeted to 5' ends of open reading frames. *Nucleic Acids Res* **28**: 4709–4716.
- Bensasson D, Zhang D, Hartl DL, Hewitt GM. 2001. Mitochondrial pseudogenes: Evolution's misplaced witnesses. *Trends Ecol Evol* **16**: 314–321.
- Bermejo R, Capra T, Gonzalez-Huici V, Fachinetti D, Cocito A, Natoli G, Katou Y, Mori H, Kurokawa K, Shirahige K, et al. 2009. Genome-organizing factors Top2 and Hmo1 prevent chromosome fragility at sites of S phase transcription. *Cell* **138**: 870–884.
- Boeke JD, LaCrute F, Fink GR. 1984. A positive selection for mutants lacking orotidine-5'-phosphate decarboxylase activity in yeast: 5-Fluoro-orotic acid resistance. *Mol Gen Genet* **197**: 345–346.
- Borensztajn K, Chafa O, Alhenc-Gelas M, Salha S, Reghis A, Fischer AM, Tapon-Bretaudière J. 2002. Characterization of two novel splice site mutations in human factor VII gene causing severe plasma factor VII deficiency and bleeding diathesis. *Br J Haematol* **117**: 168–171.
- Bowen NJ, Jordan IK, Epstein JA, Wood V, Levin HL. 2003. Retrotransposons and their recognition of pol II promoters: A comprehensive survey of the transposable elements from the complete genome sequence of *Schizosaccharomyces pombe*. *Genome Res* **13**: 1984–1997.
- Bullerwell CE, Leigh J, Forget L, Lang BF. 2003. A comparison of three *S. pombe* mitochondrial genomes. *Nucleic Acids Res* **31**: 759–768.
- Chalker DL, Sandmeyer SB. 1992. Ty3 integrates within the region of RNA polymerase III transcription initiation. *Genes Dev* **6**: 117–128.
- Dai J, Chuang RY, Kelly TJ. 2005. DNA replication origins in the *Schizosaccharomyces pombe* genome. *Proc Natl Acad Sci* **102**: 337–342.
- Decottignies A. 2005. Capture of extranuclear DNA at *S. pombe* double-strand breaks. *Genetics* **171**: 1535–1548.
- Decottignies A. 2007. Microhomology-mediated end joining in fission yeast is repressed by pku70 and relies on genes involved in homologous recombination. *Genetics* **176**: 1403–1415.

- Decottignies A, Grant AM, Nichols JW, de Wet H, McIntosh DB, Goffeau A. 1998. ATPase and multidrug transport activities of the overexpressed yeast ABC protein Yor1p. *J Biol Chem* **273**: 12612–12622.
- Decottignies A, Zanzov P, Nurse P. 2001. In vivo localisation of *S. pombe* cyclin-dependent kinase cdc2p and cyclin B cdc13p during mitosis and meiosis. *J Cell Sci* **114**: 2627–2640.
- Di Micco R, Fumagalli M, Cicalese A, Piccinini S, Gasparini P, Luise C, Schurra C, Garre M, Nuciforo PG, Bensimon A, et al. 2006. Oncogene-induced senescence is a DNA damage response triggered by DNA hyper-replication. *Nature* **444**: 638–642.
- Di Rienzi SC, Collingwood D, Raghuraman MK, Brewer BJ. 2009. Fragile genomic sites are associated with origins of replication. *Genome Biol Evol* **3**: 350–363.
- Durkin SG, Ragland RL, Arlt MF, Mülle JG, Warren ST, Glover TW. 2008. Replication stress induces tumor-like microdeletions in FHIT/FRA3B. *Proc Natl Acad Sci* **105**: 246–251.
- Fersht N, Hermand D, Hayles J, Nurse P. 2007. Cdc18/CDC6 activates the Rad3-dependent checkpoint in the *S. pombe*. *Nucleic Acids Res* **35**: 5323–5337.
- Ghavi-Helm Y, Michaut M, Acker J, Aude JC, Thuriaux P, Werner M, Soutourina J. 2008. Genome-wide location analysis reveals a role of TFIIIS in RNA polymerase III transcription. *Genes Dev* **22**: 1934–1947.
- Giresi PG, Kim J, McDaniel RM, Iyer VR, Lieb JD. 2007. FAIRE (formaldehyde-assisted isolation of regulatory elements) isolates active regulatory elements from human chromatin. *Genome Res* **17**: 877–885.
- Gordon JL, Byrne KP, Wolfe KH. 2009. Additions, losses, and rearrangements on the evolutionary route from a reconstructed ancestor to the modern *Saccharomyces cerevisiae* genome. *PLoS Genet* **5**: e1000485. doi: 10.1371/journal.pgen.1000485.
- Greenwood E, Nishitani H, Nurse P. 1998. Cdc18p can block mitosis by two independent mechanisms. *J Cell Sci* **111**: 3101–3108.
- Haffter P, Fox TD. 1992. Nuclear mutations in the petite-negative yeast *Schizosaccharomyces pombe* allow growth of cells lacking mitochondrial DNA. *Genetics* **131**: 255–260.
- Hayashi M, Katou Y, Itoh T, Tazumi A, Yamada Y, Takahashi T, Nakagawa T, Shirahige K, Masukata H. 2007. Genome-wide localization of pre-RC sites and identification of replication origins in *S. pombe*. *EMBO J* **26**: 1327–1339.
- Hazkani-Covo E, Covo S. 2008. Numt-mediated double-strand break repair mitigates deletions during primate genome evolution. *PLoS Genet* **4**: e1000237. doi: 10.1371/journal.pgen.1000237.
- Hazkani-Covo E, Sorek R, Graur D. 2003. Evolutionary dynamics of large numts in the human genome: Rarity of independent insertions and abundance of post-insertion duplications. *J Mol Evol* **56**: 169–174.
- Heichinger C, Penkett CJ, Bähler J, Nurse P. 2006. Genome-wide characterization of *S. pombe* DNA replication origins. *EMBO J* **25**: 5171–5179.
- Hyman BC, Cramer JH, Rownd RH. 1983. The mitochondrial genome of *Saccharomyces cerevisiae* contains numerous, densely spaced autonomously replicating sequences. *Gene* **26**: 223–230.
- Hyvärinen AK, Pohjoismäki JL, Reyes A, Wanrooij S, Yasukawa T, Karhunen PJ, Spelbrink JN, Holt JJ, Jacobs HT. 2007. The mitochondrial transcription termination factor mTERF modulates replication pausing in human mitochondrial DNA. *Nucleic Acids Res* **35**: 6458–6474.
- Ide S, Watanabe K, Watanabe H, Shirahige K, Kobayashi T, Maki H. 2007. Abnormality in initiation program of DNA replication is monitored by the highly repetitive rRNA gene array on chromosome XII in budding yeast. *Mol Cell Biol* **27**: 568–578.
- Ivessa AS, Lenzmeier BA, Bessler JB, Goudsouzian LK, Schnakenberg SL, Zakian VA. 2003. The *Saccharomyces cerevisiae* helicase Rrm3p facilitates replication past nonhistone protein-DNA complexes. *Mol Cell* **12**: 1525–1536.
- Kim D-U, Hayles J, Kim D, Wood V, Park H-O, Won M, Yoo H-S, Duhig T, Nam M, Palmer G, et al. 2010. Analysis of a genome-wide set of gene deletions in the fission yeast *Schizosaccharomyces pombe*. *Nat Biotechnol* **28**: 617–623.
- Labib K, Hodgson B. 2007. Replication fork barriers: Pausing for a break or stalling for time? *EMBO Rep* **8**: 346–353.
- Leister D. 2005. Origin, evolution and genetic effects of nuclear insertions of organelle DNA. *Trends Genet* **21**: 655–663.
- Lucas I, Palakodeti A, Jiang Y, Young DJ, Jiang N, Fernald AA, Le Beau MM. 2007. High-throughput mapping of origins of replication in human cells. *EMBO Rep* **8**: 770–777.
- Maudrell K. 1993. Thiamine-repressible expression vectors pREP and pRIP for fission yeast. *Gene* **123**: 127–130.
- Meister P, Poidevin M, Francesconi S, Tratner I, Zanzov P, Baldacci G. 2003. Nuclear factories for signalling and repairing DNA double strand breaks in living *S. pombe*. *Nucleic Acids Res* **31**: 5064–5073.
- Meister P, Taddei A, Vernis L, Poidevin M, Gasser SM, Baldacci G. 2005. Temporal separation of replication and recombination requires the intra-S checkpoint. *J Cell Biol* **168**: 537–544.
- Nishitani H, Nurse P. 1995. p53cdc18 plays a major role controlling the initiation of DNA replication in *S. pombe*. *Cell* **83**: 397–405.
- Nugent JM, Palmer JD. 1991. RNA-mediated transfer of the gene *coxII* from the mitochondrion to the nucleus during flowering plant evolution. *Cell* **66**: 473–481.
- Pamilo P, Viljakainen L, Vihavainen A. 2007. Exceptionally high density of NUMTs in the honeybee genome. *Mol Biol Evol* **24**: 1340–1346.
- Ricchetti M, Fairhead C, Dujon B. 1999. Mitochondrial DNA repairs double-strand breaks in yeast chromosomes. *Nature* **402**: 96–100.
- Ricchetti M, Tekaia F, Dujon B. 2004. Continued colonization of the human genome by mitochondrial DNA. *PLoS Biol* **2**: e273. doi: 10.1371/journal.pbio.0020273.
- Richly E, Leister D. 2004. NUMTs in sequenced eukaryotic genomes. *Mol Biol Evol* **21**: 1081–1084.
- Sacerdot C, Casaregola S, Lafontaine I, Tekaia F, Dujon B, Ozier-Kalogeropoulos O. 2008. Promiscuous DNA in the nuclear genomes of hemiascomycetous yeasts. *FEM Yeast Res* **8**: 846–857.
- Schafer B, Merlos-Lange AM, Anderl C, Welser F, Zimmer M, Wolf K. 1991. The mitochondrial genome of *S. pombe*: Inability of all introns to splice autocatalytically, and construction and characterization of an intronless genome. *Mol Gen Genet* **225**: 158–167.
- Segurado M, de Luis A, Antequera F. 2003. Genome-wide distribution of DNA replication origins at A+T-rich islands in *Schizosaccharomyces pombe*. *EMBO Rep* **4**: 1048–1053.
- Shay JW, Werbin H. 1992. New evidence for the insertion of mitochondrial DNA into the human genome: Significance for cancer and aging. *Mutat Res* **275**: 227–235.
- Shimada K, Oma Y, Schleker T, Kugou K, Ohta K, Harata M, Gasser SM. 2008. Ino80 chromatin remodeling complex promotes recovery of stalled replication forks. *Curr Biol* **18**: 566–575.
- Singleton TL, Levin HL. 2002. A long terminal repeat retrotransposon of *S. pombe* has strong preferences for specific sites of insertion. *Eukaryot Cell* **1**: 44–55.
- Szilar RK, Jacques P-E, Laramée L, Cheng B, Galicia S, Bataille AR, Yeung MT, Mendez M, Bergeron M, Robert F, et al. 2010. Systematic identification of fragile sites via genome-wide location analysis of γ -H2AX. *Nat Struct Mol Biol* **17**: 299–305.
- Tanny RE, MacAlpine DM, Blitza HG, Bell SP. 2006. Genome-wide analysis of re-replication reveals inhibitory controls that target multiple stages of replication initiation. *Mol Biol Cell* **17**: 2415–2423.
- Tao L, Dong X, Leffak M, Zannis-Hadjopoulos M, Price G. 2000. Major DNA replication initiation sites in the c-myc locus in human cells. *J Cell Biochem* **78**: 442–457.
- Tourmen Y, Baris O, Dessen P, Jacques C, Malthiery Y, Reynier P. 2002. Structure and chromosomal distribution of human mitochondrial pseudogenes. *Genomics* **80**: 71–77.
- Tourriere H, Pasero P. 2007. Maintenance of fork integrity at damaged DNA and natural pause sites. *DNA Repair* **6**: 900–913.
- Triant DA, DeWoody JA. 2007. Extensive mitochondrial DNA transfer in a rapidly evolving rodent has been mediated by independent insertion events and by duplications. *Gene* **401**: 61–70.
- Turner C, Killoran C, Thomas NS, Rosenberg M, Chuzhanova NA, Johnston J, Kemel Y, Cooper DN, Biesecker LG. 2003. Human genetic disease caused by de novo mitochondrial-nuclear DNA transfer. *Hum Genet* **112**: 303–309.
- Willett-Brozick JE, Savul SA, Richey LE, Baysal BE. 2001. Germ line insertion of mtDNA at the breakpoint junction of a reciprocal constitutional translocation. *Hum Genet* **109**: 216–223.
- Wood V, Gwilliam R, Rajandream MA, Lyne M, Lyne R, Stewart A, Sgouros J, Peat N, Hayles J, Baker S, et al. 2002. The genome sequence of *Schizosaccharomyces pombe*. *Nature* **415**: 871–880.
- Wuarin J, Buck V, Nurse P, Millar JB. 2002. Stable association of mitotic cyclin B/Cdc2 to replication origins prevents endoreduplication. *Cell* **111**: 419–431.
- Yamaguchi S, Decottignies A, Nurse P. 2003. Function of Cdc2p-dependent Bub1p phosphorylation and Bub1p kinase activity in the mitotic and meiotic spindle checkpoint. *EMBO J* **22**: 1075–1087.
- Yieh L, Kassavetis G, Geiduschek EP, Sandmeyer SB. 2000. The Brf and TATA-binding protein subunits of the RNA polymerase III transcription factor IIIB mediate position-specific integration of the gypsy-like element, Ty3. *J Biol Chem* **275**: 29800–29807.
- Yu X, Gabriel A. 1999. Patching broken chromosomes with extranuclear cellular DNA. *Mol Cell* **4**: 873–881.

Received December 23, 2009; accepted in revised form June 30, 2010.